Selecting Stocks Using A Genetic Algorithm: A Case of Real Estate Investment Trusts (REITs)

Kaveepong Lertwachara

ABSTRACT

Many investors invest according to analysts' recommendations, but some studies detect sizable abnormal returns associated with those recommendations. A tool that imitates the mechanism of the recommendation process should enable individual investors to formulate a profitable investment strategy and to select stocks that analysts have not covered. This paper presents a genetic algorithm (GA) method for identifying attractive stocks. The GA method has shown promising performance as a tool with which investors can develop a profitable investment strategy. GA is able to pick stocks with a "buy" rating at 62.50%, compared to 36.45% for stocks picked at random.

Key words: analysts' recommendation, genetic algorithm, cluster analysis, real estate investment trusts

INTRODUCTION

Many investors rely on analysts' recommendations. Givoly and Lakonishok (1984) even praised stock recommendations as the most notable output of financial analysts. Francis and Soffer (1997) indicated that individual investors trade stocks according to analysts' recommendations, while Barber and Loeffler (1993) and Hirchey *et al.* (2000a, 2000b) reported significantly abnormal returns and volumes following "buy" announcements from analysts. Hirchey *et al.* (2000a) added that the impact of analysts' recommendations on individual investors is even more immediate and significant with the advent of Internet technology, which makes such information readily available and accessible to anyone with Internet access at little or no cost. A tool that imitates the mechanism of the recommendation process should enable individual investors to formulate an optimal investment strategy before the analysts' recommendation is publicized, giving the investor a significant advantage over other investors. Even with the potential of such a tool, few studies have attempted to identify one.

Although the process of formulating a stock recommendation is sophisticated (Breton and Taffler, 2001), only linear models have been examined in previous studies. Through multiple regression and correlation analysis, Mear and Firth (1987) investigated what information financial analysts incorporate into the decision-making process. Breton and Taffler (2001) also investigated what information financial analysts use to formulate stock recommendations and found that qualitative information such as management and strategy, rather than financial information, is most relevant. However, they cautioned that the qualitative information extracted from analysts' reports is subjective and can be deviated from by

College of Innovative Education, Thammasat University, Bangkok 10200, Thailand.
e-mail: kaveepong@gmail.com

anyone experienced in the field. The lack of expertise in the field makes the decision process even more complicated (Mear and Firth, 1987). Since the relationship between financial information and the decision is sophisticated, a linear model may not fully capture it. In addition, a linear model requires certain probability distributions, which are normally unrealistic in finance problems.

In this study, I examined two heuristic approaches, the neural network (NN) and the genetic algorithm (GA), in order to address (1) the sophisticated relationship between financial information and stock recommendation, and (2) violations of probability distribution which usually occur in finance problems. The NN has been identified as a tool for bankruptcy prediction (Tam and Kiang, 1992; Altman *et al.* 1994). NNs "learn" from the data set and adjust the estimated parameters as observations pass through them. The success of NNs occurs because (1) a non-linear function produced by NNs is suitable to a multi-modal data set, (2) NNs are capable of adaptively adjusting the model according to change in the real-world data, and (3) NNs do not assume any probability distribution and do not require any form of input and output (Tam and Kiang, 1992).

However, Tam (1991) pointed out that NNs consume enormous computational time and that, while there is no formal way to set up an optimal configuration for them, their performance varies based on the configurations. In addition, NNs can be easily stuck in a local optimum. On the other hand, the GA, a parallel search technique, is known for finding a global optimum (Varetto, 1998). GAs emulate genetic production and search for solutions to optimization problems (Holland, 1992a). They have been found to be better than traditional techniques (Varetto, 1998; Chiou and Lan, 2001) and better than NNs (Sexton and Dorsey, 2000).

In this study, I configured NNs and GAs to replicate analysts' process of making stock recommendations. To avoid subjectivity, the NN and the GA in this study analyzed only quantitative variables from financial reports. I also compared the NN and the GA to the k-means algorithm (KM), a traditional clustering technique used in selecting stocks for Real Estate Investment Trusts (REITs) on the basis of correct classification rate, using the actual analysts' recommendation as a benchmark. I analyzed REITs because: (a) the industry is regulated and unique in the sense that it is required to distribute at least 90% of taxable income to shareholders, (b) the industry performs better on average than companies in other industries during economic downturns, and (c) REITs fluctuate less than other non-regulated industries since the supply side of this industry is limited. In addition, I created homogeneity within the dataset to avoid an industry effect that might influence the analysts' opinion, which also embedded another level of difficulty to the clustering task of this paper. Finally, I analyzed the probability that the KM and the GA can correctly pick a "buy" stock.

Background

K-Means algorithms (KM)

The K-means algorithm (KM) classifies observations into groups based on some similarity. KMs begin with a pre-determined starting centroid (or seed) for each cluster. Observations are then grouped on the basis of their distance from the seeds. In some KMs, each observation is placed into the cluster with the nearest centroid. After all observations are assigned to clusters, the centroids are recalculated. In other KMs, the centroids are recalculated after each observation is assigned to a cluster.

Research on KMs appears in a wide range of disciplines. For example, Slater and Olson (2001) performed KMs on firms' marketing strategies and found four marketing strategies: aggressive marketers, mass marketers, marketing minimizers, and value marketers. They also find that firms perform well if specific marketing strategies are matched with specific business strategies. Barrett and Wilkinson (1985) applied a KM to Australian manufacturing firms to eliminate problems in exporting their

products and services.

Unfortunately, the KM is sensitive to the sequence of the data, the specified number of clusters, and initial seeds (Punj and Stewart, 1983), so it requires a pre-specified number of clusters and starting points based on the desired number of clusters. These requirements usually cause the two most common problems in cluster analysis: error in determining the numbers of clusters, and error in assigning observations to clusters. The well established practice is to use another traditional technique, prior knowledge, and/or an associated theory to estimate the number of clusters, then to select the initial seeds randomly.

Neural network (NN)

The NN mimics a mechanism of the brain. A single NN consists of at least two layers: input and output. Any layer between the input and output layers is called a hidden layer, and as many hidden layers as desired can be added between the input and output layers. Each layer consists of a number of processing units, which are called neurons or nodes and are basically computing devices. Each neuron in a hidden layer receives inputs from other neurons in the previous layer—input, output or hidden—and sends outputs to neurons in the next layer. Each signal, whether input or output, is multiplied by a weight before it is passed on to the next layer; then each neuron applies an activation function to the signals from the neurons in the previous layer. The NN "learns" a cluster structure from training data set by adjusting weights for each node in the network to fit the data on the basis of either external or internal measurements.

NNs have been compared to many traditional clustering algorithms, although results on how well they stack up are mixed. In many studies, NNs were found to be superior to traditional clustering procedures. For example, Chen *et al.* (1995) compared NNs to seven traditional clustering algorithms, randomly generating and varying their data sets used in the comparison on four factors: number of clusters, number of variables, relative dispersion within the clusters, and number of observations. The results indicated that NNs are superior to traditional clustering algorithms, especially when the level of dispersion increases.

Hruschka and Natter (1999) compared the NN to the KM for cluster-based market segmentation and analyzed the usages of household cleaners brands in different situations. In their study, the NN outperformed the KM based on the Davies-Bouldin index (Davies and Bouldin, 1979). The NN suggested a two-cluster structure while the KM failed to recover the true cluster structure on the basis of an external criterion.

Tam and Kiang (1992) studied the success of NNs in finance problems, especially bankruptcy prediction, by applying a NN to bankruptcy predictions problems. Their results suggested that a NN should be preferable to linear discriminant analysis because (1) a non-linear function produced by a NN is suitable to a multi-modal data set, (2) a NN is capable of adaptively adjusting the model accordingly to a change in the real-world data, and (3) NNs do not assume any probability distribution and do not require any form of input and output.

Nevertheless, NNs do not always outperform the traditional approaches (Markham and Ragsdale, 1995). Balakrishnan *et al.* (1994) compared NN results to the performance of a KM and found that the KM generated less misclassification than did the NN. Balakrishnan *et al.* (1994) compared performances of a NN and a KM using simulated data and real-life data. With simulated data, the NN's performances were very sensitive to the number of clusters, number of attributes, and error levels, while the KM's performances were sensitive only to error levels according to their analysis of variance. With the set of brand choice data, the NN provided clusters with similar sizes and high interpretability. However, the NN generated misclassified members more often than did the KM for the brand choice data.

Tam and Kiang (1992) pointed out some of the disadvantages of NNs, primarily that there is no

formal procedure in configuring NNs and that they consume a tremendous amount of computational time. Balakrishnan *et al.* (1994) added the observation that the performance of NNs worsens as the number of clusters increases. In addition, NNs are not effective in finding a global optimum (Pinter and Pesti, 1991).

Genetic algorithms (GA)

Genetic Algorithms (GA) emulate genetic production and are used to find solutions to optimization problems (Holland, 1992a). Members in each generation are usually called chromosomes, which represent a possible solution to the problems. Each chromosome consists of basic elements, called genes. As described by Goldberg (1989) and many others, the processes of GA are as follows. The initial population (or generation) of chromosomes is usually randomly selected and serve as the parents of the next generation. However, the probability of being selected is based on a chromosome's success in the first generation, since one with a higher evaluating value based on pre-determined criteria has a greater chance of being selected. Next, the selected chromosomes pass through one or more processes of crossover, mutation, and inversion. As shown in Figure 1, crossover is simply a process of swapping parts of the two selected chromosomes with a randomly selected crossover point. Then, all genes behind the crossover point of the two selected chromosomes are swapped. As shown in Figure 2, mutation deviates randomly selected genes. In the mutation process, target genes are randomly selected; then values of the selected genes are changed. The third process, inversion, flips the series of genes, as illustrated in Figure 3. With inversion, the GA randomly selects a series of genes, then reverses the series so it will be in backward order. The new chromosomes that result from these three operations substitute chromosomes with a low evaluating value from the previous generation, so the new generation consists of chromosomes with high evaluating values

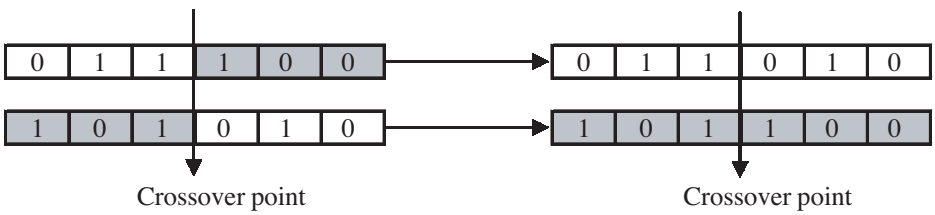

Figure 1 Crossover.

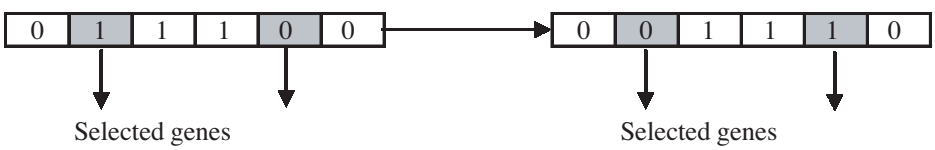

Figure 2 Mutation.

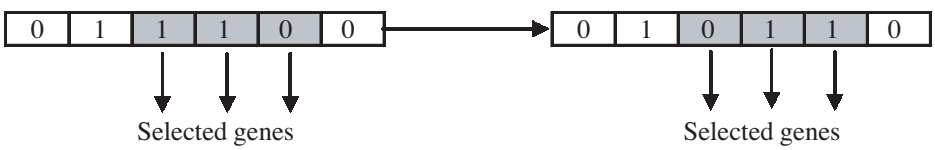

Figure 3 Inversion.

and new chromosomes. The process is repeated until the improvement of the evaluating value cannot be found or is less than a pre-determined value.

GAs are well known for finding a global optimum and are also widely compared to and combined with NNs to solve classification problems (Faulkenauer, 1998). The performance of GAs has been found to be promising in previous studies, such as Varetto (1998), who compared GAs to linear discriminant analysis (LDA) with data consisting equally of insolvent and solvent firms. In Varetto's study, LDA yielded a slightly better accuracy but consumed more time than the GA did.

Sexton and Dorsey (2000) configured three GA models and three NN models and examined six machine learning models in ten different real-world data sets: cancer, card, diabetes, gene, glass, heart, heart-c (heart data set without incomplete observations), horse, soybean, and thyroid. All three GA models generally outperformed all three NN models in terms of an average classification error.

Chiou and Lan (2001) investigated the clustering abilities of three configurations of the GA by comparing the clustering performances of the GA to that of a traditional clustering method. The results indicated that the best-configured GA, in their study, performed better than the hierarchical clustering techniques when the sample size was medium to large. However, the GA required tremendous storage space relative to the hierarchical clustering technique.

Methodology

The KM, the NN, and the GA all attempt to maximize the within-group variance and were evaluated on SAS version 8. The FASTCLUS procedure in SAS was used as the KM, which was allowed a maximum of 500 iterations and provided randomly selected seeds from the data set. The researchers also supplied the KM with a true number of clusters.A true number of clusters was also supplied to KM.

The fully connected feed-forward NN in this study consisted of three layers—input, hidden, and output layers—and was modified from Sarle's (1994) NN. Sarle (1994) provided a prototype of a simple and supervised NN using SAS's PROC NLP, while the NN in this study was supervised NN. The number of input nodes was exactly the same as the number of attributes, and the number of nodes in the hidden layer was equal to the number of desired or expected clusters. Once the initial weights were randomly selected, the datum was fed into the system through the input layer. The hidden layer applied a logistic function, sometimes called "softmax function," to the datum,

$$Z_i = e^{Y_i} \quad (1)$$

where Y_i is a linear combination of attributes for the i^{th} hidden node. Then, the output layer turned Z_i into a probability function using the multinomial logistic formula:

$$Prob_i = Zi \; / \; \Sigma(Z_i) \quad (2)$$

where $Prob_i$ represents a probability that the datum belongs to group *i*. Next, the output layer applied a competitive rule allowing a competitive node with the highest probability to win and assigned the datum to the winning node, which represented a cluster. The procedure was repeated until there was no data left unassigned.

Like the KM, the NN was also allowed a maximum of 500 iterations. Similar architecture is found in Bentz and Merunka (2000), who configured what was referred to as a "neural network with softmax output," except that their NN was a generalized form of the multinomial logistic function. The distinction between the NN and the multinomial logistic function is that the multinomial logistic function is a function for classification problems where group memberships are known before a clustering process begins.

The GA consisted of 10 chromosomes. Each chromosome, a possible solution for the cluster structure, comprised the means of the clusters, so each chromosome consisted of *pxk* digits, where *p* is the number of attributes and *k* is number of clusters. Accordingly, the fitness value—the reverse of the sum of square error—for each chromosome

was calculated and compared. Parents of new chromosomes in the next generation were selected through "the roulette-wheel selection," where opportunity to be selected was determined by the fitness value. (A chromosome with a high fitness value is assigned a higher chance of being selected as a parent for new members in the next generation. This process is also called "mating.") The reproduction process incorporated crossover, mutation, and inversion. A crossover procedure randomly mates two chromosomes, where the probability of each chromosome's being selected is calculated based on the chromosome's fitness value. Next, two crossover points are randomly selected and the two chromosomes are swapped between two crossover points, generating two new chromosomes are generated to replace the worst two chromosomes. Mutation points were randomly selected at a rate of 10% and mutated based on the range of the variable. (For example, I had $p \times k$ genes, so I selected 10% of $p \times k$.) Then, if the selected gene had a value of 2.74 on a variable that ranged from 0 to 10, the gene took 10 - 2.74 = 7.26 as a new value. Subsequently, a series of genes was randomly selected and inversed at a 10% rate. The production process was repeated until at least eight chromosomes with the same fitness value were present in the same generation or until the maximum number of iterations was reached.

Unlike the KM and the NN, the GA was allowed only 50 maximum iterations because it is extremely slow (Chiou and Lan, 2001). Besides reaching the maximum number of iterations, the GA could also stop if at least 8 out of 10 chromosomes indicated similar fitness value. In another words, if 7 other chromosomes provided fitness values within 300 units of the best fitness value, the system could stop.

Data and variables

Although a substantial body of research involves analysts' recommendations, determinant variables used to derive a recommendation have received little attention. Financial ratios were usually used in previous studies, since they are publicly available and are basic tools for financial analysts (Bouwman *et al.* 1987). Previous studies focused only on abnormal returns following the announcements of the recommendations; for example, Hirschey *et al.* (2000a) examined the effects of online recommendations on stock price and found significant stock price increases on "Buy" recommendations and decreases on "Sell" recommendations. Barber and Loeffler (1993) provided descriptive characteristics of four portfolios—pros' picks, dartboard stocks, S&P 500, and NYSE firms—and compared them in terms of growth, dividend yield, PE ratio, monthly volume, and beta. In general, variables used in previous studies represented growth, risk, market value, liquidity, and volume.

Data for this analysis were collected from Compustat's Research Insight for the year 2000. The analysis incorporated five variables—five-year average growth (X1), beta (X2), PE ratio (X3), dividend payout (X4), and volume (X5)—and five categories—buy, buy/hold, hold, sell/hold, and sell. Compustat's Research Insight provided 231 companies in the real estate investment trust (REITs) industry (SIC 6798) in the year 2000. I investigated the comparative performances of the underlying techniques in the REITs because (1) the industry is lucrative during economic downturns, when the interest rate is relatively low, (2) the industry effect can be avoided, and (3) the industry is regulated, so published information should be comparable to the information available to analysts. The average analysts' recommendations and five-year growth rate as of December 2002 were collected from Yahoo!Finance's stock screener website, which included 201 REITs. After eliminating the incomplete observations and providing alignment, 107 REITs remained.

Table 1 provides descriptive statistics on the data set. No "sell" recommendation exists in this industry, so it is not included in the table. Two large and two small clusters are present in the data set. Variables are on different scales. For example, X1

Table 1 Descriptive statistics.

Group		Buy	Buy/Hold	Hold	Sell/Hold	Sell
N		4.000	35.000	59.000	9.000	0.000
Max	X1	0.122	0.651	0.391	0.230	0.000
	X2	0.865	0.865	0.510	0.390	0.000
	X3	23.922	127.500	94.222	28.560	0.000
	X4	310.739	1601.000	523.962	293.241	0.000
	X5	757.400	10493.400	15502.400	2739.800	0.000
Min	X1	0.026	-0.445	-0.169	0.000	0.000
	X2	0.080	-0.132	-0.164	-0.075	0.000
	X3	4.439	5.630	5.429	6.172	0.000
	X4	47.703	46.448	46.379	53.509	0.000
	X5	32.800	38.400	219.600	206.300	0.000
Mean	X1	0.052	0.077	0.070	0.084	0.000
	X2	0.293	0.205	0.178	0.179	0.000
	X3	10.312	20.126	17.768	15.248	0.000
	X4	133.455	201.029	149.402	135.257	0.000
	X5	289.650	1840.206	3381.793	809.567	0.000
Std. Dev	X1	0.047	0.151	0.080	0.075	0.000
	X2	0.382	0.198	0.153	0.158	0.000
	X3	9.229	23.668	12.114	7.243	0.000
	X4	123.156	286.609	93.201	65.917	0.000
	X5	339.351	1983.974	3678.184	786.307	0.000

ranges from -0.445 to 0.651, while X5 ranges from 32.800 to 15,502.400. Mean ranges from 0.052 for X1 in the "Buy" category to 3381.793 for X5 in the "Hold" category. Standard deviation varies from group to group, so cluster sizes are not constant. The cluster structure in this data set is obviously complex, and clusters are not well separated.

Since the first and fourth groups are small or empty, I collapsed the first and second groups and the third and fourth groups to make two larger groups, labeled "Buy" and "Not-Buy." The descriptive statistics for the new data set, not shown, still indicated that variables are not on the same scale, that cluster sizes are not equal, and that clusters are not well separated. This may worsen the KM's performance.

Table 2 reports the correct classification. Overall, correct classification alone indicated that the GA was the best and the KM was the worst among the three algorithms in their ability to replicate the analysts' recommendation. However, if I look closely, the NN did not classify at all, either because none of the stocks appealed to it or because it could not reach an optimum within the given number of iterations. The second explanation seems to be realistic since, in previous studies, NNs were allowed to run for at least 10,000 iterations. Thus, the NN was the worst among the three algorithms, given the number of iterations. It should also be noted that the KM consumed the least and the GA consumed the most amount of time.

Next, I tested whether there was a better than even chance of overall correct classification with the KM and the GA using (3). Both yielded classification accuracies statistically higher than 50%, with p-values of 0.031 and 0.00, respectively.

Table 2 Correct classification.

Group	Actual	KM		NN		GA	
		Buy	Not-Buy	Buy	Not-Buy	Buy	Not-Buy
Buy	39	1	38	0	39	5**	34
Not-Buy	68	6	62	0	68	3	65
		7	100	0	107	8	99
Overall	107	63**	68***	70***			
% success	36.45	14.26***	00.00	62.50***			

** significant at 0.05 level

*** significant at 0.01 level

$$z = \frac{\bar{p} - p_0}{s_{\bar{p}}} = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \qquad (3)$$

Next, I compared the correct classifications of the GA and the KM using (4). Based on the overall correct classifications for the KM and the GA, I did not find statistical evidence that one classified better than the other. I then considered the effectiveness of the KM and the GA in identifying stocks with the "buy" recommendation and found that the KM recommended seven stocks to buy while analysts recommended only one. Thus, the percent correct classification of the KM for stocks with a "Buy" recommendation was 1/7 or 14.29%. On the other hand, the GA selected eight stocks to buy, while analysts recommend five, giving the GA a correct classification rate of 62.50%. This test indicated that the GA outperforms the KM at the 0.05 significance level, since the p-value for the test was 0.0464.

When the analysts' recommendation was not available, I tested whether the correct classification of the KM and the GA were statistically better than a random "buy" stock selection. The results showed that the KM yielded a chance of success significantly lower than a random selection did at the 0.01 level, since the p-value was 0.000003. On the other hand, the GA was superior to a random selection at any reasonable significance level, since the p-value is 0.000048.

$$z = \frac{\bar{p}_1 - \bar{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_1}}} \qquad (4)$$

The results presented in Table 2 should be considered persuasive, even though I found no significant difference between the overall correct classifications by the KM and the GA. To take a long position, investors consider only stocks with a "buy" rating, so the percent of times the process can correctly classify stocks in the "buy" category is more important than the overall rate of correct classifications.

Although the GA provides a success rate significantly better than random, while the KM is statistically worse than random, one should keep in mind that the GA's performance is compared to a random chance. Once the analysts' recommendation becomes available, the KM's performance becomes 100% and the superiority of the GA is nullified. Therefore, the GA is useful only before the analysts' recommendation becomes public. Investors should also be aware of the efficiency market hypothesis (Block and Hirt, 2000), which contended that the market is efficient and all investors are informed of the same set of information. This suggests that all investors can obtain the GA recommendation of GA and, as a result, all investors would bid for "buy" stocks and ignore "not-buy" stocks in order to achieve an optimal investment portfolio. Thus, the "buy" stocks become less attractive and "not-buy" stocks become more attractive.

Finally, I formulated three equally weighted portfolios: The first portfolio contained 39 stocks with a "buy" or "buy/hold" rating recommended by analysts; the second portfolio contained stocks recommended by the KM and the third portfolio contained eight stocks with a "buy" rating recommended by the GA. Using Research Insight's Compustat, I recorded the monthly return for the portfolios from January to November 2001 and obtained their short-term return on investment, shown in Table 3. The maximum ROI of the first and third portfolios was the same, while the minimum differed. Still, the third portfolio, containing eight stocks with a "buy" rating recommended by the GA, yielded the highest average return and the lowest variation among the three portfolios. The GA was best able to distinguish the best and the worst stocks from all 39 stocks that the analysts recommended, but the returns on investment of the first and the third portfolios were not significantly different, since the p-value was 0.3375.

CONCLUSION

Many investors invest according to analysts' recommendations, and some studies, indeed, detect abnormal return associated with analysts' recommendations. A tool that imitates the mechanism of the recommendation process should enable individual investors to formulate a profitable investment strategy and to select stocks that analysts have not covered—or to select stocks before analysts announce their recommendations. In previous studies, researchers have acknowledged that the decision-making process is sophisticated and have examined primarily linear models (Hirchey *et al.* 2000a, 2000b). This paper presents the GA's effectiveness in identifying attractive stocks, since the GA can overcome (1) a sophisticated relationship between financial information and stock recommendations, and (2) violations of probability distribution which usually occur in finance problems.

Along with linear models, previous studies have frequently analyzed financial ratios since they are publicly available and are basic tools for financial analysts (Bouwman *et al.* 1987). Qualitative variables have been largely ignored because the qualitative information extracted from analysts' reports is subjective and can be affected by the researcher's experience in the field (Breton and Taffler, 2001). A lack of expertise in the field makes the decision process even more complicated (Mear and Firth, 1987).

This study's original five recommendation categories have been collapsed into two groups because of low frequencies in some categories. Compared to the KM and the NN, the GA demonstrates a promising performance as a tool with which investors can develop a profitable investment strategy. An investment portfolio selected by a GA yields a higher chance of selecting a stock to buy than a KM, a NN, or a random process. The GA exceeds 62% in picking a "buy" stock versus 36.45 and 14.26% for a random process and a KM, respectively. However, it is worth noting that the KM consumes the least amount of time, while the GA consumes the most amount of time. Another caution is that the publicity of analysts' recommendations would undermine the promising performance of the GA, since the 36.45% performance of a random selection would then become a 100% "buy" recommendation. In respect to the efficiency market hypothesis, investors should also be aware that other investors could also obtain the recommendation of

Table 3 Average return on investment.

	Mean return (%)	p-value	Standard deviation	Maximum	Minimum
1	11.73		0.1797	37.92	-39.77
2	-0.06	0.4968	0.1781	35.86	-37.26
3	14.70	0.3375	0.1652	37.92	-30.63

the GA, so over-buying would render the "buy" stocks less attractive and the "not-buy" stocks more attractive.

LITERATURE CITED

Altman, E. I., G. Marco and F. Varetto. 1994. Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks (The Italian experience). *Journal of Banking and Finance* 18: 505-529.

Balakrishnan, P. V., M. C. Cooper, V. S. Jacob and P.A. Lewis. 1994. A Study of the Classification Capabilities of Neural Networks using Unsupervised Learning: A Comparison with K-means clustering. *Psychometrika* 59(4): 509-525.

Barber M. B. and D. Loeffler. 1993. The "Dartboard" Column: Second-Hand Information and Price Pressure. *Journal of Financial and Quantitative Analysis* 28(2): 273-284.

Barrett, N. J. and I. F. Wilkinson. 1985. Export Stimulation: A Segmentation Study of the Exporting Problems of Australian Manufacturing Firms. *European Journal of Marketing* 19(2): 53-72.

Bentz, Y. and D. Merunka. 2000. Neural Networks and the Multinomial Logit for Brand Choice Modellnig: A Hybrid Approach. *Journal of Forecasting* 19(3): 177-200.

Block, S. B. and G. A. Hirt. 2000. *Foundations of Financial Management, 9th ed.* Boston: Irwin McGraw-Hill.

Bouwman, M. J., P. A. Frishkoff and P. Frishkoff. 1987. How do Financial Analysts Make Decision? A Process Model of the Investment Screening Decision. *Accounting, Organizations and Society.* 12(1): 1-30.

Breton, G. and R. J. Taffler. 2001. Accounting Information and Analyst Stock Recommendation Decisions: A Content Analysis Approach. *Accounting and Business Research* 31(2): 91-101.

Chen, S. K., P. Mangiameli and D. West. 1995. The Comparative Ability of Self-organizing Neural Networks to Define Cluster Structure. *Omega* 23(3): 271-279.

Chiou, Y. and L. W. Lan. 2001. Theory and Methodology: Genetic Clustering Algorithms. *European Journal of Operational Research* 135: 413-427.

Davies, D. L. and D. W. Bouldin. 1979. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1: 224-227.

Falkenauer, E. 1998. *Genetic Algorithms & Grouping Problems.* New York: Wiley.

Francis, J. and L. Soffer. 1997. The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecast Revisions. *Journal of Accounting Research* 35(2): 193-211.

Givoly, D. and J. Lakonishok. 1984. The Quality of Analysts' Forecasts of Earnings. *Financial Analysts Journal* 40(5): 40-47.

Goldberg, D. E. 1989. *Genetic Algorithms in Search.* New York: Addison-Wesley.

Hirschey, M., V. J. Richardson and S. Scholz. 2000a. Stock-Price Effects of Internet Buy-Sell Recommendtions: The Motley Fool Case. *The Financial Review* 35: 147-174.

_____ . 2000b. How "Foolish" are Internet Investors? *Financial Analysts Journal* 56(1): 62-69.

Holland, J. H. 1992. *Adaptation in Natural and Artificial Systems, 1992 ed.* Cambridge: MIT Press.

Hruschka, H. and M. Natter. 1999. Comparing Performance of Feedforward Neural Nets and K-means for Cluster-Based Market Segmentation. *European Journal of Operational Research* 114: 346-353.

Markham, I. S. and C. T. Ragsdale. 1995. Combing Neural Networks and Statistical Predictions to Solve the Classification Problem in Discriminant Analysis. *Decision Sciences* 26(2): 229-242.

Mear, R. and M. Firth. 1987. Cue Usage and Self-

Insight of Financial Analysts. *The Accounting Review* 62(1): 176-182.

Punj, G. and D. W. Stewart. 1983. Cluster Analysis in Marketing Research: Review and Suggestions for Application. *Journal of Marketing Research* 20(2): 134-148.

Sarle, W. S. 1994. Neural Network Implementation in SAS® Software. *Proceedings of the Nineteenth Annual SAS Users Group International Conference.* 1-28.

Sexton, R. S. and R. E. Dorsey. 2000. Reliable Classification using Neural Networks: A Genetic Algorithm and Backpropagation Comparison. *Decision Support Systems* 30(1): 11-22.

Slater, S. F. and E. M. Olson. 2001. Marketing's Contribution to the Implementation of Business Strategy: An Empirical Analysis. *Strategic Management Journal* 22(11): 1055-1067.

Tam, K.Y. 1991. Neural Network Models and The Prediction of Bank Bankruptcy. *Omega* 19(5): 429-445.

Tam, K.Y. and M. Y. Kiang. 1992. Management Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science* 38(7): 926-947.

Varetto, F. 1998. Genetic Algorithms Applications in the Analysis of Insolvency Risk. *Journal of Banking & Finance* 22: 1421-1439.