STOCHASTIC MULTI CRITERIA DECISION ANALYTICS AND ARTIFICIAL INTELLIGENCE IN CONTINUOUS AUTOMATED TRADING FOR WEALTH MAXIMIZATION

Gordon H. Dash Jr.*[1], Nina Kajiji[2], and John Forman[3]

[1]College of Business Administration, University of Rhode Island
[2] The NKD Group, Inc. and, Computer Science and Statistics, University of Rhode Island
[3]Thomson Reuters, Boston, MA

Abstract: Recent technological and regulatory advances have coalesced to usher in an era where both automated and algorithmic trading routinely characterize a wealth maximization process managed by the continuous trading of equity securities. Under this approach to wealth maximization, automated trading focuses on the process determining directional trades for individual securities based upon the receipt and interpretation of new data. This paper presents a stochastic price formation algorithm that implements a cognitive decision making system modeled by twin radial basis function artificial neural networks to produce a high frequency automated trading system for individual equity securities listed on U.S. exchanges. The overall effectiveness and efficiency of the automated trading system is calibrated by estimating non-parametric quasi elasticity coefficients for individual firm fundamental characteristics. We find that automation driven by cognitive science can effectively auto-trade securities and produce changes to individual wealth that equals or exceeds the performance generated by a simple buy-and-hold strategy. We also identify four fundamental firm factors that explain the ability of the automated trading algorithm to produce a measured level of percent-positive trades.

Keywords: Stochastic Models, Artificial Neural Networks; Cognitive Science; Automated Trading

1 Introduction

How wonderful it would be to perfectly predict stock prices. In such a world the individual seeking to maximize wealth would only be required to choose an appropriate risk level. Owing to a myriad of different factors it is well known that consistent prediction of future stock price levels is not possible. As a consequence, researchers have turned to quantitative modeling as a means by which to describe a comprehensive view of individual stock price behavior. But, the evolving complexities of an ever expanding 24/7 global market has contributed to the need for decision-makers to understand and include behavioral consequences into the formulation of mathematical models of market price behavior.

The purpose of this paper is to integrate cognitive science, or behavioral decision theory (BDT), with an artificial intelligence (AI) based multiple criteria decision aiding model (MCDA) to explain how individual’s trade equity securities in their attempt to maximize wealth while mitigating perceived risks to the wealth accumulating process. To achieve this objective, we present a BDT information system that is engaged in the continuous receipt and evaluation of new information. More directly, the system presented in this paper implements an MCDA artificial intelligence-based automated trading system that includes a nested decision subsystem.

Automated trading of equity securities is a complex system that continuously integrates various cognitive and algorithmic processes that receive and interpret new market data in order to update the individual’s wealth share decision under the assumption of wealth maximization. The efficient automated trading system depends on its ability to use a set of mathematical models and market rules to automatically identify optimal market decisions. These decisions include choices of whether to open, sell short, hold or close a specific equity position within the framework of the continuous trading evaluation. To evaluate one such automated trading information system, this paper introduces the **Wi**nORS **N**euroeconomic **K**nowledge-based market trading **S**ystem (WINKS). The WINKS system is a dual radial basis function (RBF) artificial neural network (ANN) stock price (return) forecasting system that combines end-of-day (EOD) forecasts with intra-day forecasts to determine the quantities for intra-day or EOD transaction decisions.

The paper proceeds as follows. Section 2 introduces the stochastic integral model as a theory of trading an equity security for profitability. Section 3 presents the relevant algorithms.[z] Section 4 provides

* Corresponding author: Gordon H. Dash, Jr. at GHDASH@URI.EDU

a snapshot of how firm fundamental variables contribute to the production of profitable trades within the context of the WINKS high-frequency trading system. Section 5 provides a summary and conclusion.

2 The Stochastic Integral as a Model of Equity Trading Profits

Stochastic calculus has rapidly become the language of financial modeling (see Brock et. al. (1992), for a discussion). In this section we present a characterization of the Shreve (2004) framework for use of the stochastic integral to characterize uncertain stock trading. The methodology is developed in a manner that is designed to support the WINKS framework.

Consider X_t to be the random variable of a stock's market price at time t. As in prior research, we assume that the price process X follows a geometric Brownian motion with a constant drift and volatility (For a in-depth development see Tsay (xxx)). Next we define a trading strategy θ that determines the quantity $\theta_t(\omega)$ of each security held in each state $\omega \in \Omega$ and at each time t. The trading system envisioned by this approach assumes a market that is not characterized by the no-risk unlimited profit arbitrage effects of trading on advanced knowledge. That is, θ is adapted and corresponds to the necessary restriction that the trading strategy can only make use of the available information at any time t. This prevents the possibility of unlimited gains through high frequency trading or flash-crash trading (For description of May 6th, 2010 flash crash see: http://en.wikipedia.org/wiki/May_6,_2010_flash_crash) . Note, the condition that θ is adapted implies that the stochastic integral will not diverge when calculated as a limit of Riemann sums. Hence, given a price process X and a trading strategy θ that satisfies the no arbitrage conditions, the total financial gain $\int_s^t \theta_u dX_u$ between any times $s, t \geq 0$ is defined by Ito's stochastic integral. Since this is a continuous-time stochastic process, it is assumed that there is an underlying filtered probability space $(\Omega, V, (V_t)_{t\geq 0}, P)$. The increasing sequence of σ-algebra of V,$\{V_t: t \in [0, \infty]\}$, determines the relevant timing of information. That is, V_t represents the information available up until time t, and is loosely viewed as the set of events whose outcomes are certain to be revealed to investors as true or false by, or at, time t. Finally, the trading strategy θ is adapted if $\theta_t(\omega)$ is V_t measurable.

2.1 Buy and Hold

The buy-hold (BH) strategy is a short-horizon element of the WINKS trading strategy captured by θ. Under the BH strategy an investor initiates a position immediately after some stopping time T and closes it at some later stopping time U. For a position size $\theta_t(\omega)$ that is V_t measurable, the trading strategy θ is defined by $\theta_t = 1_{(T<t\leq U)}\theta_t(\omega)$. By definition, the gain from the BH trade strategy is the position size multiplied by the interim price change, or $\int_0^U \theta_t dX_t = \theta_t(X_U - X_T)$.

2.2 The N-Dimensional Trading Strategy

A typical financial model allows for n different securities, with price process $X_1, \ldots, X_n$. The investor can choose an associated N-dimensional trading strategy $\theta = (\theta_1, \ldots, \theta_n)$ for which the total gain from the equity trading process is: $\int \theta_t dX_t \equiv \sum_{i=1}^n \int \theta_{it}\, dX_{it}$. The technical restrictions that define the stochastic integrals can be augmented for the allowable set θ to include budget limits, credit constraints, short-sales restrictions or various other managerially imposed investment constraints.

3 The Automated Trading System and the Production of Profitability

The prediction and mapping capabilities of ANNs in general, and the RBF topology specifically, has resulted in an extraordinary amount of interest in applying various ANN algorithmic topologies to stock market prediction and forecast behavior (for example, see Refenes, P., et.al. (1996)). The usefulness of RBF ANNs continues to be exploited in complex financial optimization and mapping studies as this particular ANN topology does not require a parametric system model and tends to be relatively insensitive to chaotic appearing data patterns (Dash and Kajiji (2008)) . The objectives for this section of the paper are twofold. The first objective is to specify the K4-RBF ANN that provides both EOD and 20-minute ahead forecasts for individual securities is described. The second objective for this section is to define the WINKS decision algorithm system in pseudo detail. Like many auto trading systems WINKS relies upon the prediction of price at time period $t+1$ given a price observation at time t with a known information set, θ. Within WINKS the functional form of all prediction models (both the EOD and 20-minute ahead forecasts) is as follows:

$$x_{t+1} = f_\theta\left(x_t, P_{1,t}, P_{2,t}, \dots, P_{k,t} \mid \theta\right) \quad (1)$$

Where x_t is the price of the target security at time *t* and $P_i, \dots, P_k$ captures the set of *k* exogenous predictor variables.

3.1 High-Frequency Neuroeconomics for Stock Price Forecasting

The automated trading algorithm employed by WINKS is based upon the K4-RBF ANN (2001). WINKS employs this ANN to produce an analytic approximation for the next period stock return by mapping the noisy exogenous data stream which may be described as follows,

$$\{[x(k), y_i : [\mathbb{R}^n, \mathbb{R}]\}_{k=1}^{m} \quad (2)$$

where *x(k)* is the input vector for predictor *k*, y_i is the output for stock *i*, and *n* is the dimension of the input space, and *m* is the number of basis functions. The data is drawn from the noisy set:

$$\{[y_i = f[x(k) + \epsilon]\}_{k=1}^{m} \quad (3)$$

As shown in figure 1, the RBF ANN topology is defined by three layers: the input layer, the hidden layer (linear layer) and the output layer. The input layer has no particular calculating power; its primary function is to distribute the information to the hidden layer of the RBF network. The hidden, or middle, layer embraces computing units, or hidden nodes. Each hidden node is defined by a center. The center, *c(k)*, is a parameter vector of the same dimension as the input data vector, *x(k)*, and calculates the Euclidean distance between the center and the network input vector *x* defined by $\|x(k) - c(k)\|$. The results are passed through a nonlinear activation function, $\emptyset(k)$, to produce output from the hidden nodes. The Gaussian basis function shown in eq. 4 is a widely used approach to establish the activation function.

$$\emptyset(k) = exp\left(\frac{\|x(k) - c(k)\|^2}{\sigma_j^2}\right), \quad j = 1 \dots m \quad (4)$$

where σ_j is a positive scalar and is referred to as the width of the center. The output layer is a linear combiner with the *i*th output of the network model being a weighted sum of the hidden nodes:

$$\hat{y}_i = \sum_{j=1}^{m} \emptyset(k)\, w_j, \quad i = 1 \dots p \quad (5)$$

where *p* is the number of outputs (generally *p* = 1), *w* represents output layer weights, and $\hat{y}$ is the network output to estimate the target *y*. For generalizations see Haykin (1994).

The Kajiji (2001) extension to the traditional RBF ANN specification introduced multiple objectives within a Bayesian RBF ANN framework. By adding a weight penalty term to the SSE optimization objective, the modified SSE is restated as the following cost function:

$$C = \sum_{i=1}^{p} (\hat{y}_i - f(x_i))^2 + \sum_{j=1}^{m} v_j w_j^2 \quad (6)$$

where: v_j are regularization parameters or weight decay parameters. Under this specification the function to be minimized is stated as:

$$C = \frac{argmin}{v} \left(\zeta \sum_{i=1}^{p} (y_i - f(x_i | \overline{v}))^2 + \sum_{j=1}^{m} v_j w_j^2\right) \quad (7)$$

In early implementations of the RBF topology, iterative techniques were commonly employed to compute the weight decay vector $\overline{v}$. The extensions embraced by the K4-RBF ANN allow the dual-objective, multiple criteria decision analytic (MCDA) algorithm to directly attack the twin evils that deter efficient ANN modeling: the “curse” of dimensionality (multicollinearity or over-parameterization) and inflated residual sum of squares (inefficient weight decay). The benefit to WINKS is straightforward. First, the excellent mapping capabilities of the RBF topology are applied to the generalized- volatility forecasting problem inherent in all price forecasting systems. Second, the algorithmic speed generated by K4-RBF ANN enhancements permit the computational algorithm to operate in a high-frequency forecasting environment for thousands of securities.

3.2 The WINKS Automated Trading Algorithm

The inherited computational commonality among decision theory, the cognitive sciences, artificial intelligence (AI) and operations research has been well established in the literature (see, Zimmerman (1991) for a review). As shown in the flow-chart (figure 2), the WINKS automated trading algorithm is an N-dimensional cognitive decision making engine that generates probability judgment(s) and an information search in response to a diagnosis of human hypotheses when presented with new trading-oriented data. The process is based on a high-frequency (daily) and ultra-high frequency (intraday) system that executes one RBF ANN to produce a 20-minute BH investment decision which is compared to the daily EOD investment horizon forecast that is generated by a separate RBF ANN model over the course of a one-day investment horizon.

3.3 WINKS Performance

The complete results of comparing the buy-and-hold strategy (BH) to the continuous process of automated trading (e.g., WINKS) are left for a longer version of this work. Table 1 presents an abbreviated comparison of the results produced by BH and WINKS. We note that trading efficiency is defined as: ((WINKS trade profit – BH trade profit) / Initial $ Investment). Positive and negative trades are indicated by +ve / -ve, respectively.

Table 1: Comparison of Buy-Hold v/s Trading: June 01, 2009 to March 19, 2010

Equity Ticker	BH Profit	BH Annualized RoR	WINKS Trading Profit	WINKS Annualized RoR	WINKS +ve Trades	WINKS -ve Trades	WINKS Trading Efficiency
TVL	$1925	140%	$2402	172%	45	14	48%
WLL	$700	55%	$912	71%	40	17	22%
PLCE	$229	18%	$850	66%	26	16	63%
CMED	$(339)	-29%	$806	62%	41	21	116%
ILMN	$24	2%	$632	49%	26	16	61%

4 Firm Fundamental Factors that Produce Efficient Positive Trades

In this section of the research we investigate the relationship between firm fundamental characteristics and percent positive trades. Over the past 40 years, a large body of research has evolved to explore specific characteristics (beyond market beta) that are known to have a significant explanatory effect on average market returns. For the purpose of estimating the joint roles contributed by market beta and firm specific variables the extant literature is uniquely reliant on the classic study of Fama and French (1993). The result of these studies has established a role for a size factor, and a book-to-market factor. Carhart (1997) extended cross-section modeling to include a fourth factor, a momentum effect, to encapsulate market risk. We specifically note recent applications of the Carhart model to identify international market factors (see, Lam, Li and So, (2009)). Unlike financial econometric studies that focus on portfolio theoretic implications, this research seeks to identify individual firm-level factors that efficiently produce profitable trading opportunities when merged with BDT modeling. To this end, we invoke the artificial intelligence approach encapsulated in a RBF neural network to extract nonparametric quasi-elasticity estimates of the relative contribution provided by firm fundamental variables to the process of producing trading profitability. However, the experiment provided below is guided by the portfolio-theoretic findings as reviewed.

4.1 Elasticity

The derivative is commonly used to compute the percentage rate of change of a function. It is well known that for the function $y = f(x)$, the unit free average elasticity, $E_{y/x}$, of the variable y with respect to the variable x is given by the ratio: $E_{y/x} = \frac{\%\Delta y}{\%\Delta x}$. For estimation purposes we implement

$$\ln(y) = \ln A + \sum_{i=1}^{n} \beta_i \ln(x_i)\,,\ \beta_i > 0 \tag{8}$$

In a manner that is consistent with the properties stated above, all inputs are interchangeable and each input must be used in strictly positive amounts to obtain a positive output.

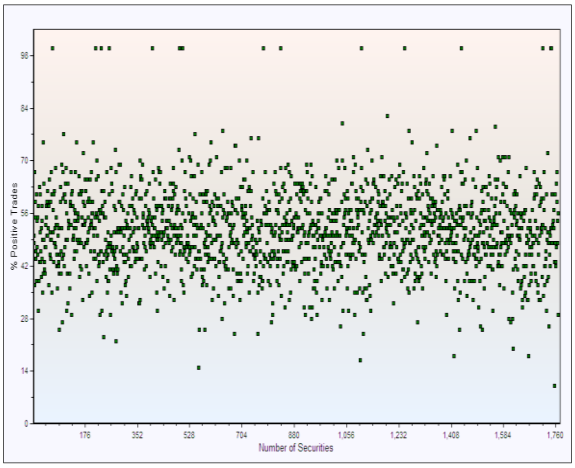
4.2 Nonparametric Estimation of Fundamental Trading Predictors

For the primary analysis, we begin with a time series of 20-minute price observations for 2,225 securities from all U.S. trading dates from June 01, 2009 to March 19, 2010 inclusive. To eliminate structural biases in the econometric analysis of trading performance we eliminate all non-equity stocks and those equities that do not have Yahoo! supported fundamental characteristics. Upon completing data reduction procedures a full-sample size of 1,765 is produced. For efficient cross-sectional modeling we sample from within the full content population. The data sampling is guided by the use of the target variable of the study – *percent positive trades* (PPT):

$$\%\ \text{Positive Trades} = \text{Round}\left(\frac{\#\ \text{positive trades}}{\text{total}\ \#\ \text{trades}}\ \text{x}\ 100\right) \tag{9}$$

The scatter plot of percent positive trades is shown in figure 3. Immediately obvious is the implied lower/upper bands at approximately 30% and 80%, respectively.

Figure 3: Percent Positive Trades by Security on March 19, 2010



4.2.1 Neural Network Modeling and Cognitive Decision Theory

As reported by Patterson (1996) ANNs are complex computational algorithms that trace their roots to modeling low-level structures of the human brain. Recent research authored by Lewicki, Hill, and Czyzewska (1992) in the cognitive science of non-conscious information processing has demonstrated the efficiency of the human brain to learn simple input-output co-variations from extremely complex stimuli. To prepare for the K4-RBF ANN, we create a training sample from the cross-sectional observations. The full sample of 1,765 securities is separated into four strata based on the value of the target variable. After strata creation, we compute the standard deviation of the % positive trades variable and then compute the sample size n at the 95% confidence interval with a minimum error rate (E) of 0.75. This process results in the selection of 793 securities for the training set.

4.2.2 Nonparametric Quasi-Elasticity Estimation

To estimate a nonparametric production theoretic model, we identify the functional form with one target variable and the set of predictor variables: $p_i = f(P_1, P_2, P_3, P_4)$, where p_i is the PPT for each *ith* security; P_1 is the individual security *Vasicek adjusted beta*, P_2 is *book/price ratio* computed from annualized book value and the last observed trade price, P_3 is the security's *current market capitalization*; and, P_4 is the Yahoo! Provided: *% change from 50 day MA (moving average)*.

4.3 Firm Fundamentals for Systemic Prediction of Profitable Trades

The overall statistical efficiency of several alternative ANN solutions is presented in table 2. The solutions are differentiated by the choice of data transformation functions implemented prior to training and validation. *Model Norm:2* is deemed to be the most efficient solution .

Table 2: K4 Analysis Using Softmax Transfer Function.

Description	*Norm:2*[a]	*Norm:1*	*STD:1*
Validation Error	7.19E-04	9.54E-03	1.11E-04
Fitness Error	8.49E-04	1.07E-02	1.02E-04
R-Square	99.12%	90.42%	99.89%
AIC	-12472	-7994	-16197
Schwarz	-12450	-7972	-16175

a: Selected data transformation

Table 3 provides an interpretation of the Norm:2 model quasi-elasticity estimate (RBF weight).

Table 3: Weights of Comparative Models

Variables / Interpretation for Norm:2	*Norm:2*	*Norm:1*	*STD:1*
Ln(Abs(Vasicek's Beta)) *A 10% increase in V_i will result in just under a 1.6% decrease in Percent positive Trades (PPT)*	-0.157	-0.060	-1.993
Ln(Abs(Book To Last Trade Price)) *A 10% increase in Vi will result in a 5.5% decrease in PPT*	-0.546	1.478	-2.592
Ln(Market Capitalization / 10000) *A 10% increase in V_i will result in a 13.9% increase in PPT*	1.387	0.544	1.691
Ln(% Change From 50 Day MA) *A 10% increase in V_i will result in a 2% decrease in PPT*	-0.200	-0.948	-0.233

5 Summary and Conclusion

First, this paper provides a synthesis of stochastic equity price behavior and the cognitive science of trading by the use of K4-RBF modeled ANN. WINKS is an MCDA trading algorithm that integrates the AI properties of two unique, but coordinated, high-frequency RBF ANNs. The results of executing WINKS over the test period produced transaction cost adjusted trading profits that exceeded those generated by the simple buy-and-hold strategy. Second, WINKS performance was modeled using firm fundamentals to uncover the factors that generate PPT. It was argued that the effective use of WINKS across global markets is potentially enhanced by the pre-selection of stocks based on the estimated quasi-elasticity estimates (RBF weights). Alternate performance comparisons are left for future research.

References

Brock, W., J. Lakonishok, et al. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*, 47(5), 1731-1764.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1), 57-82.

Dash Jr., G. H. and N. Kajiji (2008). Engineering a Generalized Neural Network Mapping of Volatility Spillovers in European Government Bond Markets. *Handbook of Financial Engineering*. C. Zopounidis, M. Doumpos and P. M. Pardalos, Springer. 18.

Fama, E. and K. French (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33(1), 3-56.

Haykin, S. (1994). *Neural Networks a Comprehensive Foundation*. New York, Macmillan.

Kajiji, N. (2001). Adaptation of Alternative Closed Form Regularization Parameters with Prior Information to the Radial Basis Function Neural Network for High Frequency Financial Time Series. *Applied Mathematics*. Kingston, University of Rhode Island.

Lam, K., F. K. Li, et al. (2009). "On the Validitity of the Augmented Fama-French Four Factor Model." from SSRN: http://ssrn.com/abstract=1343781.

Lewicki, P., T. Hill, et al. (1992). Nonconscious Acquistion of Information. *American Psychologist*, 47, 796-801.

Patterson, D. (1996). *Artificial Neural Networks*. Singapore, Prentice Hall.

Refenes, A.-P. N., Y. Abu-Mustafa, et al., Eds. (1996). *Neural Networks in Financial Engineering*. Singapore, World Scientific.

Shreve, S. E. (2004). *Stochastic Calculus for Finance*. New York, Springer.

Zimmermann, H. J. (1991). Cognitive Sciences, Decision Technology, and Fuzzy Sets. *Information Sciences*, 57-58, 287-295

[z]For all cited tables, graphs, and references see: http://www.nkd-group.com/research/EuroOpt2010/tablegraph.pdf.