

Performance Evaluation of ANFIS and GA-ANFIS for Predicting Stock Market Indices

Farnaz Ghashami¹ & Kamyar Kamyar²

¹ Ph.D. Candidate, LeBow School of Business, Drexel University, Philadelphia, USA

² Ph.D. Student, Department of Economics, The Ohio State University, Columbus, USA

Correspondence: Kamyar Kamyar, Department of Economics, The Ohio State University, Columbus OH 43210, USA

Received: March 29, 2021

Accepted: May 7, 2021

Online Published: June 5, 2021

doi:10.5539/ijef.v13n7p1

URL: <https://doi.org/10.5539/ijef.v13n7p1>

Abstract

A model of Adaptive Neuro-Fuzzy Inference System (ANFIS) trained with an evolutionary algorithm, namely Genetic Algorithm (GA) is presented in this paper. Further, the model is tested on the NASDAQ stock market indices which is among the most widely followed indices in the United States. Empirical results show that by determining the parameters of ANFIS (premise and consequent parameters) using GA, we can improve performance in terms of Mean Squared Error (MSE), Root Mean Squared Error (RMSE), coefficient of determination (R-Squared) in comparison with using solely ANFIS.

Keywords: ANFIS, GA (Genetic Algorithm), stock market prediction, hybrid optimization

1. Introduction

The stock market is among the most comprehensive and important pillars of any country's economy, as it plays a crucial role in the process of growth of industries and over time brings the economic workforce into nourishment. As a never-ending trial to acquire profits through investment into certain stock market indices, the machine learning, forecasting, and time-series literature has been captivated with applying results of particular algorithms to forecasting of stock market price indices. Consequently, over the last decades, academic literature in stock market forecasting has gained additional attention.

Gourav et al. (2020) discusses the challenges involved with the task of predicting stock market indices. The arguments presented generally involve issues resulted by the chaotic nature of stock price time series: the data is non-parametric, volatile, highly noisy, non-linear, complex, and dynamic. Further, Gholamiangonabadi et al. (2014) elaborates on the behavior of stock markets depending on numerous factors such as economic and political and/or businesses' situation, unrevealed preferences of institutional investors, other stock markets, expectation of the investors, and so on. Such factors either permanently and/or temporarily affect the stock market.

For reasons including above, scientific literature majorly relies on various types of intelligent systems to forecast stock indices, on which trading decisions will further be made on. Upon forecasting, the investment literature would potentially gain additional opportunities for gaining profits by receiving more accurate forecasts of the stock index.

Based upon ongoing research, adaptive neuro-fuzzy inference system (ANFIS) method handles the uncertainty in prediction and has widely been applied in various research contexts even including medical system, image processing, electrical system and so on. We have picked the methodology since it has majorly shown to perform a good performance for modeling complex datasets. In this paper, we first use adaptive neuro-fuzzy inference system (ANFIS) method and try to minimize the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) and find the value of Coefficient of Determination (R-Squared) for different structures of ANFIS.

Although adaptive neuro-fuzzy inference systems have been widely used in different studies, they have a few drawbacks. First, the learning process of ANFIS is gradient-based, which may lead into particular traps in local minima. Second, when the number of input data are significantly large, the number of respective rules and their tunable parameters increase exponentially, causing into computational complexities (Salleh et al., 2018) Extensive research has been conducted in order to reduce the complexity of ANFIS and to overcome its

drawbacks using different evolutionary algorithms. In this paper, we make an effort to improve its performance while avoiding the mentioned drawbacks (i.e., getting trapped at the local optima) by determining the parameters of ANFIS using and evolutionary algorithms, Genetic Algorithm, and compare the result based on error terms, R-Squared as well as running time. Despite the substantial ongoing research on ANFIS methodology and applications, our method and algorithm is novel in the context.

The structure of the paper is organized as follows: The second section is dedicated to the literature review. In the third section the detailed presentation of the used methodology, data description and experiment will be presented. In the fourth section, results and discussion will be presented; and finally, in the fifth section, the conclusion and possibilities of future work are discussed.

2. Literature Review

Few research papers have approached this question in line with our methodology. Salleh and Hussain (2016) refer to applications of ANFIS (excluding GA), noting that such methods require training parameters effectively so as to perform efficiently as the inputs count increases. In terms of applications, they write of financial applications, namely, to predict financial crises, bankruptcy, credit risk, currency, stocks prices, and gold prices. (Salleh et al., 2016)

Similar ANFIS- hybrid methods in the literature of time-series forecasting include autoregressive adaptive network fuzzy inference system (AR-ANFIS) by Sarica et al. (2016), a hybrid model of ARIMA (Auto Regressive Integrated Moving Average) and ANFIS by Barak and Sadegh (2016), and a Quantum-behaved PSO (Particle Swarm Optimization) and ANFIS by Bagheri et al. (2016).

With regards to similar contexts, Chen's (2013) hybrid ANFIS model predicts business failure using PSO and subtractive clustering. Another notable paper in forecasting stocks prices is Wei et. al. (2014), which applies a hybrid ANFIS-MATI method (moving average technical index model) based on an n-period moving average model to forecast TAIEX stock prices.

3. Model

3.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

Neural networks can be used as appropriate tools of pattern matching and can have their weights automatically adjusted to have their behavior optimized, despite their inability to deduct how to reach to certain decisions. On the other hand, Fuzzy Logic fills such caveat by deducting how to reach to the certain decision yet are unable to "train" themselves and learn automatically. As a result, a simultaneous use of these methods could perform as complements in order to overcome a number of drawbacks proposed by hybrid networks.

Among them, the Adaptive Neuro-Fuzzy Inference System (ANFIS) is one of the best possible combinations that is broadly applied in various areas of research. The adaptive neuro-fuzzy inference system was introduced by Jang in 1993. (Salleh et. al., 2018) ANFIS provides links connecting the neural network and fuzzy logic for modeling complicated and dynamic systems. (Aghbashlo et. al., 2016) Through ANFIS, "IF-THEN" rules are applied through generating a mapping between the variables of input and output. The architecture of ANFIS is composed by five layers. There are several single-processing elements called neurons that are organized in each layer. Every neuron of each layer is connected to another layer by directed links. In order to produce the value of each neuron, neurons perform a particular function on its incoming signals. Figure 1 illustrates the ANFIS model. (Salleh et. al., 2018) The five layers of ANFIS model follow:

Layer 1 (Fuzzification):

Letting y and x define the inputs and $O_{1,i}$ be the output of layer one for the i th node which is computed as follows:

$$O_{1,i} = \mu_{A_i}(x), \quad i = 1, 2 \text{ or} \quad (1)$$

$$O_{1,i} = \mu_{B_{i-2}}(y), \quad i = 3, 4$$

$$\mu(x; c, \sigma) = e^{-\frac{1}{2} \left(\frac{x-c}{\sigma} \right)^2} \quad (2)$$

μ_{A_i} and μ_{B_i} are Gaussian membership functions specified by two parameters with center c and width σ , given that A_i and B_i are the members values. We proceed with the Gaussian membership function since it achieved better results.

Layer 2 (Product):

This layer accepts input values from first layer, the output of Layer 2 which is known as the firing strength of a

rule is as follows:

$$O_{2,i} = \mu_{A_i}(x) \cdot \mu_{B_{i-2}}(y), i = 1,2 \tag{3}$$

Layer 3 (Normalization):

The output of Layer 3 is computed by normalizing the firing strength of a rule from the previous layer by calculating the ratio of the *i*th rule's firing strength to the sum of all rules' firing strength as below:

$$O_{3,i} = \bar{w} = \frac{w_i}{\sum_{i=1}^n w_i} \tag{4}$$

A normalized firing strength of a rule denoted by \bar{w}

Layer 4 (Defuzzification):

The output of Layer 4 is calculated as follows:

$$O_{4,i} = \bar{w}f_i = \bar{w}_1(p_i x + q_i y + r_i), i = 1,2 \tag{5}$$

Where r_i , q_i , and p_i define the consequent parameters of the node *i*.

Layer 5 (Overall output):

Layer 5 contains only one node and the output of layer 5 computed as follows:

$$O_{5,i} = P_i \bar{w} f_i \tag{6}$$

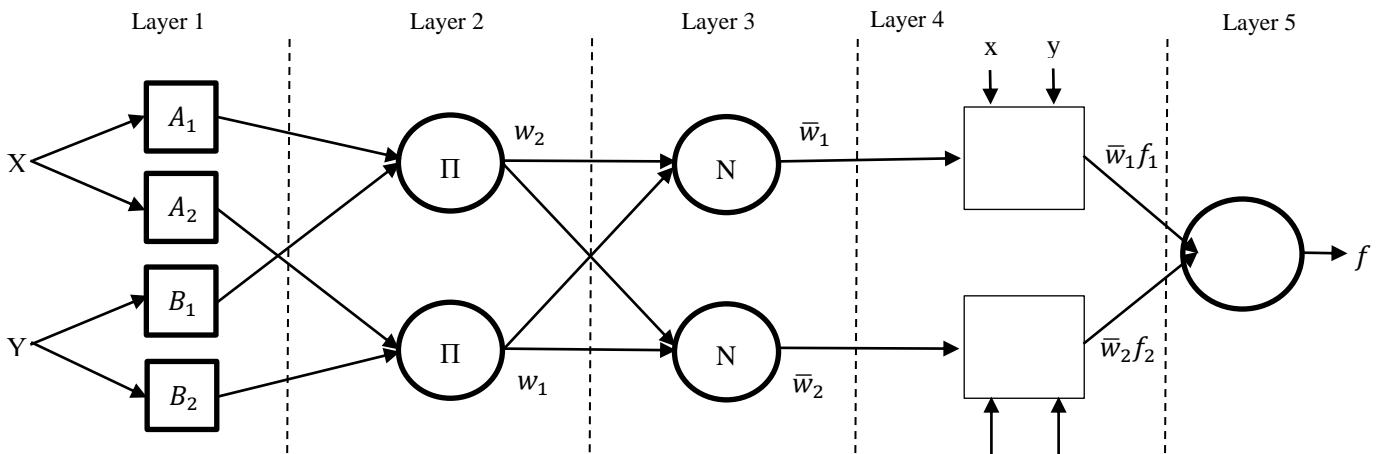


Figure 1. Adaptive Neuro-Fuzzy inference system

The adaptive neuro-fuzzy inference system trains its parameters for minimizing the error term between actual and predicted output. The procedure follows: during forward pass, consequent parameters (r_i , q_i , and p_i) are updated by the least square estimator (LSE); and, during backward pass, the premise parameters (c and σ) are updated by gradient descent and neural networks trains.

ANFIS uses three methods, namely grid-partitioning, subtractive clustering, and Fuzzy C-Mean to partition the input and output values into rule patches. It is appropriate to use grid-partitioning only for problems with any number of input variables below six. Fuzzy C-Mean algorithm allows one set of data to be in two or more clusters. (Samat & Salleh, 2016) By considering each method's objective, we applied subtractive clustering method. Subtractive clustering is an appropriate method for more than six number of input variables, and reduces the computation time by finding the center of clustering by using the data itself. (Le & Altman, 2011)

3.2 Genetic Algorithm

Initially developed by John Holland, Genetic Algorithm (GA) is a metaheuristic inspired by the process of natural selection, mutation, and crossover. Genetic Algorithm can be implemented on the multidimensional problems with large-size search space and with a great number of variables. (Haznedar & Kalinli, 2016) Genetic Algorithm starts with a finite population of solutions and evolves it from one generation to the next. Basic steps of the Genetic Algorithm follow:

- 1) Random populations of the n individual generated.
- 2) Evaluation: the fitness of each individual is evaluated.
- 3) Parent Selection: based on their fitness certain pairs of solutions (parents) are selected.
- 4) Crossover: one or two new solutions (offspring) produced by combining each pair of parents.
- 5) Mutation: some of the offspring are randomly modified.
- 6) Population Selection: by considering their fitness, a new population is selected replacing some or the whole original population by an identical number of offspring.

3.3 ANFIS Trained by GA

In this paper, binary and continuous genetic algorithms were implemented for training adaptive neuro-fuzzy inference system's parameters such that during the training process, the genetic algorithm makes a binary vector which contains seven digits (seven input variables) with a total of 127 possible combinations. This vector defines the input parameters used throughout the training process of the model. ANFIS has two parameters type which require updating. These are premise parameters (c, σ) and consequent parameters (r_i, q_i, p_i) as explained in the method part. Genetic algorithm is used to update these parameters. In order to know the performance of the model we are using MSE and RMSE which gives values that we want to optimize. The computation follows.

$$MSE = 1/n \sum_i^n (y_i - \hat{y}_i)^2 \quad (7)$$

$$RMSE = \sqrt[2]{1/n \sum_i^n (y_i - \hat{y}_i)^2} \quad (8)$$

In the equation above y_i is the actual output and \hat{y}_i is the predicted output and n is the number of observations.

3.4 Data Description and Empirical Analysis

We have used NASDAQ-100 index data sourced from *Yahoo! Finance*. The NASDAQ-100 index is arguably the most widely followed index in the United States. More precisely, the used data is a modified capitalization-weighted index consisted of 103 equity securities issued by 100 of the most significant non-financial companies listed on the NASDAQ stock market. (Yahoo! Finance) The data points are of daily frequency and extend from 21 October 2020 through 17 March 2021. We considered the short-term historical stock prices (seven prior days) as the inputs dataset. So, considering the output to be $y(t)$ (the stock price at time t), with n being the number of historical data, the equation looks like:

$$y(t) = F(y(t-1), y(t-2), y(t-3), \dots, y(t-n)) \quad (9)$$

In the next step the adaptive neuro-fuzzy inference system is developed for the dataset and the best results based on R-Squared, MSE and RMSE, running time and parameters setting are shown in Table 1.

Table 1. Adaptive Neuro-Fuzzy inference system

Adaptive Neuro-Fuzzy Inference System					
Method	RMSE	MSE	Running Time	Method	R-Squared
ANFIS	270.16	72988	0.9955 sec	Subtractive Clustering	0.89
Parameters Setting	Max. epochs	Error Goal	Initial Step-Size	Step-Size Decrease Rate	Step-Size Increase Rate
—	50	0	0.01	0.9	1.1

In the following step the different architectures of hybrid genetic algorithm and adaptive neuro fuzzy inference system (GA-ANFIS) method is developed for the dataset and the best results based on R-Squared, MSE and RMSE, running time and parameters setting are shown in Table 2.

Table 2. GA-ANFIS

GA-ANFIS									
Method	R-Sqr	RMSE	MSE	Running Time	Iteration Number	Population size	Crossover Percentage	Mutation Percentage	MutationRate
GA-ANFIS	0.9703	140.76	19813	19.79 sec	5	5	0.9	0.4	0.01

In the final step we compare the best adaptive neuro-fuzzy inference system (ANFIS) architecture with hybrid

genetic algorithm-adaptive neuro-fuzzy inference system (GA-ANFIS) model in terms of R-Squared, MSE, RMSE, Running time.

Table 3. Comparison of results

Comparison of Results									
Method	R-Sqr	RMSE	MSE	Running Time	Iteration Number	Population size	Reproduction rate	Crossover rate	Mutation Rate
GA-ANFIS	0.9703	140.76	19813	19.79 sec	5	5	0.9	0.4	0.01
Method	R-Sqr	RMSE	MSE	Running Time	Max. epochs	Error Goal	Initial Step-Size	Step-Size Decrease Rate	Step-Size Increase Rate
ANFIS	0.89	270.16	72988	0.9955 sec	50	0	0.01	0.9	1.1

4. Discussion

Comparison of the results of both the proposed hybrid GA-ANFIS method (genetic algorithm and adaptive neuro-fuzzy inference system) and the standard adaptive neuro-fuzzy inference system (ANFIS) to predict stock market indices is shown in Table 3. Based on the results, it can be concluded that the proposed method outperforms the standard ANFIS. As an example, by examining the results of Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), it can be observed that the proposed model effectively achieves slightly smaller values. Additionally, based on the results the proposed method achieves better R-Squared to compare with the standard ANFIS. Despite the length of time required to perform a computational process for GA-ANFIS being slightly above the one of ANFIS, the GA-ANFIS method performs better than ANFIS in R-Squared, MSE and RMSE and this indicates the better quality of the proposed method (GA-ANFIS).

This paper proposed a modified version for the hybrid GA-ANFIS method (Genetic Algorithm and Adaptive Neuro-Fuzzy Inference System). Throughout the process, we initially use an adaptive neuro-fuzzy inference system (ANFIS) method and tried to minimize the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) and find the value of Coefficient of Determination (R-Squared) for different structures of ANFIS. Then, binary and continuous genetic algorithms (GA) are implemented for training adaptive neuro-fuzzy inference system's parameters (premise and consequent parameters). Finally, we test the performance of the ANFIS and GA-ANFIS method by applying on NASDAQ-100 index data from 21 October 2020 through 17 March 2021.

Experimental results showed that the GA-ANFIS architecture proposed in the study improve the Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Coefficient of Determination (R-Squared) and can be applied in different forecasting applications. Related areas of study include application of different techniques to tune parameters for genetic algorithm or utilizing other optimization methods to updates the parameters of ANFIS, as well as using different architectures to improve the slow convergence speed of GA-ANFIS model. Finally, we can effectively compare the results of this study with different hybrid techniques which applied on forecasting stock market indices.

References

- Aghbashlo, M., Hosseinpour, S., Tabatabaei, M., Hosseini, S. S., Najafpour, G., & Younesi, H. (2016). An exergetically-sustainable operational condition of a photo-biohydrogen production system optimized using conventional and innovative fuzzy techniques. *Renewable Energy*, *94*, 605-618. <https://doi.org/10.1016/j.renene.2016.03.061>
- Bagheri, A., Mohammadi, P. H., & Akbari, M. (2014). Financial forecasting using ANFIS networks with quantum-behaved particle swarm optimization. *Expert Systems with Applications*, *41*(14), 6235-6250. <https://doi.org/10.1016/j.eswa.2014.04.003>
- Barak, S., & Sadegh, S. S. (2016). Forecasting energy consumption using ensemble ARIMA-ANFIS hybrid algorithm. *International Journal of Electrical Power and Energy Systems*, *82*, 92-104. <https://doi.org/10.1016/j.ijepes.2016.03.012>
- Chen, M. (2013). A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. *Information Sciences*, *220*, 180-195. <https://doi.org/10.1016/j.ins.2011.09.013>
- Ghashami, F., Kamyar, K., & Riazi, S. A. (2021). Prediction of Stock Market Index Using a Hybrid Technique of

- Artificial Neural Networks and Particle Swarm Optimization. *Applied Economics and Finance*, 8, 1. <https://doi.org/10.11114/aef.v8i3.5195>
- Gholamiangonabadi, D., Mohseni, T. S. D., Mohammadi, A., & Menhaj, M. B. (2014). Investigating the performance of technical indicators in electrical industry in Tehran's Stock Exchange using hybrid methods of SRA, PCA and Neural Networks. *2014 5th Conference on Thermal Power Plants (CTPP) IEEE*. <https://doi.org/10.1109/CTPP.2014.7040698>
- Haznedar, B., & Kalinli, A. (2016). Training ANFIS using genetic algorithm for dynamic systems identification. *International Journal of Intelligent Systems and Applications in Engineering*, 44-47. <https://doi.org/10.18201/ijisae.266053>
- Howard, D., Mark, B., & Martin, H. (2009). *Neural Network Toolbox 6 User's Guide for MATLABfi*.
- Kheirollahi, R., Tahmasebifar, R., & Dehghanpour, E. (2017). Genetic Algorithm Based Optimal Coordination of Overcurrent Relays Using a Novel Objective Function. *International Electrical Engineering Journal (IEEJ)*, 7(10), 2403-2414.
- Kumar, G., Jain, S., & Singh, U. P. (2021). Stock Market Forecasting Using Computational Intelligence: A Survey. *Arch Computat Methods Eng.*, 28, 1069-1101. <https://doi.org/10.1007/s11831-020-09413-5>
- Le, T., & Altman, T. (2011). A new initialization method for the Fuzzy C-Means Algorithm using Fuzzy Subtractive Clustering. *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- Salleh, M. N. M., & Hussain, K. (2016). A review of training methods of anfis for applications in business and economics. *International Journal of U- and E- Service, Science and Technology*, 9(7), 165-172. <https://doi.org/10.14257/ijunesst.2016.9.7.17>
- Salleh, M. N. M., Talpur, N., & Talpur, K. H. (2018). A Modified Neuro-Fuzzy System Using Metaheuristic Approaches for Data Classification, Artificial Intelligence - Emerging Trends and Applications, Marco Antonio Aceves-Fernandez. *IntechOpen*. <https://doi.org/10.5772/intechopen.75575>
- Samat, N. A., & Salleh, M. N. M. (2016). A study of data imputation using fuzzy c-means with particle swarm optimization. *International Conference on Soft Computing and Data Mining*. Springer, Cham. https://doi.org/10.1007/978-3-319-51281-5_10
- Sarıca, B., Eğrioğlu, E., & Aşıkçıl, B. (2018). A new hybrid method for time series forecasting: AR–ANFIS. *Neural Comput & Applic.*, 29, 749-760. <https://doi.org/10.1007/s00521-016-2475-5>
- Wei, L., Cheng, C., & Wu, H. (2014). A hybrid ANFIS based on N-period moving average model to forecast TAIEX stock. *Applied Soft Computing*, 19, 86-92. <https://doi.org/10.1016/j.asoc.2014.01.022>
- Yahoo! Finance. (n. d.). *Yahoo! Finance - Stock Market Live, Quotes, Business and Finance News*. Retrieved from <https://finance.yahoo.com/>
- Yarpiz. (2021). ANFIS Training using Evolutionary Algorithms and Metaheuristics, MATLAB Central File Exchange.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).