International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:6, 2016

A Case-Based Reasoning-Decision Tree Hybrid System for Stock Selection

Yaojun Wang, Yaoqing Wang

***Abstract*—**Stock selection is an important decision-making problem. Many machine learning and data mining technologies are employed to build automatic stock-selection system. A profitable stock-selection system should consider the stock's investment value and the market timing. In this paper, we present a hybrid system including both engage for stock selection. This system uses a case-based reasoning (CBR) model to execute the stock classification, uses a decision-tree model to help with market timing and stock selection. The experiments show that the performance of this hybrid system is better than that of other techniques regarding to the classification accuracy, the average return and the Sharpe ratio.

***Keywords*—**Case-based reasoning, decision tree, stock selection, machine learning.

I. Introduction

STOCK selection is regarded as a challenging task in portfolio optimization. As the role of equities is becoming more important, selecting attractive stocks for the short- and long-term investment has been the most important decision. Therefore, a reliable tool in the selection process can be helpful to investors. An effective and efficient system gives investors the competitive edge over others as they can identify the stocks with good performance with minimum efforts. Both academics and practitioners have used trading strategies and rules based on fundamental and technical analysis for decision-making. Trading strategies can be transformed into computer languages to exploit the- processing power of the computer. This machine program significantly reduces the time and efforts in finding attractive stocks. Many papers have focused on fundamental indicators to learn about how they can affect the future earnings and the stock prices and developing an investment strategy based on fundamental indicators results in substantial abnormal returns [1]-[3]. Also, analysts' recommendations, stock-market rumors, and earnings surprises can lead to abnormal returns.

In the financial markets, fundamental analysis and technical analysis are the two primarily used stock analysis methods. Fundamental analysis refers to the analysis of various aspects, including macroeconomics, company's main business, competition facing the company's internal management level and so on. The technical analysis, on the other hand, analyzes the historical trading data of a stock and uses this data to predict the trend [4].

Yaojun Wang is with Institute of Computing Technology, Chinese Academy of Science; University of Chinese Academy of Sciences, Beijing, China (phone: +86 015210996629; e-mail: wangyaojun@ict.ac.cn).

Yaoqing Wang is with Zhejiang University of Science and Technology School of Economics and Management Hang Zhou, China (e-mail: wangyq1021@163.com).

As for the fundamental analysis, stock valuation based on fundamentals aims at making an estimate of the intrinsic value of the stock, based on the predictions of the future cash flows and profitability of the business. Evidences demonstrate that ratio of measures of fundamental value to market value systematically predicts profits from stock investment. These indicators are calculated on the basis of the data about the company's fundamental to estimate the "intrinsic" value of the stock. They range from simple ratios like earnings per share (EPS) to ratios based on more sophisticated valuation models such as Ohlson [5]. Many studies find that investment returns of stocks are relatively high in firm-years with relatively low values of cash-flow-to-price, earnings-to-price, book-to-market, and value-to- market ratios [6], [7].

In addition to fundamental analysis, market timing is also crucial to a profitable stock investment. The stock profits depend not only on the company's fundamentals but also on the timing of the transaction. Some researchers consider market timing to be reasonable in certain situations, such as an apparent bubble [8], [9]. However, because the market is a complex system that contains many factors, even in the situation that the company has splendid fundamentals, it remains difficult to make a big profit when the market as a whole has a downward trend. Proponents of the market timing deem it the theme of trading. They firmly believe that all profitable trades are inseparable from market timing, regardless of what kind of products they choose to trade [10]. We suppose that market trend is an important factor needing to be considered when we prepare to trade stocks and a strategy or system that will result in more profits in long term, compared with consistently investing over an extended period regardless of market trends. Based on these findings, we can develop a model for selecting winning stocks by taking into account the two aspects of fundamental analysis and market timing. There have been many studies using machine learning (ML) techniques in stock selection and stock-price prediction [11]-[20]. Most of these studies pay attention to the stock market index and the individual-stock forecast [13]-[20]. Recent studies on the stock selection showed encouraging results by using data-mining techniques such as rule induction, neural network, and the combination of classifiers [12], [14], [17], [21]. CBR technique is one of the popular methodologies in knowledge-based systems. It is a novel paradigm that solves a new problem by recalling and reusing specific knowledge from experience. Concurrently, it is already an established and powerful methodology for creative problem solving and has been used for developing a variety of applications. Due to its strengths, researchers have successfully applied CBR to many areas: Supply chain management and scheduling, bond rating, business failure prediction, business control system development, bankruptcy prediction and credit analysis, and stock market forecasting [22]-[27].

This study presents a hybrid stock selection system based on CBR model and decision tree model. In this system, a fundamental indicator is used as an input in the CBR model for stock preselection and technical indicators is used as a contribution in the decision tree algorithm for market timing. For effective stock selection, CBR model and decision tree models are trained separately by corresponding algorithms with fundamental data and technical indicators, and then hybrid used in stock selections process. Portfolio managers focus on long-term portfolio management. Therefore, they try to choose fundamentally strong stocks without trading time consideration. On the other hand, some researchers developed investment system based on fundamental and technical analyses and they always combine these two types of indicators into one model [28].

The primary contribution of this study is to show that CBR can be used for stock preselection. Another contribution is that we separated the fundamental and technical indicators.

The rest of this paper is organized as follows: Section II briefly describes the related work. Section III describes the CBR model and decision tree algorithm. Section IV presents the proposed stock selection model. Section V gives results. The last part presents the conclusion.

II. Related Work

In the past few decades, various financial researchers have developed many conventional numeric forecasting models [29]. According to Ballings et al., common methods are the autoregressive method, the moving average model, the auto regressive moving average model [30] and the threshold autoregressive model, while Engle's autoregressive conditional heteroscedasticity model [31], Bollerslev's generalized ARCH model [32] and Box and Jenkins' autoregressive integrated moving average model [33], being some of the most cited. However, the stock market is a complex system with noisy, non-stationary and chaotic data series [34]. Therefore, an automatic decision-making systems that employ a soft computing approach has been supposed by many researchers in recently tried to predict stock market movements [11]. These systems include fuzzy logic and neural networks for the advantages that these models are tolerant of imprecision, uncertainty and approximation. To surpass this, other researchers have been interested in predicting the future price of a stock market index, stock, or technical indicator. Hafezi et al. adopted an Artificial Neural Network multi-agent system with four layers, by applying both fundamental and technological data using as a case study the DAX stock market and compared the outcomes with the results of other methods [12]. Kao et al. presented a hybrid approach by integrating wavelet-based feature extraction with multivariate adaptive regression splines and support vector regression, and the model outperformed five extra competing approaches [13]. Nair et al. also adopted a hybrid approach to a Genetic Algorithm tuning the parameters of an ANN at the end of each trading session [14]. Svalina et al. used an ANFIS for the Zagreb Stock Exchange with a separate fuzzy inference system for every day while every daily input variable was differently created from previous closing prices and the output was the closing price five days in advance [15]. Ticknor integrating the ANN architecture with the fuzzy time series, while Chang et al. integrated an ANFIS with a fuzzy time series model to predict the stock market [16].

Another way for the prediction of future prices is through technical pattern recognition, although Zapranis and Tsinaslanidis claim that is only a small fraction of the existing literature [17]. The advantages of being able to predict future prices are evident. Thus, this is one of the reasons that there is such active research motivation in this area. However, all previous studies did not contain any mechanism to propose the day or price where the investor should enter or exit the market. Thus, various researchers tried to predict the actual buy/sell signals.

Pereira and Tettamanzi used a plethora of technical indicators and an evolutionary algorithm to create a fuzzy predictive model, to produce a go shortly, go long or do nothing trading signal [18]. In contrast, Dourra and Siyused only three technical indicators for their method, which maps these ratios as inputs that can be fed into a fuzzy logic system to facilitate the decision to buy/sell a stock, when certain price movements or price formations occur [19]. Further, Ijegwa et al. used four technical indicators as inputs, and they developed a similar fuzzy system which outputs a buy, sell or hold signal, with satisfactory results [20]. However, every system should employ technical indicators with similar characteristics to strengthen the output and avoid their contradiction with each other. Therefore, the idea to use "as many as possible" indicators, might end up with a system which produces inconsistent signals. If the evaluation period is short, the signals might seem satisfactory although the results might be the opposite in another or a longer period. Buy/sell signals are produced by other researchers as well although their approaches differ.

The appropriate compilation of technical analysis with soft computing techniques in the area of stock market prediction can offer significant advantages since it can combine the hardly gained by traders experience with the unique characteristics of soft computing to face the problems of complexity, uncertainty and nonlinearity of the markets. The main disadvantage of fuzzy systems is that the knowledge about a problem must be known in advance, for a qualitative fuzzy inference system to be defined [15].

Unlike many other soft computing techniques, CBR systems avoid the reliance only on quantitative data, since they require decisions which are taken using subjective assessments. Through querying the case database and compare with all the cases to get the most profit stocks, CBR work in a more natural way, which represents in a better way the decision-making process during the stock selecting. This can be reverted to a significant advantage, since the knowledge, experience and intuition of an expert trader, which are easily handled by CBR systems, can be ensured in advance.

III. CBR and Decision Tree

A. CBR Approach

The CBR technique is one of the popular methodologies in knowledge-based systems. While other artificial intelligence techniques depend on generalized relationships between problem descriptors and conclusions, CBR utilizes specific knowledge of previously experienced problem situations. It solves a problem by retrieving the past case database and then reusing the solution which has the most comparable situation with the current case. Revising and retaining previous cases based on their degree of match and similarity to the current case. This is done by partial matching of the past cases with the current situation, and by ranking across case dimensions until a smaller set of matching and useful cases are retrieved [35]. The usefulness of former cases for the current situation, on the other hand, may be assessed by assigning weights. A high degree of similarity presents a good reason for adaptation. CBR methodology has been used in a broad range of domains to capture and organize experience and to learn how to deal with the new case from previous solutions. In general, a CBR system can be viewed as a composition of two modules, i.e., a case library and a problem solver [36]. The case library, which contains historical problems and their corresponding solutions, acts as a source of knowledge. Given a new problem, the problem solver performs two actions, i.e., (i) retrieves similar cases from the case library based on some similarity measure; and (ii) adapts the retrieved cases so that a solution to the new problem can be proposed. CBR system is composed of four sequential steps which are called into action each time that the new issue is to be solved [35], [36]. CBR involves four major steps which are a Retrieve, Reuse, Revise and Retain.

The aim of the retrieval step is to search the case library to select existing cases sharing significant features with the new case. The key issues in this step are computing case similarity to match the best case, and adopting a similar solution to fit the new problem. Thus, the success of a CBR system largely depends on an efficient retrieval of useful prior cases for the problem. The nearest neighbor method has been widely used for case retrieval. The method involves the assessment of similarities between stored cases and the current input case, based on matching a weighted sum of features. Once one or more cases are identified in the case base as being very similar to the new problem, they are selected for the solution of this particular challenge.

The CBR system tries to reuse the information and knowledge of the previously retrieved cases for solving the new issue. Once matching cases are retrieved from the case base, they should be adapted to the requirements of the current case. This process is called the revision process for CBR. This solution is revised (if possible) and finally the new case is stored. Cases can also be deleted if they prove to be inaccurate; they can be merged to create more generalized ones, and they can be modified. In the terminal step, the new solution is retained as part of a new case likely to be useful for future problem solving [37].

Efficiency and accuracy of case retrieval highly depend on the determination of weight for each feature. In many cases, subjective weighting values are given by the user, and thus, the retrieved solutions cannot always be guaranteed. Therefore, several case indexing methods have been proposed for effective case retrieval [38]-[40]. These are the nearest neighbor, induction, fuzzy logic, rough set theory, kernel methods and database technology. Nearest neighbor is the most commonly used case indexing method. It is a direct method that uses a numerical function to compute the degree of similarity. In this study, we use a numerical evaluation function which measures the distance taking into account the importance of features to calculate the degree of match in retrieval. A typical numerical function is [38]:

$$DIS_{ab} = \sqrt{\sum_{i=1}^{n} W_i \cdot \left(f_{ai} - f_{bi} \right)^2} \qquad (1)$$

where DIS is the matching function using the Euclidian distance between cases, with the weight of the feature i, and n is the number of features. In nearest-neighbor, every feature in the input case is matched to its corresponding feature in the stored case, and the similarity of each pair is computed using the matching function given in (1). Then, based on the importance assigned to each feature, an aggregate match score is calculated. Cases are ordered according to their scores [39]. The importance of each field shows us how much attention to pay to the respective match. Although researchers suggest several ways of assigning the importance values such as knowledge of a human expert, statistical evaluation, ML techniques and fuzzy logics [39], it was difficult to tell a priori regarding which set of weights would be the most efficient to solve a specific problem. According to [40], one way is to have a human expert assign the importance values as the case library. To the expert, it is required to have the knowledge and experience to decide which dimensions make good predictors. Another way to assign weight is to do a statistical evaluation of known cases to determine which dimensions predict the solutions best. The correlation coefficient between each input case and the one stored cases as an indicator to evaluate each input when computing the distance measure for a new example. ML can be used as an alternative approach to learning the optimal weights from historical cases using evolutionary search technique. By evaluating the fitness of diverse weight vectors, good solutions can be found for CBR system.

B. Decision Tree Approach

A decision tree is a basic classification and regression methods. The decision tree is a set of decision rules organized with a tree structure. In classification problems, an instance of decision represents a procedure based on the feature categorize. It can be considered a collection of If-then rules, also can be regarded as a conditional probability defined in the feature space and spatial distribution of class. Its main advantage is easy to interpret and process with high classification speed. Decision Tree Learning typically comprises three steps: feature selection, decision tree generation and pruning the decision tree. The learning ideas mainly from the CART algorithm ID3 algorithm and C4.5 algorithm [41], [42].

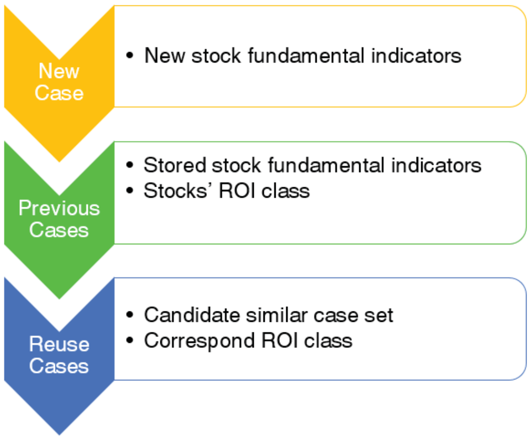

Fig. 1 CBR stock pre-select model overview

TABLE I
COMMONLY USED FUNDAMENTAL ANALYSIS INDICATORS [43]

Feature Name	Description	Formula*
Earnings	Simply put, earnings are how much profit (or loss) a company has made after subtracting expenses.	
EPS	EPS used to make earnings comparisons across companies.	Net earnings / number of outstanding shares of stock
P/E	The price-to-earnings ratio, more commonly known as the P/E ratio, to figure out how much the market is willing to pay for a company's earnings.	Price per share/EPS
PEG	A ratio that considers a stock's projected earnings growth.	P/E ratio / expected percentage earnings growth for the next year
Dividend Yield	The dividend yield measures the percentage of return which a company pays out to its shareholders in the form of dividends.	Amount of dividends paid per share/ stock's price
Book Value	The book value of a company is the company's net worth, as measured by its total assets minus its total liabilities.	
Price / Book	This measure is known as the price-to-sales ratio (P/S or PSR).	Stock's current price/total sales per share past year
ROE	Return on equity (ROE) shows you how much profit a company generates in comparison to its book value.	after-tax income/book value

*Formula units Net earnings and number of outstanding shares of stock is counted in years

The essential of decision tree learning is to sum up a set of classification rules from the training dataset. None or more than one trees can be obtained from the training dataset which not in contradiction with the training data. The needed decision tree is that has less contradiction with the training data but has good generalization ability. From another perspective, decision tree learning by training data set is a process of estimate a conditional probability model. There is an infinite number of conditional probability model for the classification based on feature space divide. We choose these models that not only have a good fit with the training data but also are good predictors to unknown data.

Loss function has been introduced in the decision tree learning model; the learning strategy can come down to the process of minimizing the loss function. When the loss function is determined, the learning problem becomes an optimization procedure at the expense of function significance. Select the best tree from all possible decision tree is an NP-complete problem, so in practically, decision tree learning algorithms typically employ heuristic methods to approximately solving this optimization problem. Thus, the training result is always a sub-optimal one.

The heuristic methods can be seen as recursively select the optimal features from the training data and reduced the classification problem to the best sub-classification process. This process corresponds to the division of the feature space, also corresponds decision tree construction. Specific steps are as follows:

(1) At the beginning of decision tree construct, all the training data are placed in the root node.
(2) Select an optimal feature, depending on the feature's value the elements of training data will be divided into subsets such that each subset been the best classified in current conditions.
(3) If a subset of these have been classified correct, and all elements can be assign to one class, then construct leaf node and assigned to the subsets corresponding to the leaf node; else, this subset continues to be divided through selecting the new best distinguishing feature and create a branch node for this subgroup.
(4) Do above step recursively until all elements of training data can be correct classified. Finally, each subset is assigned to the child node that has a clear classification.

The method of generating the above decision tree may have an excellent classification ability on training data but to the unknown data may not have a proper classification capability as overfitting. We need to use bottom-up tree pruning technology to simplify the tree structure so that it has better generalization ability. Specifically, it is to remove too finely divided leaf node.

IV. HYBRID APPROACHES

A. Stock Pre-Select Model Based on Fundamental Indicators

Fundamental analysis is a method used to evaluate the value of stock by analyzing the corresponding company's financial data. Moreover, in general, fundamental analysis takes into consideration only those variables that are directly related to business operations, such as its earnings, its dividends, and its sales. It focuses exclusively on the company's business to determine whether or not the stock should be bought or sold.

Fundamental analysis is useful to evaluate the investment value of the company's stock. A company's financial statements (its income statement, its balance sheet and its cash flow statement) will be indispensable resources for company analysis. In this study, we use nine fundamental indicators (See Table I) as the case features to build CBR model (Fig. 1).

All research data in this study are obtained from the Yahoo finance, and the adjusted stock price is used. The time series data set ranges from 1991-01-01 to 2016-01-01. All the fundamental indicators counted based on the open data from the listed company's annual report. We calculated stock return on investment (ROI) for every two natural year cycle. If the performance of a stock is greater or equal to 20%, then it is labeled as a positive class (+1) and the others are labeled as negative class (−1). This stock selection problem is then formulated as a two-class pattern recognition task. We represent the fundamental indicators for the i-th firm as a vector of the predictor variable $S_i = (x1, x2, ..., xn)$ (for our case n = 9). The expected future return of the stock is a binary dependent variable yi= ±l, where +1 represents high return stocks and −1 as normal stocks. Therefore, a training set (x, y) of n firms will be the following pairs:

$$\{(x1, y1), (x2, y2), ..., (xn, yn)\}$$

Several researchers have suggested using genetic algorithms (GAs) to determine the most appropriate feature weights of CBR approach [44], [45]. According to Moon and Sohn, the GA is a popular used algorithm to optimize the results by constantly enumerate various possible solutions with the reproduction, crossover and mutation operations [46].

The objective of GAs is to determine the set of weighting values that can best formalize the match between the input case and the previously stored cases. In other words, our objective is to retrieve more relevant cases that can lead to the correct solution. This can be achieved by increasing the classification accuracy. In this study, the fitness function is defined as the classification accuracy of the training set. The fitness function is expressed as:

$$F = \frac{1}{n}\sum_{i=1}^{n} CR_i \quad (i = 1, 2, \ldots n) \qquad (2)$$

B. Stock Select Model Based on Technical Indicators

Technical analysis is also necessary for a stock transaction; Technical indicators are used as a measurement to evaluate the supply and demand of securities within technical analysis. Those indicators (see Table II) confirm price movement and the probability that the move will continue. The direction prediction of a given asset's momentum is the goal of every short-term trader and to attempt to profit from it. For long-term investors, the timing of the transaction is also crucial. When the market trend in a downward, the buy-up will be trading losses in a high probability, even if the stock's fundamentals are stunning. Here, we devise a decision tree algorithm to select the time of the transaction. The entire process consists of the following four steps:

(1) Grab all the data we needed and calculate the indicators.
(2) Then calculate the variable which was looking to predict.
(3) Build a tree using the training data (indicators and class).
(4) Prune the tree that has already been built to minimize the cross-validated error.

A quick note on interpreting the tree (see Fig. 2). The nodes in the graph represent a split, with the left branch reflecting a "yes" answer and the right branch represent a "no" answer. The number in the final "leaf" is the percentage of instances that were corrected classified by that node

We use an example contains 1000 cases to test the efficiency of the decision tree model. From Table III, it can be found that the predictive accuracy can be up to 70%

C. The Hybrid Stock Selection System

The entire stock selection system consists of two models: one is CBR model, and the other is Decision Tree model. Firstly, the CBR model evaluates fundamental indicators of all stocks of the market and output with potentially high-yield stocks as a preselected stock set. Secondly, input this stock set, together with its corresponding technical indicators, into the Decision Tree model; The output stock set from the Decision Tree model is the final result of this selection system (Fig. 3).

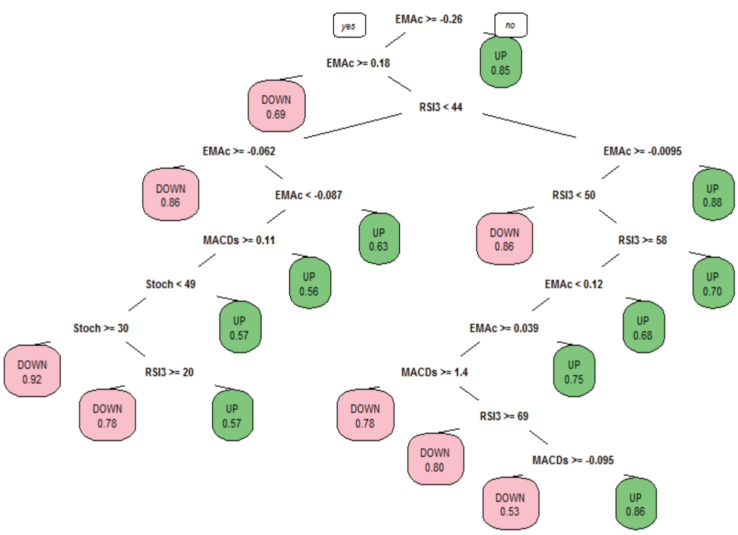

Fig. 2 Part of initial decision tree

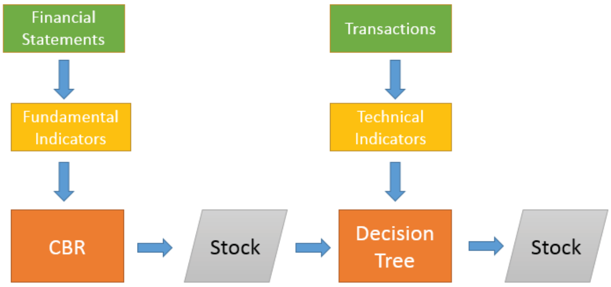

Fig. 3 The hybrid stock selection system overview

V. Experiment Results

The dataset used in this paper is the stocks that were traded on NASDAQ. The dataset contains 1200 observations. Stock prices are used to calculate the return on investment (ROI). Given the stock prices and indicators, it is a prediction problem that involves discovering useful patterns in the dataset and applying that information to classify the stocks.

To evaluate the effectiveness of the CBR-decision tree hybrid system (**CBR_DTree**), We set up three groups of experimental control group model:

- **CBR:** a single CBR stock selection model which included all the indicators CBR_DTree's CBR module has.
- **DTree:** a single decision tree based stock selection model which included all the indicators CBR_DTree's DTree module have.
- **DTree_CBR:** A hybrid system, but first use decision tree model and fundamental indicators to build pre-select model and then CBR model for timing.

We compare the return generated by the selected stock from CBR_DTree with other three models CBR, DTree, DTree and CBR.

We have used the Sharpe ratio, average return and ideal profit ratio for out-of-sample (test sample) comparison. The Sharpe ratio can be defined as the mean return of the trading strategy by its standard deviation. In other words, The Sharpe ratio measures return to the risk taken, and higher positive values are preferred. A Higher value of the Sharpe ratio ratio measures the return of the corresponding method against a perfect predictor. The range of the ideal ratio is $[-1, 1]$, yet a positive value is desirable. Table IV presents the Sharpe ratio, ideal profit ratio and average return values for each technique. CBR_DTree approach has the highest average return while DTree method has the lowest average return. Ideal profit ratio indicates that CBR_DTree approach is better than the other techniques. The Sharpe ratio values for each technique are positive and the highest value is given for CBR_DTree approach. Based on these three tests, we can conclude that CBR_DTree approach outperforms the other techniques for long-term stock selection.

VI. Conclusions

Stock selection problem is important and widely studied topic due to its significant impact on the profitability of portfolio. This paper presents a new hybrid approach work tandemly by CBR technique and decision tree model, namely CBR_DTree, which is an efficient and robust artificial intelligent (AI) paradigm for problem-solving. Experimental results showed that CBR_DTree was significantly better than other models regarding stock selection. We have shown that this approach increases overall classification accuracy rate significantly.

TABLE II
Commonly Used Technical Analysis Indicators [47]

Feature Name*	Description
OBV	The on-balance volume indicator (OBV) is used to measure the positive and negative flow of volume in security, about its price over time.
A/D	One of the most commonly used indicators to determine the money flow of security is the accumulation/distribution line (A/D line).
ADX	The average directional index (ADX) is a trend indicator used to measure the strength and momentum of an existing trend.
MACD	The moving average is used to signal both the trend and momentum behind a security.
RSI	The relative strength index (RSI) is used to signal overbought and oversold conditions in security.
EMA	An exponential moving average (EMA) is a type of infinite impulse response filter that applies weighting factors which decrease exponentially.

*Feature Name is all in abbreviation form

TABLE III
Results of Test Set on Decision Tree Model

Indicators	Methods			
	CBR_DTree	CBR	DTree	DTree_CBR
Average return	0.23	0.16	0.09	0.13
Sharpe ratio	0.51	0.30	0.26	0.38
Ideal profit	0.62	0.43	0.17	0.39

*Methods: include 4 models

TABLE IV
Performance Evaluation of Stock Selection Models

	Actual	
Predicted	Down	UP
Down	386	163
Up	126	325

*Actual: actual data labeled class; Predicted: the predicted class.

References

[1] Zhou R T, Lai R N. Herding and information based trading (J). Journal of Empirical Finance, 2009, 16(3): 388-93.
[2] Da Z, Schaumburg E. Relative valuation and analyst target price forecasts (J). Journal of Financial Markets, 2011, 14(1): 161-92.
[3] Atsalakis G S, Valavanis K P. Forecasting stock market short-term trends using a neuro-fuzzy based methodology (J). Expert Syst Appl, 2009, 36(7): 10696-707.
[4] Hanssens D M, Parsons L J, Schultz R L. Market response models: Econometric and time series analysis (M). Springer Science & Business Media, 2003.
[5] Dechow P M, Hutton A P, Meulbroek L, et al. Short-sellers, fundamental analysis, and stock returns (J). Journal of Financial Economics, 2001, 61(1): 77-106.
[6] Lau, Sie Ting, Chee Tong Lee, and Thomas H. McInish. "Stock returns and beta, firms size, E/P, CF/P, book-to-market, and sales growth: evidence from Singapore and Malaysia." Journal of multinational financial management 12.3 (2002): 207-222.
[7] Chan, Louis KC, and Josef Lakonishok. "Value and growth investing: Review and update." Financial Analysts Journal 60.1 (2004): 71-86.
[8] Graham, John R., and Campbell R. Harvey. "Market timing ability and volatility implied in investment newsletters' asset allocation

recommendations." Journal of Financial Economics 42.3 (1996): 397-421.
[9] Henriksson, Roy D. "Market timing and mutual fund performance: An empirical investigation." Journal of business (1984): 73-96.
[10] Thompson G. Markets, hierarchies and networks: the coordination of social life (M). Sage, 1991.
[11] Bao D, Yang Z. Intelligent stock trading system by turning point confirming and probabilistic reasoning (J). Expert Syst Appl, 2008, 34(1): 620-7.
[12] Hafezi R, Shahrabi J, Hadavandi E. A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price (J). Applied Soft Computing, 2015, 29(196-210.
[13] Kao L-J, Chiu C-C, Lu C-J, et al. A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting (J). Decision Support Systems, 2013, 54(3): 1228-44.
[14] Nair B B, Sai S G, Naveen A, et al. A GA-artificial neural network hybrid system for financial time series forecasting (M). Information Technology and Mobile Communication. Springer. 2011: 499-506.
[15] Svalina I, Galzina V, Lujić R, et al. An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices (J). Expert Syst Appl, 2013, 40(15): 6055-63.
[16] Chang P-C, Liu C-H. A TSK type fuzzy rule based system for stock price prediction (J). Expert Syst Appl, 2008, 34(1): 135-44.
[17] Zapranis A, Tsinaslanidis P E. A novel, rule-based technical pattern identification mechanism: Identifying and evaluating saucers and resistant levels in the US stock market (J). Expert Syst Appl, 2012, 39(7): 6301-8.
[18] Da Costa Pereira C, Tettamanzi A G. Fuzzy-evolutionary modeling for single-position day trading (M). Natural computing in computational finance. Springer. 2008: 131-59.
[19] Dourra H, Siy P. Investment using technical analysis and fuzzy logic (J). Fuzzy sets and systems, 2002, 127(2): 221-40.
[20] Ijegwa A D, Rebecca V O, Olusegun F, et al. A predictive stock market technical analysis using fuzzy logic (J). Computer and Information Science, 2014, 7(3): 1.
[21] Han J, Kamber M, Pei J. Data mining: concepts and techniques (M). Elsevier, 2011.
[22] Kwon, Ohbyung, Ghi Paul Im, and Kun Chang Lee. "MACE-SCM: A multi-agent and case-based reasoning collaboration mechanism for supply chain management under supply and demand uncertainties." Expert Systems with Applications 33.3 (2007): 690-705.
[23] Kim, Kyung-Sup, and Ingoo Han. "The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases." Expert Systems with Applications 21.3 (2001): 147-156.
[24] Li, Hui, and Jie Sun. "Business failure prediction using hybrid 2 case-based reasoning (H 2 CBR)." Computers & Operations Research 37.1 (2010): 137-151.
[25] Allen, Bradley P. "Case-based reasoning: Business applications." Communications of the ACM 37.3 (1994): 40-43.
[26] Bryant, Stephanie M. "A case based reasoning approach to bankruptcy prediction modeling." Intelligent Systems in Accounting, Finance and Management 6.3 (1997): 195-214.
[27] Chun, Se-Hak, and Yoon-Joo Park. "Dynamic adaptive ensemble case-based reasoning: application to stock market prediction." Expert Systems with Applications 28.3 (2005): 435-443.
[28] Hall C A. Integrating Concepts and Models from Development Econoimcs with Land use Change in the Tropics (J). Environment, Development and Sustainability, 2006, 8(1): 19-53.
[29] Cheng C H, Chen T L, Wei L Y. A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting (J). Inform Sciences, 2010, 180(9): 1610-29.
[30] Ballings M, Van Den Poel D, Hespeels N, et al. Evaluating multiple classifiers for stock price direction prediction (J). Expert Syst Appl, 2015, 42(20): 7046-56.
[31] Engle R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United-Kingdom Inflation (J). Econometrica, 1982, 50(4): 987-1007.
[32] Bollerslev T. Generalized Autoregressive Conditional Heteroskedasticity (J). J Econometrics, 1986, 31(3): 307-27.
[33] George Box. Box and Jenkins: Time Series Analysis, Forecasting and Control (M)// A Very British Affair. Palgrave Macmillan UK, 2013.
[34] Ticknor J L. A Bayesian regularized artificial neural network for stock market forecasting (J). Expert Syst Appl, 2013, 40(14): 5501-6.
[35] Aamodt A, Plaza E. Case-based reasoning: Foundational issues, methodological variations, and system approaches (J). AI communications, 1994, 7(1): 39-59.
[36] Ahn H, Kim K-J. Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach (J). Applied Soft Computing, 2009, 9(2): 599-607.
[37] Yang B-S, Han T, Kim Y-S. Integration of ART-Kohonen neural network and case-based reasoning for intelligent fault diagnosis (J). Expert Syst Appl, 2004, 26(3): 387-95.
[38] Leigh W, Purvis R, Ragusa J M. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support (J). Decision Support Systems, 2002, 32(4): 361-77.
[39] ST Phane N, Hector R. Effective retrieval and new indexing method for case based reasoning: application in chemical process design (J). Engineering Applications of Artificial Intelligence, 2010, 23(6): 880-94.
[40] Shin K-S, Han I. A case-based approach using inductive indexing for corporate bond rating (J). Decision Support Systems, 2001, 32(1): 41-52.
[41] Quinlan J R. Induction of decision trees (J). Machine Learning, 1986, 1(1):81-106.
[42] Breiman L I, Friedman J H, Olshen R A, et al. Classification and Regression Trees (CART) (J). Lecture Notes in Computer Science, 1984, 40(3):17–23.
[43] Using Stock Fundamental Analysis to Value a Company. Web hyperlink: http://www.investorguide.com/article/11632/valuation-methods-used-in-fundamental-analysis-of-stocks-igu/ . 2016.4
[44] Beddoe G, Petrovic S. Determining feature weights using a genetic algorithm in a case-based reasoning approach to personnel rostering (J). 2004.
[45] Jarmulak J, Craw S. Genetic algorithms for feature selection and weighting (J). Proceedings of the Ijcai'99 Workshop on Automating the Construction of Case Based Reasoners, 1999:28--33.
[46] Sohn S Y, Moon T H, Seok K J. Optimal pricing for mobile manufacturers in competitive market using genetic algorithm (J). Expert Systems with Applications, 2009, 36(2):3448-3453.
[47] Top 7 Technical Analysis Tools. Web hyperlink: http://www.investopedia.com/slide-show/tools-of-the-trade/ 2016.4