International Journal of Information Technology & Decision Making
Vol. 5, No. 1 (2006) 227–240

A DECISION TREE-BASED CLASSIFICATION APPROACH TO RULE EXTRACTION FOR SECURITY ANALYSIS

N. REN, M. ZARGHAM and S. RAHIMI*

Department of Computer Science
Southern Illinois University Carbondale
Mailcode 4511, Carbondale, IL, 62901-4511, USA
**rahimi@cs.siu.edu*

Stock selection rules are extensively utilized as the guideline to construct high performance stock portfolios. However, the predictive performance of the rules developed by some economic experts in the past has decreased dramatically for the current stock market. In this paper, C4.5 decision tree classification method was adopted to construct a model for stock prediction based on the fundamental stock data, from which a set of stock selection rules was derived. The experimental results showed that the generated rules have exceptional predictive performance. Moreover, it also demonstrated that the C4.5 decision tree classification model can work efficiently on the high noise stock data domain.

Keywords: Stock selection rules; stock prediction model; decision tree; data mining; C4.5 decision tree algorithm.

1. Introduction

The stock market is a complex system affected by various factors which make it very hard to be predicted. Research has been conducted to predict the future behavior of the stock market using various techniques and approaches for which computer technology has played a very important role. Today, there are various stock analyzer programs available as commercial products or research projects. They range from pure mathematical models, databases, and expert systems to neural networks and fuzzy systems. These software packages are based on either a technical analysis approach or a fundamental analysis approach. Stock selection rules are the true spirit of most of the fundamental analysis based portfolio selection systems.

In the real world, when an investor selects stocks to construct an optimal portfolio, some rules of thumb are usually used to evaluate stocks. For instance, usually stocks with a stable earning record, low price-to-earning ratio, low debt, and/or high dividend yield are selected. Without any doubts, there are some principal rules associated with choosing the right stocks, although the rules may differ from one era of economics to the other. Finance experts have set up stock selection rules mainly based on some finance and economic theories such as Peter Lynch investment rules, Warren Buffet investment rules, Benjamin Graham investment rules, Philip Fisher investment rules, and T. Rowe Price investment rules.[1] Many of the above rules have had good performances during some period of time in the past. For example, Graham's ten rules had their best performance before 1976. However, with the fast development of the stock market and the globalization of the world economy, present stock markets have changed dramatically as compared to the market prior to 1980. The applicability of many existing rules is unsatisfactory and their performance has decreased substantially. Therefore, the main objective of this study is to produce a new set of rules, which are to be used to select stocks with possible high return.

In recent years, some data mining based techniques such as classification, rule induction, etc., have been utilized to analyze the stock market which have led to the development of promising stock market prediction models.[2–5] John and Miller[2] introduce the Recon system, which has been developed to construct the long/short term portfolios, based on the rule-induction system developed at Lockheed-Martin. In Refs. 3–5, a minimal rule generation and contextual features analysis algorithm from IBM research project R-MINI have been presented. It can be observed from these research works that combining some data mining techniques with appropriate finance fundamental factors can produce a promising stock prediction model, which guarantees high investment return.

Classification in the context of data mining is defined as learning a function that classifies a data item into one of predefined classes. Classification rules can be considered as particular kinds of prediction rules. Many classification algorithms and models have been developed; among them, decision tree is a classical and popular classification model which was adopted and used in many different areas. From a decision tree model, a set of rules can be produced to make predictions. The C4.5 decision tree algorithm,[6] a modified version of the ID3 tree algorithm,[7] is proven to be accurate, efficient, and robust by many researches.[6,8–10] It is capable of generating simple and concise rules while its construction cost is lower than the construction cost of other computational models such as neural networks and Bayesian inference. In addition, some significant features of the C4.5 algorithm make it suitable for some particular domains, such as securities analysis, in which data contain a high amount of noise. For example, it can classify records that have unknown attribute values by estimating the probability of the various possible results. Moreover, it can deal not only with the attributes that contain discrete values but also with the attributes that represent numbers,[9] which is very important for security analysis applications. In this work, the C4.5 decision tree model was implemented based on the fundamental stock data from which a set of stock selection rules was derived for portfolio construction. The experimental results, discussed in later sections, show that the generated rules have an exceptional predictive performance.

The rest of the paper is organized as follows. After discussing the pertinent works on stock prediction in Sec. 2, the decision tree model construction and rule extraction procedure are presented in Sec. 3. The rule validation and performance analysis are reported in Sec. 4. To further illustrate the applicability of the C4.5 decision tree model on the fundamental stock data domain, a second set of rules was extracted and validated in Sec. 5. Finally, the discussion is given in Sec. 6.

2. Pertinent Works on Stock Prediction

In the past several decades, numerous researches have been done on the stock market predication and many techniques and methodologies have been successfully applied in this area. The statistical-based approaches, neural network-based approaches, and classification rule-based approaches are the most striking and promising ones among others.

Statistics has been applied for analyzing the behavior of the stock market for more than half a century. The majorities of the classical statistical stock market models focus on the stock time series prediction and have achieved some acceptable results. However, due to the random-walk process that the stock market follows and the non-linearity in the stock data set, the time series models usually cannot reach very high prediction accuracy. The hit rate of 54% is considered as a satisfying result for stock prediction in this class of approaches.[11]

To reach better hit rates, other techniques such as neural networks were applied for stock prediction. The most obvious advantage of neural network systems used in stock markets is that they can outperform the classical statistical methods with 5–20% higher accuracy rate.[12,13] In addition, neural network systems show better flexibility for dealing with uncertain and missing data. However, the neural network method has its limitations. For example, it is not easy to find the best network architecture for a specific problem. As the networks become more complicated, the reliability of results may decrease because it is hard for the norm of the gradient of the error function to reach a very low value when complex networks are trained.[14] Since a neural network is a multiple layered graph, the predication pattern in the training data set can be buried in both the structure of the network and the weights assigned to the links between the nodes. Therefore, although the prediction result can be reported clearly, it is hard to express prediction patterns in a straight forward way. Also, it may take very long time to train a neural network.

Moreover, some other research, such as in Refs. 2 and 4, has also shown that the stock market domain can be effectively modeled by classification rules which are induced from available historical data using data mining techniques. In contrast to neural networks, the classification rules, usually in disjunctive normal form (DNF), are capable of expressing the prediction patterns in a straight forward way. Rule induction algorithms could generate simple and efficient rule sets that cover all the training data, while analysts can refine the induced rule set easily. Although many classification rule induction algorithms have been successfully applied to the stock market such as,[2,4,5] they all require specific data model, and work well only during some specific period of time. Since, except time, many other factors and variables may affect the behavior of the stock market, the classification models have to consider a large number of independent variables in order to find high-performance prediction rules.

Based on the above discussion and our objective (finding out an efficient methodology to construct a prediction model, from which to generate a new set of stock prediction rules for portfolio construction), the classification rule based approach is adopted in this work. In discussion section, we have compared the results of our approach against the results of some other statistics, neural network, and rule induction approaches.

3. Decision Tree Classification Model and Rule Extraction

3.1. *C4.5 decision tree classification model*

Decision tree, a popular classification methodology, is capable of classifying a pattern by a sequence of questions. In this context, a pattern is a data record in a data set, which is defined by several attributes. The classification of a particular pattern begins at the root node, which asks for the value of a particular attribute of the pattern. Based on the answer, the appropriate branch to a subsequent or descendent node is followed. This procedure is repeated until a leaf node is reached, which has a category label. Numerous decision tree construction methods have been introduced in the past several decades. Among these methods, the C4.5 algorithm is considered to be a robust, accurate, and efficient algorithm capable of constructing simple and effective decision trees from which classification rules can be extracted.

The fundamental principle underlying decision tree construction is that of simplicity: trying to create a simple and compact tree with few nodes. Therefore, the decision tree construction methods focus on deciding which attribute should be used to split the training data set at each node such that it can create a simple model that explains the data appropriately. Therefore, the splitting attribute at each node should make the data reach the immediate descendent nodes as pure as possible.[15] Several different mathematical measures of impurity have been proposed to help select the splitting attribute, such as entropy impurity, and Gini impurity.[15] C4.5 adopts entropy impurity and automatically selects the attribute that provides the highest information gain[7] as the splitting attribute such that the splitting attribute can partition the data set with the best improvement on purity. However, this splitting attribute selection criterion favors the attribute with the largest number of values. In the extreme case that every data record has a unique value for the first splitting attribute, the decision tree would stop growing at depth one. Usually, such a decision tree is not desired. In order to avoid such cases, gain ratio[6] instead of information gain is used. When creating the decision tree by the C4.5 algorithm, the splitting attribute for root node is selected first and the training data set is divided into subsets by the selected attribute. After that, more splitting attributes are selected recursively for each produced subset to construct the corresponding subtree. The algorithm uses heuristics for pruning based on the statistical significance of splits.

3.2. *Data collection and preparation*

In this work, the data set used to construct the C4.5 decision tree and to test the produced rules' performance is the S&P 500 data from 1993 to 2002. The data were downloaded from the Standard and Poor's compustat database.[16]

Table 1 illustrates the format of the input data set, in which each record has five predicting attributes and one target attribute. The predicting attributes used in this model are fundamental factors including price to book ($PtoB$), earning growth (EG), earning stability (ES), price earning ratio monthly (PEM), and current ratio (CR).[17] The target attribute used here is the one year total return. This target attribute is a numerical value in the raw data set and is categorized based on the following criteria: if its value is less than zero, then the return is "bad"; if its value is greater than or equal to zero and less than 10, then the return is "ok"; if its value is greater than or equal to 10 and less than 30, then the return is "good"; if its value is greater than or equal to 30 and less than 100, then the return is "very good"; and finally if its value is greater than or equal to 100, then the return is "excellent".

3.3. *Rule extraction*

The decision tree classification model in this paper was constructed from S&P 500 data from 1993 to 1998 (the training data set). Figure 1 illustrates part of the constructed decision tree. The root node of the constructed tree is "price to book". So first, the stocks in the training data set are divided into two subsets based on their price to book value. One of the subsets is made up from the stocks with the price to book value greater than 4.065 and the other subset has the stocks with the price to book value no greater than 4.065. After the first level splitting, all the subsets are recursively split in the same way until the leaf nodes are reached. Each path in the decision tree generates a classification rule. After simplification and

Table 1. An example of partial input data set.

Price to Book	Earning Growth	Earning Stability	Price Earning Ratio Monthly (PEM)	Current Ratio	One Year Total Return
2.39	−48.276	5	71.233	6.534	Bad
2.171	−53.363	5	104.462	5.247	Bad
7.517	−29.377	5	32.565	1.396	OK
1.545	41.615	5	12.963	0.608	OK
4.366	60.312	5	31.965	1.506	Good
2.366	81.159	5	21.5	2.264	Very good
3.447	156.25	4	20.528	0.769	Very good
7.742	243.503	5	65.158	11.496	Excellent

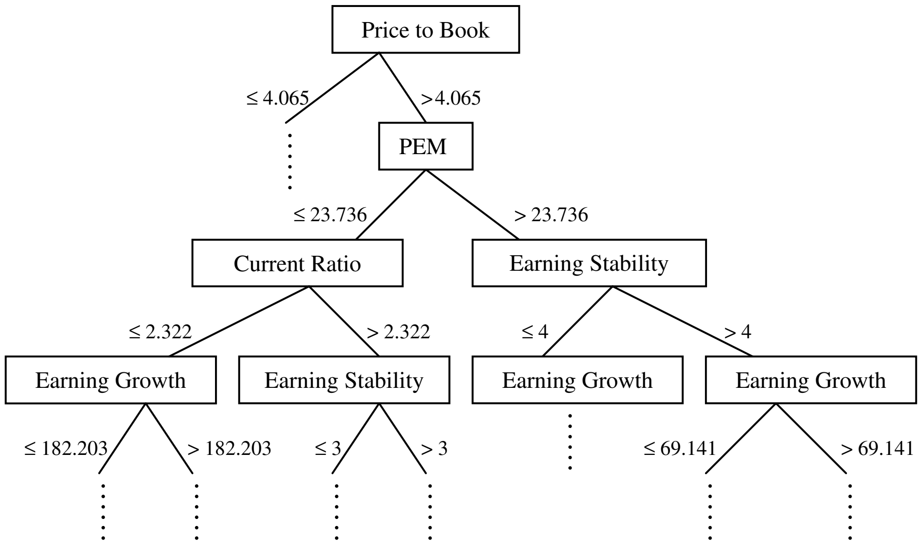

Fig. 1. The partial constructed decision tree.

generalization, the following rules were derived from the constructed decision tree:

Rule 1: If $PEM > 23.736$ and $ES \geq 4$ and $EG > 69.141$ and $PtoB > 9.462$, then the return is "very good" (return $\geq 30\%$).

Rule 2: If expected return is "excellent" or "very good", the Current Ratio should be greater than 0.5 and less than 6.2.

Rule 3: If expected return is "good" (return $\geq 10\%$), earning stability should be no less than 3.

Rule 4: If expected return is "very good" (return $\geq 30\%$), earning stability should be no less than 4.

4. Rule Validation and Performance Analysis

To evaluate the proposed decision tree classification model, the derived rules in the previous section are validated and their performances are analyzed in this section. Validation of a rule is achieved by applying the extracted rules to the training data set while performance evaluation is done by applying the rules to a more recent S&P 500 test data set.

4.1. *Rule validation*

In order to verify that Rule 1 is the rule hidden in the training data set, this rule was used to select stocks from the S&P 500 data set for the years 1993–1998 (the training data set). A total of 90% of the selected stocks, using Rule 1, have returns greater than 30% and 98% of them have returns greater than 10%. Based on the selected stocks, Rule 1 is illustrated by Figs. 2 and 3. On the return axis of

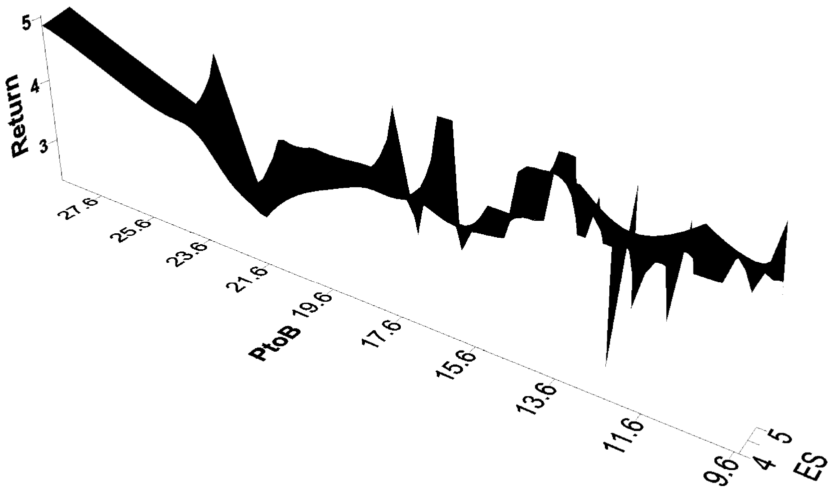

Fig. 2. Relationship between price to book, earning stability, and return (for Rule 1).

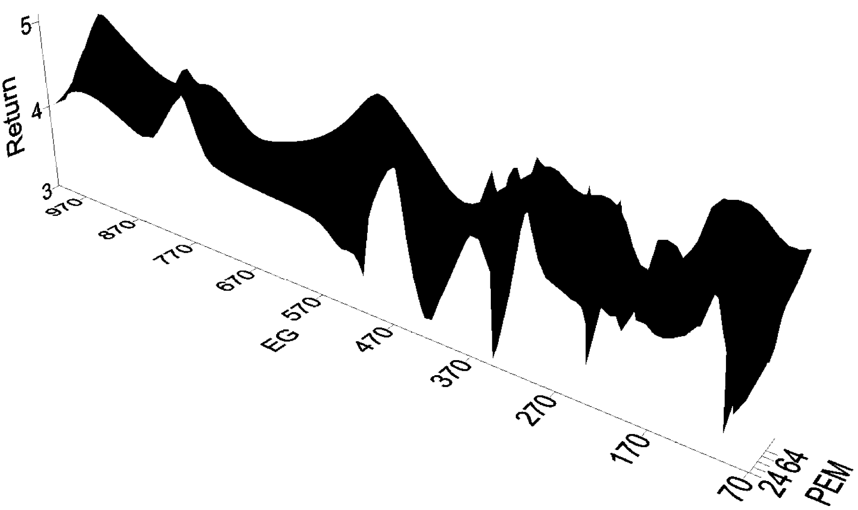

Fig. 3. Relationship between earning growth, price earning monthly, and return (for Rule 1).

these figures, number 3 means a "good" return (10% $\leq$ return $<$ 30%), number 4 means a "very good" return (30% $\leq$ return $<$ 100%), and number 5 means an "excellent" return (return $\geq$ 100%). It can be observed from Fig. 2 that the majority of the stocks with the price to book (*PtoB*) value greater than 9.6 and the earning stability (*ES*) value no less than 4 have an "excellent" or "very good" return. When the selected stocks are plotted in terms of earning growth (*EG*), price earning ratio monthly (*PEM*), and return, it becomes clear that the stocks with a "very good" or "excellent" return have the earning growth value greater than 69.141 and the price earning ratio monthly value greater than 23.736, as shown in Fig. 3. Therefore, Rule 1 is proven to be an effective rule in the S&P 500 data from 1993 to 1998. The combination of *PtoB*, *PEM*, *EG* and *ES* in Rule 1 looks promising in the selection of stocks with "very good" returns.

Rule 2 is visualized and validated in Figs. 4 and 5. These figures are stack column graphs in which each column consists of several segments, each of them representing the number of the stocks with a specific return. For example, the second column in Fig. 4 shows that, among the stocks selected by Rule 1, when the current ratio (*CR*) is between 1.02 and 1.41, there are four stocks with "excellent" returns (return $\geq$ 100%), eight stocks with "very good" returns (30% $\leq$ return $<$ 100%), and three stocks with "good" returns (10% $\leq$ return $<$ 30%). Figure 4, based on the stocks selected by Rule 1 from the S&P 500 1993–1998 data set, shows that the stocks with different returns have very similar distribution patterns regarding the current ratio. It also shows that the favorable current ratio range is between 0.5 and 6.2 in order to select the stocks with expected good return. This pattern can also be revealed from Fig. 5, which shows very similar distribution of the stocks with different returns in the S&P 500 1994–2001 data set. Therefore, it can be concluded that the current ratio is not a determinant attribute for separating the stocks with good returns from the ones with bad returns because of their similar current ratio distribution pattern. According to a well-known stock selection criterion, Graham's stock selection rules,[18] the current ratio should be 2 in order to select stocks that

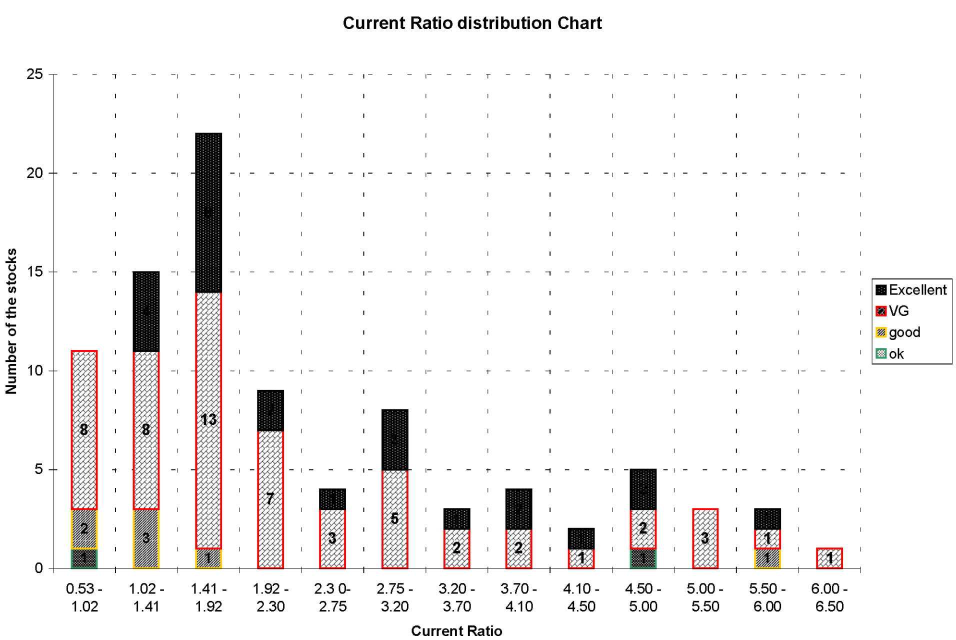

Fig. 4. Current ratio distributions for the selected stocks by Rule 1.

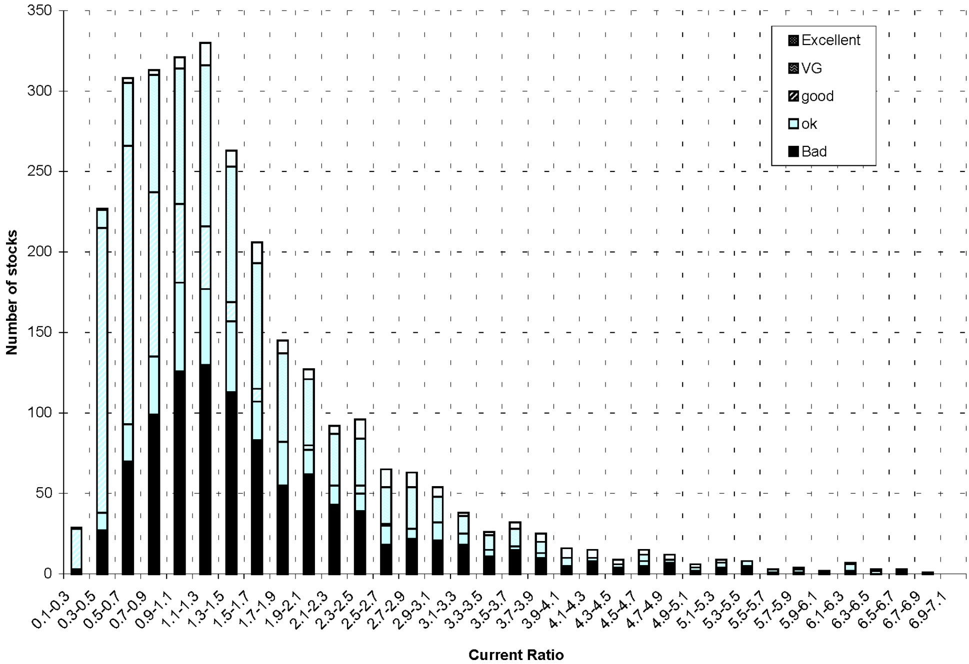

Fig. 5. Current ratio distribution for S&P 500 stocks from 1994 to 2001.

have a good return. However, it can be observed from Figs. 4 and 5 that, in order to get a high accumulated return, the current ratio should vary in a specific range instead of being a fixed value, as Graham's rules indicated.

It is observed that all the selected stocks by Rule 1 have the earning stability (ES) no less than 4. This observation proved that Rule 4 is a valid stock selection rule. In addition, it is found that the majority of stocks with the return greater than 10% in the training data set have the earning stability no less than 3. Therefore, Rule 3 is proven a valid rule as well. Although Rules 3 and 4 cannot be used individually, they could be used with other rules to construct optimal portfolios.

4.2. *Rule performance analysis*

In order to test the performance of Rule 1, four 1-year portfolios were constructed based on the testing data set of S&P 500 from 1999 to 2002. The statistical information of the constructed portfolios is stated in Table 2, one record for each portfolio. For instance, 48 stocks were selected to form the 1999 portfolio, from which 72.92% (35 stocks) had the return greater than 30% and 81.25% (39 stocks) had the return greater than 10%. The cumulative total return for the 1999 portfolio was 6019.51% and the average total return per corporation was 125.41% while the S&P 500 index return in 1999 was 19.53%. The selected stocks for 2001's portfolio are presented in Table 3. The 1-year total return in Table 3 is calculated using the following

Table 2. Statistical information of the portfolios constructed by Rule 1.

	Number of Selected Stocks	% of Selected Stocks with Return > 30%	% of Selected Stocks with Return > 10%	Cumulative Total Return (%)	Average Total Return Per Corp. (%)	S&P 500 Index Return (%)
1999 Portfolio	48	72.92	81.25	6019.51	125.41	19.53
2000 Portfolio	48	62.50	75.00	3277.56	68.28	−10.14
2001 Portfolio	14	50.00	57.14	276.51	19.75	−13.04
2002 Portfolio	7	28.57	28.57	154.04	22.01	−23.35
Average	29	—	—	2431.90	58.86	−6.75

Table 3. The stocks in 2001 portfolio.

Tic	Company Name	Close Price Dec. 2000	Close Price Dec. 2001	One-Year Return (%)
MXIM	MAXIM INTEGRATED PRODUCTS	67.94	44.21	−34.424
AGN	ALLERGAN INC	96.81	75.05	−22.131
PFE	PFIZER INC	46.00	39.85	−12.463
AMGN	AMGEN INC	63.94	56.44	−11.726
PAYX	PAYCHEX INC	38.43	34.65	−8.78
GDT	GUIDANT CORP	53.94	49.8	−7.671
SYK	STRYKER CORP	50.59	58.37	15.576
BBBY	BED BATH & BEYOND INC	24.63	33.4	35.635
HDI	HARLEY-DAVIDSON INC	39.75	54.31	36.983
FRX	FOREST LABORATORIES -CL A	29.62	40.85	37.914
APOL	APOLLO GROUP INC -CL A	18.14	26.247	44.698
CE	CONCORD EFS INC	21.97	32.78	49.212
ABC	AMERISOURCEBERGEN CORP	47.00	70.95	50.957
EBAY	EBAY INC	16.50	33.45	102.727
	The Portfolio Total Return			276.51

equation:

$$\text{One year total return} = ((PRCCM01 * TRFM01)/(PRCCM00 * TRFM00) - 1) * 100$$

This equation is provided by Standard and Poor's compustat database.[16] In this equation, $PRCCM01$ and $PRCCM00$ (Price-Monthly − Close) represent the absolute closed market prices for December of 2001 and 2000; $TRFM01$ and $TRFM00$ (Total Return Factor − Monthly) represent the multiplication factor for calculating the total return to shareholders for December 2001 and 2000, respectively, and include Cash Equivalent Distributions along with reinvestment of dividends and the compounding effect of dividends paid on reinvested dividends. The following observations can be made from Table 2:

1. The highest 1-year accumulated total return from the constructed portfolios is 6019.51%, which is from the 1999 portfolio. It is obvious that this is an excellent total portfolio return.

2. The lowest 1-year accumulated total return from the constructed portfolios is bounded by 154.04%, which is from the 2002 portfolio. Although this is much lower than the total return from the 1999 portfolio, it is still considered very high.
3. In some of the constructed portfolios, the percentage of the stocks with 1-year total return greater than 30% reaches as high as 72.92% and the percentage of the stocks with a 1-year total return greater than 10% reaches as high as 81.25%. This fact shows that Rule 1 has excellent prediction accuracy for this dataset.
4. The returns from the constructed portfolios are much higher than those of the S&P 500 index. The average total return from the constructed portfolios, 2431.90%, is much higher than the average S&P 500 index return, which is −6.75%.

The above observations show that Rule 1 is a very efficient stock prediction rule by which high performance portfolios can be constructed.

5. More Experimental Results

In order to further test the effectiveness of the C4.5 decision tree classification model and get more stock prediction rules, another decision tree was constructed based on the training data set of S&P 500 from 1994 to 2001. The following rules were extracted from this constructed tree:

Rule 5: If $PtoB > 4.092$ and $PEM > 40.077$ and $EG > 121.698$, then return is "very good" (return $\geq 30\%$).

Rule 6: If $PtoB > 4.721$ and $PEM > 16.377$ and $CR > 1.501$ then return is "very good" (return $\geq 30\%$).

The generated rules' prediction performance are tested on the S&P 500 1993 data set. Two portfolios were constructed by each rule respectively. The statistical information of the two constructed portfolios is stated in Table 4. From this table, it can be found that the accumulated total return from the portfolio constructed by Rule 5 is 364.02% and the accumulated total return from the portfolio constructed by Rule 6 reaches 2553.30%. The accumulated total returns of the constructed portfolios showed that these two stock prediction criteria have a high predicting performance.

This set of new rules is very similar to the rules presented in Sec. 3.3 and the analysis in this section shows that this set of rules has an excellent prediction performance as well. These experimental results further demonstrate that the C4.5 decision tree can effectively work on the stock fundamental data domain and is a robust and efficient classification approach.

Table 4. Statistic information of the portfolios based on S&P 500 1993 data.

	Number of Selected Stocks	% of Selected Stocks with Return > 30%	% of Selected Stocks with Return > 10%	Average Total Return Per Corp. (%)	Cumulative Total Return (%)	S&P 500 Index Return (1993) (%)
Portfolio based on the first rule	5	100.00	100.00	72.80	364.02	7.06
Portfolio based on the second rule	47	61.70	78.72	54.33	2553.30	

6. Discussions

By the experimental results in this study, it was shown that efficient stock prediction rules can be extracted from the fundamental stock data. These rules are capable of selecting stocks with high returns and constructing optimal portfolios. The following can be concluded from the extracted rules:

1. The constructed decision tree and the extracted rules based on the 1990's stock market data reveal that price to book, price earning ratio, earning stability, earning growth, and current ratio are important indexes to select stocks with a "very good" return.
2. Some specific combinations of these indexes, such as price to book, price earning ratio, earning stability, and earning growth, in Rule 1 are efficient stock selection criteria to construct high performance portfolios, as shown in Tables 2 and 4.
3. The extracted Rules 3 and 4 indicate that earning stability should be no less than 3 in order to select the stocks with a "good" return.
4. In order to get good portfolio returns, the current ratio should vary in a specific range instead of being a fixed value, such as from 0.5 to 6.2 as shown in Fig. 4.

In addition, this study illustrated that the decision tree based classification model can be an efficient predictive model, which can extract high performance prediction rules from noisy data domains such as the fundamental stock data set.

Since, in general, research projects in this area use different benchmarks and their portfolios are constructed under different trading conditions, during different periods, it is impossible to compare and rank them concretely. However, a rough idea can be attained by comparing their efficiency relative to their benchmarks. To compare our approach verses others, some promising experimental results from other works are presented here.

As discussed in previous sections, the statistic approach mainly adopts various regression methods and focuses on individual time series prediction. Instead of working on individual time series,[19] utilizes the standard regression technique to generate the rank measure that takes into account a large number of securities and grades them according to their relative returns. According to Hellstrom,[19] a constructed portfolio (1993–1997) based on the rank prediction method had mean annual profit of 123.6%, significantly higher than its benchmark (the Swedish General index) of 27.4%.

Furthermore, in Ref. 20, a multilayer perceptron neural network and a probabilistic neural network were implemented to predict the incline, decline, or steadiness of S&P 500 Index. The constructed portfolio (1994–1995) based on the advice of the neural network prediction shows annual return of 108.30% while its corresponding benchmark index return is 100.86%.

As another example in the area of rule induction,[21] employees the Recon system on the stock market and induces a set of classification rules to model the given data. The constructed Recon portfolio (with 1% transaction cost) had total return of 238% over a 4-year period, significantly outperforming its benchmark (a market-cap-weighted portfolio of all the stocks in the universe from which Recon was selecting), which returned 93.5% over the same period of time.

By comparing the relative efficiencies of the portfolios generated by our proposed approach and the above listed portfolios, it can be observed that the constructed portfolios in this study have exceptional performance (without transaction cost, the average cumulative total return is as high as 2431.90%) and significantly outperform the adopted benchmark (−6.75%) in the same period of time.

The stock market never stops changing. There are no rules capable of working well for ever. For instance, the performance of many stock selection rules developed by the economic experts before the 1990s, such as Graham's rules, has decreased dramatically when they are applied to the current stock market. When the training data sets formed from different years in the 1990s were applied, our constructed decision tree models came up with different prediction rules. Therefore, as time passes, the old rules need to be updated and new rules extraction should never stop.

References

1. E. Williams, *The McGraw-Hill Investor's Desk Reference* (McGraw-Hill, New York, 2000), pp. 413–453.
2. G. H. John and P. Miller, Building long/short portfolios using rule induction, in *The IEEE Conf. on Computational Intelligence in Financial Engineering*, Piscataway, New York (1996), pp. 134–140.
3. S. J. Hong, Use of contextual information for feature ranking and discretization, *IEEE Trans. Knowledge Data Eng.* **9**(5) (1997) 718–730.
4. C. Apte and S. J. Hong, Predicting equity returns from securities data, in *Advances in Knowledge Discovery and Data Mining* (AAAI Press, 1994), pp. 541–560.
5. S. J. Hong, R-MINI: an interative approach for generating minimal rules from examples, *IEEE Trans. Knowledge Data Eng.* **9** (1997) 709–717.
6. J. R. Quinlan, *C4.5: Programs for Machine Learning* (Morgan Kaufmann, 1993).
7. J. R. Quinlan, Induction of decision trees, *Machine Learning* **1**(1) (1986) 81–106.
8. J. R. Quinlan, Generating production rules from decision tree, in *Proc. 10th International Joint Conf. Artificial Intelligence*, 1987, pp. 304–307.
9. J. R. Quinlan, Improved use of continuous attribute in C4.5, *J. Artificial Intelli.* **4** (1996) 77–90.

10. S. Ruggieri, Efficient C4.5, in *Technical Report: TR-00-01*, Department Di Informatica, University Di Pisa, 2000.
11. D. J. E. Baestaens, W. M. van den Berh and H. Vaudrey, Market inefficiencies, technical trading and neural networks, in *Forecasting Financial Markets, Financial Economics and Quantitative Analysis*, ed. C. Dunis (John Wiley & Sons, Chichester, UK, 1996), pp. 245–260.
12. A. N. Refenes, A. Zapranis and G. Francis, Stock performance modeling using neural networks: a comparative study with regression models, *Neural Networks* **7**(2) (1994) 375–388.
13. G. S. Swales, Jr. and Y. Yoon, Applying artificial neural networks to investment analysis, *Financial Anal. J.* **September–October** (1992) 78–80.
14. Y. Yoon, T. Guimaraes and G. Swales, Integrating artificial neural networks with rule-based expert systems, *Decision Support Systems* **11** (1994) 497–507.
15. R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification* (John Wiley & Sons, New York, 2000), pp. 398–399.
16. Research Insight (Standard & Poor's) http://www.lib.siu.edu/cgi-bin/encore2/list_ind?key=r.
17. S. M. Bragg, *Business Ratios and Formulas: A Comprehensive Guide* (John Wiley & Sons, New York, 2002).
18. P. Blustein, Ben Graham's last will and testament, *Forbes* **August 1** (1997) 43–45.
19. T. Hellstrom, Predicting a rank measure for stock returns, *Theory Stochastic Process.* **6**(22) (2000) 64–83.
20. K. Schierholt and C. H. Dagli, Stock market prediction using different neural network classification architectures, Computational Intelligence for Financial Engineering, 1996, *Proc. of the IEEE/IAFE 1996 Conf.*, 24–26 March 1996, pp. 72–78.
21. G. H. John, P. Miller and R. Kerber, Stock selection using recon, in *Neural Networks in Financial Engineering*, eds. Y. Abu-Mostafa, J. Moody, P. Refenes and A. Weigend (World Scientific, London, 1996), pp. 303–316.