

A Hybrid Competitive Genetic Algorithm Models in Stock Market Prediction

M.P. RAJAKUMAR

Research Scholar, Sathyabama University,
Associate professor, St. Joseph's College of Engg,
Chennai-600119, Tamil Nadu, India.
E-mail: rajakumar_mp@yahoo.com

Dr. V. SHANTHI

Professor, St. Joseph's College of Engineering,
Chennai-600119, Tamil Nadu, India
E-mail: drvshanthi@yahoo.co.in

Abstract— Stock market prediction (SMP) plays an important role in the modern era for any economy which is on the development phase. Genetic Algorithm (GA) is an evolutionary algorithm that is useful for solving problems which are too complex in nature. This paper surveys different GA models that have been experimented in stock market prediction with special enhancement techniques used with them to improve the prediction accuracy. The classification is made in terms of GA with two layer, multilayer, neural network variants and modified evolutionary algorithms. Through the surveyed paper it is shown that the performance of GA excels when integrated with other machine learning algorithms.

Keywords-classification, enhancement technique, genetic algorithm, prediction accuracy

I. INTRODUCTION

Stock markets have been one of the major players in determining the economy which is the backbone of any nation and for that matter even the global economy. There is always some risk involved in investing in the stock market due to its highly unpredictable behavior. Since stock market is essentially non-parametric, dynamic, time-variant and chaotic in nature, stock market prediction is a challenging task [1]. In addition to this stock market is affected by many macro economical factors such as political events, general economic conditions, firm's policies, investors' expectations, institutional investors' choices, movement of other stock markets and the psychology of investors etc [2].

GA plays an important role in predicting the stock market prices accurately. Numerous researches on the application of GA in forecasting problems have proven its advantages over other statistical and parametric techniques. GA has the advantages like easy to understand, easy to implement, highly configurable, does not suffer from the curse of dimensionality, flexible, applicable to any problem where a fitness function is defined, easily parallelizable, output can be turned into an actual program and can be run to speed in real time, and can be used to optimize more formal models. GA also inherits certain limitations like stagnation of population, rapid predominance of certain individuals over the rest of the population, over-fitting, provide the most fit solution that was evolved not necessarily the optimal solution, hard to determine

optimal parameters. The limitations of traditional GA can be overpowered by hybrid layered GA in predicting the stock market behavior which inculcates the features like two layered GA, multilayered GA, neural network variants with GA & modified GA. These additional advantages of GA provide forecasting of stock market prices at a speedier rate with high prediction accuracy compared to traditional GA.

The purpose of this work is to review and classify the hybrid GA models to stock market prediction. The results are presented in four different categories. The first category lists two layer hybrid GA model, description and their model comparisons. The second lists three layer hybrid GA model for stock market prediction. The third presents stock markets predictions by GA with neural network variants. The last one summarizes the modified versions of simple genetic algorithms for stock market predictions.

This work focuses on the application of available GA to predict stock market indexes. GA may be applied to diverse markets to forecast the stock market indexes.

II. COMPARISON OF HYBRID GA

Improving SMP by two layer hybrid GA

In [3], a hybrid neurogenetic system for stock forecasting was proposed. This hybrid system used RNN trained by a back propagation based algorithm, to predict buy and hold strategy of 36 companies in NYSE & NASDAQ for 13 years from 1992 to 2004. Since back propagation algorithm is prone to get stuck in local minima and highly depends on the initial weights, the GA was used to optimize the NN's weight under a 2-D encoding and crossover mechanisms. The GA is parallelized on a Linux cluster system using message passing interface to reduce the time in processing the mass data. The experimental results showed that the hybrid model predicts better (34 companies among 36 companies) than both GA & SVM alone in buy and hold strategy.

In [4], a hybrid GA-SVM system was developed for predicting the future directions of the stock price. A set of technical indicators which exhibits high correlation were used as input features. The GA was used to select the set of most informative input features among all the technical indicators. The selected features were used as inputs to SVM. SVM

implemented the structured risk minimization problem, to minimize an upper bound of generalization error. The prediction performance was measured based on the hit ratio. The results were compared with the stand alone SVM. The GA-SVM hybrid model significantly outperformed the SVM.

A hybrid machine learning system based on GA-TSA model for stock market forecasting was developed [5]. The success of a trading rate depends on choosing the best parameter combination. This can be done by using GA. The GA sets the sub domain of the parameters and finds near optimal value in the sub domain with the time series analysis in a very reasonable time. This hybrid model outperformed the time model alone in terms of accuracy. The accuracy can be still improved by considering various political & economical factors that affects the stock market and incorporating market specific domain knowledge in to the system.

The stock market indices were predicted using hybrid GA-PSO with perturbation term inspired by the passive congregation biological mechanism which enables all particles in the swarm to perform the global search in the whole search space [6]. The GA is used to increase the diversity and give particles to fly in the new regions in search space. This hybrid model increases the prediction accuracy for both short term & long term stock market indices compared to GA/PSO model alone. This model can be used in other applications such as pattern classification recognition and optimization problem.

A novel method based on hybrid combination of GA-ARMA was applied for time series forecasting [7]. This hybrid model benefits from the capabilities of ARMA to identify linear trends as well as GA's ability to obtain models that capture on non linear patterns from data. The GA evolves models adopting the data without any restrictions with respect to the form of models or coefficients. The empirical result with real stock data confirmed that the hybrid approach to be a fair competitor over pure ARMA & GA. The accuracy can be still improved by combining change point detection method & modeling techniques.

In [8], a hybrid model based on GA-RST predicted the stock price. In this novel method multi-technical indicators were used to predict stock price trends. This method has employed a RST algorithm to extract linguistic rules from the linguistic technical indicator dataset and utilize GA to define the extracted rules to get better forecasting accuracy and stock return. The effectiveness of the proposed model was verified with two types of performance evaluations, accuracy and stock return for six year period of the Taiwan stock exchange capitalization weighted stock index as the experiment dataset.

b.Improving SMP by three layer hybrid GA

An intelligent decision support system which measures all the qualitative events in addition to quantitative factors that may influence the stock market were developed [9]. This novel method consists of 3 parts namely factors identification, qualitative model and decision integration. The fuzzy Delphi method was employed to capture the stock expert's knowledge and transform it to the acceptable format of GFNN. The ANN

which considers only quantitative factors is outperformed by the proposed system in the learning accuracy, buy-sell clarity and buy-sell performance. Real-number coding approach can also be applied in addition to binary coding approach.

A fusion model by combining HMM, ANN and GA to forecast financial behavior was developed [10]. These can be used for in-depth analysis of the stock market. Using the ANN the daily stock prices were transformed to independent sets of value. The GA is used to find out the optimal initial parameters for the HMM given the transformed observation sequences. This fusion model found a number of alternative data items from the historical data, which exhibit similar stock market trends. The trained HMM is used to identify and locate similar patterns in the historical data. The performance of the fusion model is better than the basic model where only a single HMM is used. It also outperforms the popular statistical forecasting tool.

The hybrid VAR-NN-GA framework has automated the decision process of prediction [11]. VAR searched for the correlated stocks and indicators automatically. Input data such as trading volume, economic growth rate and the currency exchange rate tested with VAR analysis. The NN forecast from the relevant inputs was decided by the VAR analysis. The GA is used to adjust the weights of each NN model. The VAR-NN-GA hybrid system outperformed the stand-alone neural network in prediction accuracy.

A novel approach GA-CPSO-SVR was employed to predict the financial returns of Shanghai composite index [12]. GA is used for feature selection and CPSO is used for optimizing the parameter of SVR model. CPSO method combines PSO with adaptive inertia weight factor and chaotic linear search. The accuracy of the prediction could be improved by considering other advanced searching techniques to determine suitable parameters and the number of input data. This hybrid model outperforms BPNN, ARIMA, SVR and CPSO-SVR in forecasting financial returns.

c.Improving SMP by GA with neural network variants

TDNN & GA were used together to get more accurate forecast of European option prices of NIFTY index [13]. Since back propagation algorithm is prone to get stuck in local minima and specifying the architecture at the start of the algorithm may not be optimal, the GA is used to optimize the TDNN. This approach has fixed architecture and training of TDNN was done with GA. The experimental results showed that the hybrid model predicts better than simple feed forward neural network.

A GMDHNN and genetic algorithm was developed for stock prediction of cement sector in Tehran stock exchange [14]. Using GMDHNN a model could be represented as a set of neurons in which different pairs of them in each layer were connected through a quadratic polynomial and therefore produce new neurons in the next layer. In GMDHNN the identification process used the suitable optimization method GA to find out the best network architecture. The whole architecture of the GMDHNN is designed with the help of

GA. It provides the optimal number of neurons in each hidden layer and their connectivity configuration to find the optimal set of appropriate coefficients of quadratic expressions to model stock prices. GMDHNN outperforms traditional time series method and regression based models in prediction accuracy.

In [15], the author has investigated the effectiveness of a hybrid approach based on the ATNN and the TDNN with the GA in a temporal pattern for stock market prediction tasks. TDNN permits only weights to be adopted but time delays are included but are fixed. The ATNN network adapts its time delay and weights during training to better accommodate changing temporal patterns and provide more flexibility for optimization tasks. GA supports optimization of the number of time delay and network architectural factors simultaneously for the ATNN and TDNN model. The result showed that the accuracy of this approach is higher than that of the standard ATNN, TDNN and RNN.

d. Improving SMP by modified GA

TIMTAEF method, which performs an evolutionary search for the minimum dimension in determining the characteristic phase space that generates the financial time series was presented [16]. This hybrid model composed of modular morphological neural network with a modified genetic algorithm (MGA). The modified genetic algorithm was used to improve search convergence. The prediction of the proposed model obtained a performance much better in terms of evaluation function than the TAEF & MRLTAEF model

To increase prediction accuracy and reduce search space and time for achieving the optimal solution, the combination of WNN with fuzzy knowledge was used [17]. The proposed FWNN structure is trained with differential evolution (DE) algorithm. DE includes evolution strategies and conventional GA. DE is used for minimizing non-linear and non-differentiable continuous space function. Training FWNN system by DE is much more faster than the traditional GA. The result, demonstrated that FWNN with DE has better performance than FWNN with BP, FFNN and ANFIS.

In [18], the author has investigated the chaotic behavior of NASDAQ and S&P CNX NIFTY stock markets using EDA based LLWNN. The LLWNN system was optimized using estimation of distribution algorithm which is a new class of evolutionary algorithms. EDA explicitly extract global statistical information from the selected solutions. The major issues in EDAs are selecting parents and building probability distribution model. The author used truncation selection and gaussian distribution with diagonal covariance matrix. The result showed that the accuracy of this approach is higher than that of the standard WNN. Table I lists the main features of hybrid GA.

An evolutionary neural network model was proposed to improve the performance of ANN in time series forecasting [19]. A modified GA was used to select best input features, optimize the slope of hidden nodes activation function, learning parameters and also the number of hidden layer

nodes. To evaluate the effectiveness of the proposed model FOREX rate prediction acts as a benchmark application. The result showed that the accuracy of this approach is higher than that of hybrid RNN-ARIMA, fuzzy system-ANN and ANN.

TABLE I. MAIN FEATURES OF HYBRID GA

Hybrid GA	Main features
GA-FS	reduce the processing time without loss of performance, , reduce computational cost using message passing interface, improves the generalizations, avoids the curse of dimensionality, avoids local minima
GA-SVM	resistance to over fitting problem, optimal selection of most informative input features from the technical indicator
GA-TSA	finds best parameter combination which is the key for successful trading
GA-PSO	fast convergence ,robustness
GA-ARMA	more robust, accurate model of time series
GA-RST	extract rules effectively from past stock data, classification accuracy and forecasts profit, produces more reasonable & understandable rules, provide objective suggestions
GA-FS-NN	Captures the stock expert's knowledge, decreases the training time and avoids local minima
HMM-ANN-GA	This model can be used without analyzing the dataset prior to the forecast
VAR-NN-GA	More robust, able to select variables automatically, dynamic model selection process, multiple variables can be investigated at the same time, avoids local minima
SVM –GA –CPSO	deals with features selection and SVM parameter optimization simultaneously.
TDNN-GA	adopt the network online, finds the optimum network architecture, adapts time delay to provide more flexibility
GMDHNN-GA	best optimal simplified model for inaccurate, noisy or small data sets, simple structure than traditional neural network models, higher accuracy, provides the optimal number of neurons in each hidden layer and their connectivity configuration
ATNN & TDNN-GA	supports the optimization of the number of time delays & network architectural factors simultaneously, adapts both time-delay and weights to accommodate changing temporal patterns and provides more flexibility
MGA based MMNN	Overcomes the random walk dilemma for stock market prediction, does not discard any possible correlation that exist among the time series parameters, even higher order correlations does not make any prior assumption
DE based FWNN	Combines the strengths of wavelet theory, fuzzy logic and neural networks, fast training , ability to analyze non-stationary signals to discover their local details, self learning characteristic that increases the accuracy of the prediction, less time for parameter updating
EDA based LLWNN	Learning efficiency, structure efficiency, provides more parsimonious interpolation in high dimension spaces when modeling samples are sparse
MGA based evolutionary NN	Selects best input features, optimizes the slope of the hidden nodes activation function, learning parameters and the number of hidden layer nodes.

III. CONCLUSION

This study has surveyed articles that have applied hybrid GA model to predict stock market values. This study has focused on the description of hybrid GA, model comparison and its enhanced features. Traditional GA has its own drawbacks thereby limiting its performance to a certain level of accuracy. The limitation of traditional GA can be overcome by hybrid layered GA in predicting the stock market behavior which inculcates features like two layer, multilayer & modified evolutionary algorithms.

APPENDIX: ABBREVIATION

ANFIS	Adaptive neuro fuzzy inference system
ANN	Artificial neural network
ARMA	Auto regressive moving average
ATNN	Adaptive time-delay neural network
CPSO	Chaotic particle swarm optimization
DE	Differential evolution
EDA	Estimation of distribution algorithms
FFNN	Feed forward neural network
FWNN	Fuzzy wavelet neural network
GFNN	Genetic fuzzy neural network
GMDHNN	Group method of data handling neural network
HMM	Hidden markov model
LLWNN	Local linear wavelet neural network
MGA	Modified genetic algorithm
MRLTAEF	morphological rank linear time delay added evolutionary forecasting
PSO	Particle swarm optimization
RNN	Recurrent neural network
RST	Rough set theory
SVM	Support vector machine
SVR	Support vector regression
TAEF	Time-delay added evolutionary forecasting
TDNN	Time delay neural network
TIMTAEF	Translation invariant morphological time-added evolutionary forecasting
TSA	Time series analysis
VAR	Vector auto regression
WNN	Wavelet neural network

REFERENCES

- [1] M. T. Leung, H. Daouk and A. S. Chen, "forecasting stock indices: A comparison of classification and level estimation models," *International Journal of Forecasting* 16, pp. 173-190, 2000.
- [2] K. Manish and M. Thenmozhi, "Forecasting stock index movement: A comparison of support vector machines and random forest," *Proceedings of Ninth Indian Institute of Capital Market Conference*, India, Jan 2006.
- [3] Y. Kwon and B. Moon, "A hybrid neurogenetic approach for stock forecasting," *IEEE Transactions on Neural Networks*, vol. 18, no. 3, pp. 851-864, May 2007.
- [4] R. Choudhry and K. Garg, "A hybrid neurogenetic approach for stock forecasting," *World Academy of Science, Engineering and Technology* 15, pp. 315-318, 2008.
- [5] L. Kumar, A. Pandey, S. Srivastava and M. Darhari, "A hybrid machine learning system for stock market forecasting," *Journal of International Technology and Information Management*, vol. 20, no. 1, pp. 39-47, Jan 2011.
- [6] T. Aboueldahab and M. Fakhreddin, "Prediction of stock market indices using hybrid genetic algorithm/particle swarm optimization with perturbation term," *International Conference on Swarm Intelligence*, June 2011.
- [7] A. Barbulescu and E. Bautu, "A hybrid approach for modeling financial time series," *The International Arab Journal of Information Technology*, vol. 9, no. 4, pp. 327-335, July 2012.
- [8] C. H. Cheng, T. L. Chen and L. Y. Wei, "A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting," *Information Sciences* 180, pp. 1610-1629, 2010.
- [9] R. J. Kuo, C. H. Chen and Y. C. Hwang, "An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network," *Fuzzy Sets and Systems*, vol. 118, pp. 21-45, 2001.
- [10] M. R. Hassan, B. Nath and M. Kirley, "A fusion model of HMM, ANN and GA for stock market forecasting," *Expert Systems with Applications*, vol. 33, pp. 171-180, 2007.
- [11] S. I. Ao, "A framework for neural network to make business forecasting with hybrid VAR and GA components," *Engineering Letters*, 13:1, pp. 24-29, May 2006.
- [12] Y. Fu and Y. Cheng, "Application of an integrated support vector regression method in prediction of financial returns," *International Journal of Information Engineering and Electronic Business*, vol. 3, no. 3, pp. 37-43, June 2011.
- [13] P. K. Mishra, P. Kumar, V. Sharma and J. Dhar, "Option pricing using hybrid neuro evolution approach," *International Research Journal of Finance and Economics*, issue 64, pp. 115-125, April 2011.
- [14] S. Fallahi, M. Shaverdi and V. Bashiri, "Applying GMDH-type neural network and genetic algorithm for stock price prediction of Iranian cement factor," *Applications and Applied Mathematics: An International Journal*, vol. 6, no. 2, pp. 572-591, Dec 2011.
- [15] H. Kim and K. Shin, "A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets," *Applied Soft Computing* 7, pp. 569-576, 2007.
- [16] R. A. Araujo, "Translation invariant morphological time-lag added evolutionary forecasting method for stock market prediction," *Expert Systems with Applications* 38, pp. 2835-2848, 2011.
- [17] R. H. Abiyev and V. H. Abiyev, "Differential evaluation learning of fuzzy wavelet neural networks for stock price predictions," *Journal of Information and Computing Science*, vol. 7, no. 2, pp. 121-130, 2012.
- [18] Y. Chen, X. Dong and Y. Zhao, "Stock index modeling using EDA based local linear wavelet neural network," *IEEE*, pp. 1646-1650, 2005.
- [19] M. Sheikhan, B. Movaghar, "Improvement of time series forecasting using modified GA neural hybrid architecture," *Majlesi Journal of Multimedia Processing*, vol. 1, no. 2, pp. 42-51, June 2012.