STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **12**, November 2024, pp 1934–1955.
Published online in International Academic Press (www.IAPress.org)

A Novel Hybrid ANFIS-NARX and NARX-ANN Models to Predict the Profitability of Egyptian Insurance Companies

Hanaa H. A. Aboul Ela *

Department of Statistics, Mathematics and Insurance, Faculty of Business, Ain Shams University, Cairo, Egypt

Abstract The use of fuzzy logic models with machine learning (ML) models have become common in many areas, especially insurance field. This study aims to compare between non-hybrid models such as artificial neural network (ANN) model, nonlinear auto-regressive with exogenous inputs (NARX) model, and the following hybrid models adaptive neural fuzzy inference system (ANFIS) model, (ANFIS-NARX) model and (NARX-ANN) model to predict the profits of the insurance activity which represent the important indicator of the good performance of Egypt's 39 insurance companies in the period from 1st January 2009 to 31 December 2022 per month. This prediction based on the following factors (net premiums (NP), reinsurance commissions (RC), net income from investments (NIFINV), net compensation (NC), commissions of production cost (CPC), general and administrative expenses (GAE),that help decision makers to make appropriate decisions. The results found that the(ANN) model is given good results compared with the following models (ANFIS), (NARX), hybrid (ANFIS-NARX) and (NARX-ANN) models according to the following prediction accuracy measures (RMSE, MAPE, MAE and Theil Inequality). The explanatory ability criterion (R^2)) was appeared $(0.79, 0.61)$ respectively for training and testing phases in persons insurance companies. The explanatory ability also was appeared$(0.83, 0.68)$ respectively in property insurance companies.

Keywords Insurance companies; Fuzzy logic; Membership functions; Adaptive neural Fuzzy inference system (ANFIS) model; Artificial neural network (ANN) model; Nonlinear auto-regressive with exogenous inputs (NARX) model.

AMS 2010 subject classifications: 94D05; 68T07; 68T05; 62J86; 68Q25; 70K75.

DOI: 10.19139/soic-2310-5070-2104.

1. Introduction

The insurance industry has an important role in achieving sustainable development by protecting individuals and families from falling into poverty when losses caused by insured hazards occur. Medical insurance also helps to raise the level of health care and fight diseases. Insurance activity also contributes to face the risks caused by climate changes. Property insurance helps also to retain investments of individuals, companies and successive governments. The main objective of this research is to make a comparison between the fuzzy models and the machine learning models in order to predict the surplus of insurance activity of persons insurance companies and property insurance companies in the Egyptian market. A fuzzy logic model is a technique that has a great mechanical ability to find solutions to many problems, whether scientific or applied. In addition, fuzzy logic has evolved as a result of the orientation towards the use of artificial intelligence and expert systems. As well as the evolution of the field of software and computers that can only deal with accurate and specific information and modern systems and technologies which include the fuzzy logic.

*Correspondence to: Hanaa H. A. Aboul Ela (Email: dr.hanaaaboualaela@bus.asu.edu.eg). Department of Statistics, Mathematics and Insurance. Faculty of Business, Ain Shams University,(11566). El khalifa El Mamoon St., Abbasia, Cairo, Egypt.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The main importance of this research is due to the scarcity of researches that deal with a comparison between the traditional models such as ANN and NARX versus the hybrid models such as ANFIS, ANFIS-NARX and NARX-ANN models, which is one of the modern machine learning models in insurance sector. The study hypotheses are:(1) There is a significant effect between net premiums (NP), net compensation (NC), net investment income (NIFINV) and the profits of insurance companies of persons, and there is an insignificant effect between commissions of production cost (CPC), general and administrative expenses (GAE), reinsurance commissions (RC) and the profits of insurance companies of persons. (2) The There is a significant effect between net premiums (NP), net compensation (NC), net investment income (NIFINV) and the profits of insurance companies of property, and there is an insignificant effect between commissions of production cost (CPC), general and administrative expenses (GAE), reinsurance commissions (RC) the profits of insurance companies of property. (3) The predictive and explanatory ability in hybrid models such as (ANFIS, ANFIS-NARX, and NARX-ANN) will be expected to be better than non- hybrid models (ANN), (NARX) in insurance companies of persons and property. The study data were obtained from the Financial Regulatory Authority (FRA) and the Central Authority for Public Mobilization and Statistics (CAPMAS) from 1 January 2009 to 31 December 2022. The data were monthly, and the study applied to 39 insurance companies divided into 16 persons insurance companies, and 23 property insurance companies.

2. Literature review

Worku et al[33] identified determinants of profitability of insurance companies in Ethiopian. The linear regression model estimation of ordinary least squares (OLS) was employed to identify the determinants of profitability of insurance companies in Ethiopia at 5% level of significance. It revealed that age of the company, tangibility of an asset, size of the company, managerial efficiency, leverage ratio, premium growth and GDP have a positive relationship with return on assets while loss ratio and inflation have a negative relationship with return on assets. On the other hand, age of the company size, managerial efficiency, leverage ratio, liquidity ration inflation and premium growth have statistically significant at 5% confidence interval level, whereas the other variables such tangible assets and GDP have no statistical significance at 5% confidence interval level. Anam et al[4] used machine learning methods to predict the claims of health insurance users. The results showed that the SVM method with PSO gives the greatest performance in the health claim insurance prediction. The SVM method with PSO for predicting claims on health insurance is superior to standard SVM on three evaluation metrics used from four evaluation tools, but the computation time required is longer. Khadka[15]examined the effect of firm specific and macroeconomic factors on profitability of Nepalese insurance companies using the Pearson's correlation coefficients and linear regression (LR) models. The result showed that firm size and liquidity rate, inflation and money supply have the negative impact on return on assets and return on equity. On the other hand, the study found that tangible assets, dividend per share, premium growth and gross domestic products have the positive impact on return on assets and return on equity of Nepalese insurance companies. Lalon and Das[18] used the sample of seven insurance companies of Bangladesh based on the period of 2010-2019. The study used several models such as Pooled OLS, Cross-sectional GLS, Fixed-effect, Random-effect, and One-step GMM approach. The results showed that inflation liquidity and leverage have a positive impact on ROE. Underwriting risk and size of company have affected negatively affect ROE but positively affect EPS. Reinsurance dependence had a large positive influence on EPS and ROE. In contrast. The tangible assets negatively affect both EPS and ROE. Finally, using the GMM technique found that the model fits the data well for predicting EPS and ROE as profitability measures explaining 89.38% of EPS variation and 80.38% of ROE variation using Pooled OLS. Lupačov and Stanković [20] used a type of artificial intelligence (AI) in insurance companies. The average and standard deviation for profitability of insurance companies measured by using Return on assets (ROA) for Ethiopian insurance companies was 0.117 and 0.08 , respectively. In the multivariate analysis, age of company, Firm Growth, Company Size, Leverage and market share are highly significant predictors of profitability. Furthermore, age of company, leverage and branch distribution was found the crucial factors in determining the profitability of insurance companies using multivariate analysis in Ethiopia.Vojinovi'c et al[32] used a set of standard panel regression models, such as the Pooled ordinary least squares (POLS) model, Fixed Effects model (FEM), and

Random Effects model (REM) followed by a more robust GMM estimation to uncover the relationship between selected micro-specific, macroeconomic, and institutional factors, and return of assets (ROA) and return on total premiums (ROTP). The paper found that firm size, GDP, the population growth rates, political stability, and the degree of specialization in some empirical models lead to higher profitability. On the other hand, we observe that excessive risk-taking and inflation in some specifications are inversely related to profitability. Olarewaju and Msomi[23]analyzed the profitability of 42 reinsurance companies in Sub- Saharan Africa from 1991 to 2020, revealed that various factors such as gross domestic product, competition (HHI), premium growth, investment performance, underwriting risk, and operational efficiency affect the profitability in these companies. This study is quantitative and dynamic using the system-generalized method of moments (GMM) to analysis the data.Ahmeti and Iseni[2] showed how efficiently management generates profit by utilizing all available resources. This paper used the linear regression (LR) equation, and examined the effects of specific company factors, the independent variables are: liquidity, company size, company age, tangible asset, leverage, company capital and growth of the company, on profitability represented by the return on assets (ROA) and net profit margin (NPM) as a dependent variable. The sample in this study includes eleven insurance companies for the period 2015 - 2020. The regression results indicate that size, leverage and age of company, have significant effects on the ROA. Kaur and Bassi[14] focused on investigating the efficiency of artificial neural network (ANN) and support vector machine SVM across insurance companies of CNX 500. The study's findings revealed that ANN performed best for the ICICIPRULI data model in terms of hit ratio. Whereas the performance of SVM was observed to be the best for the ICICIGI data model. In the case of pairwise comparison among the six selected Indian insurance companies from CNX 500, the extracted data evaluated and concluded that there were eight significantly different pairs based on hit ratio in the case of ANN models and nine significantly different pairs based on hit ratio for SVM models. Dhiab[7] examined the determinants of profitability in the Saudi insurance sector. The empirical analysis in this study is based on data relative to a sample of 20 Saudi insurance companies between 2009 and 2017. For robustness checks, the empirical investigation employs a wide range of econometric techniques, including the fixed-effects model (FEM), random-effects model (REM), Feasible Generalized Least Squares (FGLS), Ordinary Least Squares (OLS) with panel-corrected standard errors, Difference GMM and finally System GMM. The empirical findings suggest that the growth rate of written premium, the tangibility ratio and the fixed-assets ratio are the main factors affecting positively the profitability of Saudi insurance companies. Moreover, while the company size and the liquidity ratio are positively associated with profitability, their impacts are not statistically significant. On the contrary, the loss ratio, liabilities ratio, insurance leverage ratio, and to a less extent, the company age have negative effects on the profitability of Saudi insurance companies.Faoziyyah and Laila[9] examined the company's internal factors and macroeconomic factors partially and simultaneously on the profitability of Islamic insurance companies in Indonesia. This study applied to 36 Islamic insurance companies, which consist of Islamic general and life insurance companies in Indonesia during the period 2015-2018. The estimation results of the Fixed Effect Model (FEM) with the Weighted Least Square (WLS) method show that company size, contribution growth, retakaful, leverage, investment returns, GDP and inflation simultaneously affect the profitability of Islamic insurance companies in Indonesia. Partially, contributions growth and investment returns have a positive and significant effect on the profitability.Al-Dwiry et al[3] aimed to evaluate the major factors that affect the insurance industry in the Middle East countries through applying the principal components analysis (PCA) method. The researchers studied and analyzed 19 variables that have an impact on the insurance industry. The study revealed that the first factor (Inter developed insurance culture) was the most significant factor, followed by the second factor (Industry's dependence on traditional insurance products), and then comes the eighth factor (The weakness and strength of the appropriate legislation and regulatory systems) that apply in the Arab countries in terms of legislation, supervision and regulation of the industry.

Most previous studies examined the relationship between insurance companies' profitability and internal factors and macroeconomic factors affecting them. The current study focuses on predicting a surplus of insurance activity in insurance companies, relying on financial indicators on the income list. Previous studies have also been interested in the application of the multiple linear regression model (MLR), panel models such as (POLS, REM, FEM, FGLS and GMM), principal components analysis (PCA), support vector machine model (SVM), and artificial neural networks (ANN). The current study was characterized by the use of modern models for prediction and comparison between the non-hybrid models such as (ANN and NARX) models and the hybrid models such as (ANFIS, ANFIS-NARX and Narx -ANN) from 1st January, 2009 to 31 December 2024. Many economic events have occurred including the global crisis (2007 - 2008) and the crisis of the COVID-19 pandemic, as well as the Ukrainian-Russian crisis which has affected the Egyptian economy and the economies of many developed and developing countries.

3. Materials and methods

Before presenting the statistical methods which were used in the current study, we will define the meaning of the fuzzy set, the types of member functions and fuzzy numbers as an important part of the hybrid model which was used in the current study.

3.1 Fuzzy set

The fuzzy set is defined as a set of elements that have a degree of affiliation or membership whose range is within the period $[0,1]$. When an element has a degree of affinity (0), it means that the element belongs to a degree of zero to the fuzzy set. If the element's degree of affiliation equals one, it means that the element belongs exactly to the set. We also find that there are scores between zero and one, if the element has a degree of affinity of 0.5 , that means that the element belongs to the set at 50%, and the same proportion does not belong to the set, this element is called the balance point Klir[17].

3.2 Types of membership functions

Membership is an expression of the degree of affiliation. It is the function which uses to calculate the degree of membership of an element for the fuzzy set. The symbol of this function is $\mu_A(x)$ that shows the degree of affiliation of the variable x for the fuzzy set A. The following is the membership function: $\mu_A(x) = x \to [0,1]$, when $\mu_A(x) = 0$ this means that the variable x does not belong to the fuzzy set, when $\mu_A(x) = 1$ this means that the variable x belongs exactly to the set.

There are several types of membership functions as follows.

- Standard Membership Function is an increasing function as shown in equation 1:

$$\mu_A(x) = \left\{ \begin{array}{l} 0.........................for.........x \leq \alpha \\ 2\left[\frac{x-\alpha}{c-\alpha}\right]^2..........for.....\alpha \leq x \leq \beta \\ 1-2\left[\frac{x-c}{c-\alpha}\right]^2.....for....\beta \leq x \leq c \\ 1.......................for...........x \leq c \end{array} \right\} \tag{1}$$

Where x represents the elements of the fuzzy set. α is an element of the set that its function value is equal to (0). β is an element of the set whose function value is equal to (0.5). $\mathcal{C}$ is an element of the set whose function value is equal to (1).

-Triangular Membership Function is a linear function that is like a shape of the triangle. The base of the triangle represents the specified period, and its head represents the center of the fuzzy number, and this is illustrated by equation 2:

$$\mu_A(x) = \left\{1 - \frac{|x-a|}{s} \text{.....when.....}\ a-s \leq x \leq a+s\right\} \tag{2}$$

- Bell-Shaped Membership Function is a linear exponential function that takes the normal curve shape. This is shown in equation 3:

$$\mu_A(x) = \left\{ce^{\frac{(x-a)^2}{b}} \ldots\ldots\ldots\ldots\ldots -\infty < x < \infty\right\} \tag{3}$$

-Trapezoidal Membership Function : The characteristics of this function are determined by four parameters such as (a, b, c, d). The variable x for this function is determined by the following formula:

$$\mu_A(x) = \left\{ \begin{array}{l} 0\text{.....................}x < a \\ \frac{x-a}{b-a}\text{..........}a \leq x \leq b \\ 1\text{................}b \leq x \leq c \\ \frac{d-x}{d-c}\text{.........}c \leq x \leq d \\ 0\text{.....................}x > d \end{array} \right\} \tag{4}$$

- Gaussian Membership Function is an econometric model for describing many phenomena. It depends on two parameters (σ, c). They are the function's center and its width according the following formula:

$$\mu_A(x) = e^{\frac{-(x-c)^2}{2\alpha^2}} \tag{5}$$

- Generalized Bell Membership Function: The characteristics of this function are determined by three parameters (a, b, c). The parameters of this function are determined by the following formula:

$$\mu_A(x) = \frac{1}{1 + \left|\frac{(x-c)}{a}\right|^{2b}} \tag{6}$$

3.3 Artificial neural networks (ANN) model

An artificial neural network is a computational model that follows the behavior of the human brain. Haykin and Lippmann[11]. ANN can be defined as structures comprised of densely neurons or nodes) that are capable of performing massively parallel computational data processing and knowledge representation. Basheer and Hajmeer[5]. Every neuron in the network computes a weighted by w_{ij} that represents the sum of its p inputs signal. y_i, for $i = 1, 2, 3, \ldots\ldots\ldots\ldots n$ hidden layer and then applies a non-linear activation function to produce an output signal u_j. The form of this function is as follows.

$$u_j = \sum_{i=0}^{n} w_{ij} y_i \tag{7}$$

The ANN method most commonly used for prediction. It has been the multilayer feedforward neural network (MLF) with the backpropagation (BP) algorithm by Yadav and Chandel[34]. Qazi et al[26]and Rezrazi et al[28] This method is popular due to its ability for modeling problems that are not linearly separable. The MLF consists of an input layer, an output layer and usually one or more hidden layers. In practice, only a three-layer feedforward neural network (FFNN) is usually necessary, as seen in Figure.1, Sitharthan and Rajesh[29].

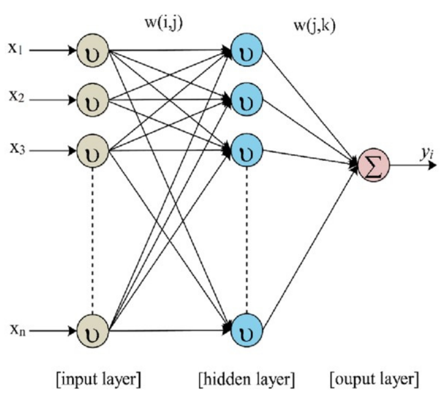

Figure 1. Artificial neural network (ANN) structure

where the first layer is the input layer representing input variables (i), the second layer is the hidden layer (j), and the third layer is the output layer (k). Each layer is interconnected by weights w_{Ij} and w_{jk}, and every unit sums of its inputs, adds a bias or threshold term to the sum and non linear transforms the sum to produce an output.This nonlinear transformation is called the activation function of the node. The output layer nodes often have linear activations. In MLFs, the logistic sigmoid function in equation 8 and linear function in equation 9 are generally used in the hidden and output layer respectively.Rezrazi et al[28].

$$f(w) = 1/\left(1 + e^{-w}\right) \tag{8}$$

$$f(x) = x \tag{9}$$

where w is the weighted sum of the input and x is the input to the output layer. The procedure for updating synaptic weights is called backpropagation (BP). BP refers to the way the error computed on the output side is propagated backward from the output to the hidden layer(s) and finally to the input layer. Esmaeelzadeh et al.[8]. The error is minimized across many training cycles called epochs. During each cycle, the network reaches a specified level of accuracy. Generally, the error estimator used here is the sum of the squared error (SSE). In conjunction with the BP procedure, the following algorithm can be used as a second training algorithm: Levenberg-Marquardt backpropagation (trainlm). The selection of an appropriate training algorithm, the transfer function, and the number of neurons in the hidden layer are fundamental characteristics of the ANN model. Each training algorithm has its own characteristics that must be adjusted according to a particular model. Quej et al.[27].

3.4 Adaptive neuro-fuzzy inference system (ANFIS) model

This a combined system of artificial neural networks and fuzzy inference performed that use the numerical data to predict the output , and this system represents an influential tool. ANFIS is a type of neural network focused on Takagi-Sugeno fuzzy inference system. It is an AI technique currently using in hydrological processes. Bisht and Jangid[6]. ANFIS was first introduced by Jang [12]. It is based on the first-order Sugeno fuzzy model. ANFIS commonly uses either back-propagation or a combination of back-propagation and least square estimation for prediction of membership function parameter, Jang et al[13]. Takagi-Sugeno type fuzzy inference system is used in ANFIS where every rule's output can be a constant term or can be a linear combination of input variables addition to a constant term. The weighted average of every rule's output is the final output. Sonmez et al[30].The basic architecture of ANFIS which has several inputs and one output is presented in Figure 2, Stefenon et al[31].

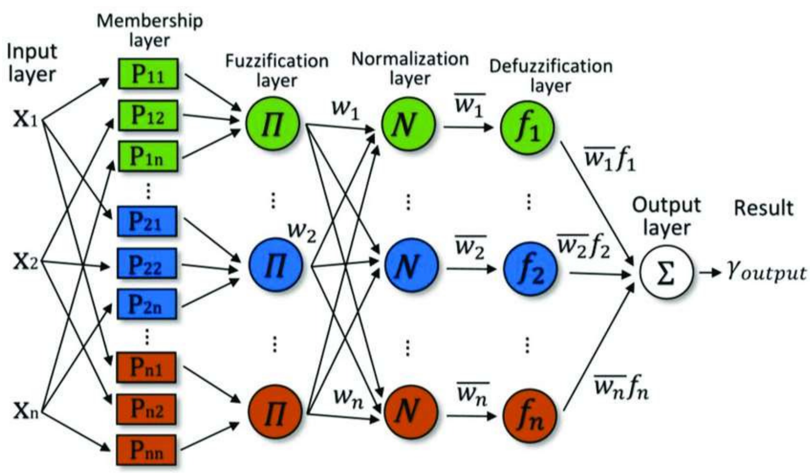

Figure 2. The structure of the ANFIS model.

The rule base of ANFIS contains two Takagi-Sugeno type if-then rules as given below:

$$\text{Rule 1 : If} x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{; then } f = p_1 x + q_1 y + r_1 \tag{10}$$

$$\text{Rule 2 : If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{; then } f = p_2 x + q_2 y + r_2 \tag{11}$$

where A_1, A_2, B_1 and B_2 are non-linear parameters while p_1, p_2, q_1, q_2, r_1 and r_2 are linear parameters.

The structure of ANFIS model contains five layers that we can explain them as follows:

Layer 1 : is the fuzzification layer in which x is the input of A_1 and A_2 nodes and y is the input of B_1 and B_2 nodes. A_1, A_2, B_1 and B_2 are used in fuzzy theory to allocate membership functions as linguistic labels. The membership relationship between the input and output functions of this layer can be shown below.

$$\begin{cases} Q_{1,i} = \mu_{Ai}(x), \ldots\ldots i = 1,2 \\ Q_{1,j} = \mu_{Bj}(y), \ldots..j = 1,2 \end{cases} \tag{12}$$

where μ_{Ai} and μ_{Bj} indicate the membership functions and $Q_{1,i}$ and $Q_{1,j}$ indicate the output functions.

Layer 2: is the product layer which includes two fixed nodes labeled with Π. The outputs of this layer are w_1 and w_2. These outputs are the weight functions of layer 3 and the product of the input signal can be shown as follows:

$$Q_{2,i} = w_i = \mu_{Ai}(x)\mu_{Bi}(y) \text{ where } i = 1,2 \tag{13}$$

Layer 3: is the normalized layer, which includes two fixed nodes labeled with N. w_1 and w_2 are the outputs of this layer. The normalizing of the weight function is the task of this layer in the next process:

$$Q_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2} \quad \text{ where } i = 1,2 \tag{14}$$

Layer 4: is the defuzzification layer which includes two adaptive nodes. The relationship between the inputs and output of this layer can be shown as follows:

$$Q_{4,i} = \overline{w_i}\left(p_i x + q_i y + r_i\right) \quad \text{ where } i = 1,2 \tag{15}$$

where p_i, q_i and r_i are the linear parameters of the node and $Q_{4,i}$ is the output of this layer.

Layer 5: is the output layer that includes a fixed node labeled $\sum$. The output of this layer comprises all the input components, which denotes the cleaning rates results. The output can be defined as follows:

$$Q_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad \text{where} \quad i = 1,2 \tag{16}$$

3.5 Nonlinear auto-regressive with exogenous inputs (NARX) model

The NARX model is used for various directions, for example, control and system identification. Also, it is able to predict the output real time data. The NARX algorithm can be expressed as follows, Peng et al.[25]

$$y(t) = \frac{B\left(Z^{-1}\right)}{A\left(Z^{-1}\right)} u(t) + \frac{\xi\left(z^{-1}\right)}{A\left(z^{-1}\right)} \tag{17}$$

where

$$A\left(z^{-1}\right) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n} \tag{18}$$

$$B\left(z^{-1}\right) = b_0 + b_1 z^{-1} + \ldots + b_n z^{-(n-1)} \tag{19}$$

After neglecting the noise error and defining the $B\left(z^{-1}\right)$ and $A\left(z^{-1}\right)$:

$$y(t) = f\left[y(t-1), \ldots., y\left(t - n_a\right), \ldots u\left(t - n_k\right), .., u\left(t - n_k - n_b + 1\right)\right] \tag{20}$$

n_a& n_b represent the previous value of input and output respectively. While, the n_k represents the input delay. Finally, f represents the nonlinear function that can be performed using intelligent methods such as the neural networks. Mohammed and Darus [21]

3.6 The hybrid models of ANFIS-NARX and NARX-ANN

The current study generally uses both hybrid and non-hybrid models to solve deficiencies in traditional methods. This is one of the best methods used to predict financial aspects in Egyptian companies. The ANFIS model is also an integrative model that links between the fuzzy logic and the eductional rules of neural networks, as well as the fuzzy rules help in interpreting of the results . Also,the combination of models gives better results. We note that to predict the profitability values in insurance companies in the ANFIS-NARX hybrid model, the predictive values of the ANFIS model will be used as a dependent variable in the NARX model and then follow the inputs variables effect on them and then predict it. We also note that to predict profitability values in insurance companies in the NARX-ANN hybrid model, the predictive values of the NARX model will be used as a dependent variable in the ANN model and then follow the inputs variables effect on them and then predict it.

3.7 Measures of prediction accuracy

To evaluate the accuracy of the forecast data, some error measures are used to evaluate the forecast procedure. Four of the widespread errors that are used to evaluate precision are the RMSE, MAE, MAPE, and Theil inequality. All of these measures can be computed by using the following equations. Abbasi et al.[1]. Oroian [24]. RMSE is a standard error index statistic used to determine the difference between the predicted model values and those of the model observed, Lin et al[19];Nayak et al[22] and is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - y_t^*)^2} \tag{21}$$

MAE is measured utilizing a term-by-term comparison of the relative error in the variable's actual prediction and defined as:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - y_i^*| \tag{22}$$

where y_t and y_t^* indicate samples of current and predicted model data, respectively. The sample size is n . Khalil et al[16]

$$MAPE = 100\% \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - y_t^*}{y_t}\right| \tag{23}$$

where y_t and y_t^* are the observed and computed values, respectively, and n is the number of sets. Gkountakou and Papadopoulos [10]. Theil's inequality coefficient is a measure of the accuracy of regression and is defined as:

$$\text{Theil' } sU = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t^* - y_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t^*)^2 + \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t)^2 +}}} \tag{24}$$

4.Results and discussion

This study used statistical models such as the artificial neural network (ANN) model, the adaptive neural fuzzy inference system (ANFIS) model, and the nonlinear auto-regressive with exogenous inputs model (NARX). In addition,the following hybrid models (ANFIS-NARX) and (NARX-ANN) to predict the profits of the insurance activity represent the important indicator of the good performance of Egypt's 39 insurance companies in the period from 1st January 2009 to 31 December 2022 per month. This prediction was based on the following factors (net premiums (NP), reinsurance commissions (RC), net income from investments (NIFINV), net compensation (NC), commissions of production cost (CPC), general and administrative expenses (GAE) that help decision makers to make appropriate decisions. The current study data have been analyzed using MATLAB R2014a. The previous statistical models have been compared at the level of people's insurance companies.

4.1 Measures of prediction accuracy of persons insurance companies

Table 1 shows that according to the RMSE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (15.74, 12.23) ,the (NARX-ANN) model comes in the second rank where its value reached (46.79, 30.08) ,then the (NARX) model comes in the third rank where its value reached (64.17, 55.09), then the (ANFIS) comes in the fourth rank where its value reached (572.71, 1377253), the (ANFIS-NARX) model comes in the last rank and Its values reached (649.13, 137636).

Table 1: Prediction accuracy measures of persons insurance companies

		RMSE	MAE	MAPE	Theil Ineq.	R^2
ANFIS	Training	572.7096	254.2805	0.03073	0.00037	0.7889
	Testing	1377253	1238057	51.00509	0.400284	0.4046
NARX	Training	64.17328	32.63138	0.008537	$4.02E-05$	0.7966
	Testing	55.09453	21.5501	0.000973	$1.17E-05$	0.5912
ANN	Training	15.74298	11.81158	0.005578	$1.02E-05$	0.7880
	Testing	12.23064	9.554831	0.00043	$2.62E-06$	0.6080
ANFIS-NARX	Training	649.1275	263.1972	0.056035	0.000409	0.7964
	Testing	137635.6	26889.04	2.893439	0.015377	0.4488
NARX-ANN	Training	46.79394	25.73497	0.005743	$2.94E-05$	0.8013
	Testing	30.08226	19.07486	0.000847	$6.41E-06$	0.5233

Then according to the MAE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (11.81, 9.55),the (NARX-ANN) model comes in the second rank where its value reached (25.73, 19.08),then the (NARX) model comes in the third rank where its value reached (32.63, 21.55), then the (ANFIS) comes in the fourth rank where its value reached (254.2805, 1238057), the (ANFIS-NARX) model comes in the last rank and Its values (263.197, 26889.04). Then according to the MAPE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (0.005578, 0.00043),the (NARX-ANN) model comes in the second rank where its value reached (0.005743, 0.000847),then the (NARX) model comes in the third rank where its value reached (0.008537, 0.000973), then the (ANFIS) comes in the fourth rank where its value reached (0.03073 , 51.00509), the (ANFIS-NARX) model comes in the last rank and Its values (0.056035, 2.893439). The Theil Inequality criterion shows that the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (1.02E-05 , 2.62E-06),the (NARX-ANN) model comes in the second rank where its value reached (2.94E-05 , 6.41E-06),then the (NARX) model comes in the third rank where its value reached (4.02E05 , 1.17E05), then the (ANFIS) comes in the fourth rank where its value reached (0.00037,0.400284), the (ANFIS-NARX) model comes in the last rank and Its values (0.000409, 0.015377). The (R^2) criterion measures the explanatory ability for the models. The model (ANN) comes in the first rank in the training and testing phases where it achieved the highest values reached (0.79, 0.61) ,the (NARX) model comes in the second rank reached (0.80, 0.59), then the (NARX-ANN) model comes in the third rank reached (0.80, 0.5233). The (ANFIS) model comes in fourth rank where its values reached (0.80, 0.41), the (ANFIS-NARX) model comes in the last rank where its values reached (0.80, 0.44). It is clear from the previous presentation according to predictive ability standards that the best model is ANN and then (NARX), (NARX-ANN),(ANFIS) and (ANFIS-NARX) respectively.

4.2 Analysis of ANN model of persons insurance companies

Figure 3 shows the architecture of the (ANN) developed in this study. This structure includes the input layer, hidden layers, and output layer. The networks were trained many times in an automatic way and stop when the best network is obtained. The best network obtained from insurance companies of persons was ANN (7-8-1). The researcher evaluated the accuracy of the model by using MSE and R-value to determine the degree of association between the predicted and actual values. In figure. 3 shows the values of the mean squares error and coefficient correlation that were $(247.8, 0.89)$ in training phase. In testing phase the values of the MSE and R-value that are $(149.59, 0.84)$ respectively. It is clear from the previous data that there is a direct strong correlation relationship between the predicted and actual values.

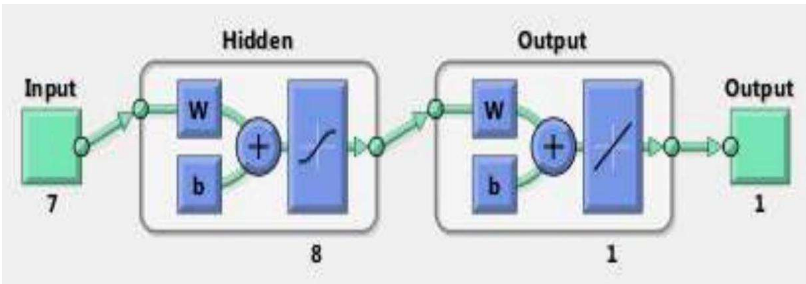

Figure 3. Structure of the FFNN model of persons insurance companies

Table 2 shows that the study depends on 168 periods that divided into 113 periods in the training phase and 55 periods in the testing phase. Also, the table 2 shows the impacts of independent variables on the dependent variable (Profit of insurance companies for persons) were as follows the impact of net premiums (NP) was 0.142, the impact of reinsurance commissions (RC) was 0.086, the impact of net income from earmarked investments (NIFINV) was 0.226 , the impact of net compensation (NC) was 0.033, the impact of production cost commissions (CPC) was 0.190 and the impact of general and administrative expenses (GAE) was 0.323. According to the importance relatively, the general and administrative expenses variable (GAE) achieved 100% so it came in the first rank, the net income from earmarked investments (NIFINV) variable achieved 70.1% so it came in the second rank, then the production cost commission (CPC) variable achieved 59% so it came in the third rank, the Net premiums (NP) achieved 44% in the fourth rank, then the reinsurance commissions (RC) and net compensation (NC) variables came in fifth and sixth ranks and achieved (26.5%, 10.4%) respectively.

Table 2: Summary model and independent variables importance

Sample	N	Percent
Training	113	66.7%
Testing	55	33.3%
Valid	168	100%
Excluded	0	
Total	168	
Variables	Importance	Normalized Importance
NP	.142	44%
RC	.086	26.5%
NIFINV	.226	70.1%
NC	.033	10.4%
CPC	.190	59%
GAE	.323	100%

Figure 4 shows the performance graph of the neural network model that was created during its training. The training phase stopped after 1000 epochs because the validation error increased. It is a useful diagnostic tool to plot the training, validation, and test errors to check the progress of training. The result shows a good network performance because the test set error and the validation set error have similar characteristics, and it doesn't appear any significance over fitting has occurred. After initial training phase of neural network model,it was retrained for 1000 epochs and the performance MSE was obtained 514.06 in training phase.

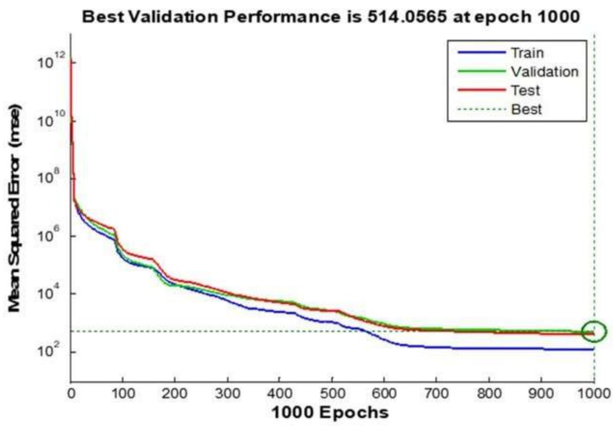

Figure 4. The performance plot of persons insurance companies

There are six input parameters into the network and only one output parameter. Different networks with different numbers of hidden neurons were used; the number of neurons varied from 5 to 30. For training the networks, the input vectors and target vectors have been randomly divided into three sets as follows: 70% used for training, 15% used to validate that the network is generalizing and to stop training before overfitting, and remaining 15% used as a completely independent test of network generalization. Based on figure 5, the value of R equals 1. This means that the output value produced by the network is closely similar to the target value. So, the model is satisfactory.

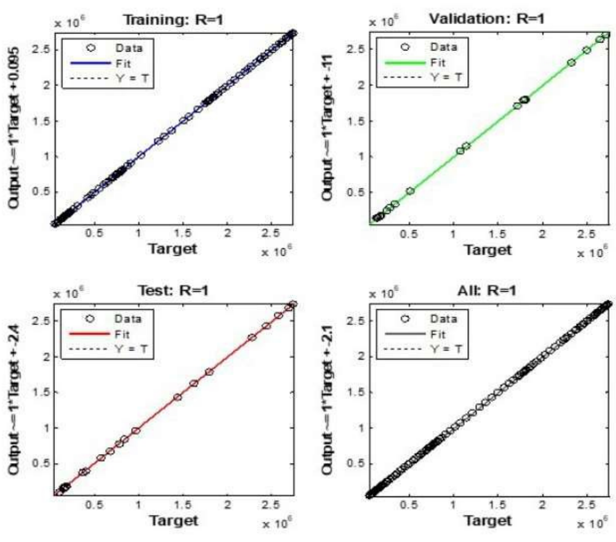

Figure 5. Regression plot of persons insurance companies

4.3 The profitability values of persons insurance companies in training and testing phases

Table 3 and figure 6 show that profitability in the insurance companies of persons in the training phase. We note that the target values for the profitability increased from 50340 in January 2009 and reached to 155270 in March 2010. As well as the output values of these companies increased from 50341 in January 2009 and reached to 155259 in March 2010 , then the error was slight reached to (-1) in January 2009 and 11 in March 2010. Then the target values fell from 154582 in April 2010 to 139546 in April 2011. the output values of these companies decreased from 154564 in April 2010 to 139531 in April 2011 , then the error between the target and the output values was slightly reached (15). After that the targets increased another time from 141262 at May 2011 to 1804213 in September 2018. The output values of these companies' profitability also increased from 141221 on May 2011 and reached 1804208 in September 2018. We note that the error between the target and output values also was slightly at (41 , 4) in May 2011 and September 2018.

Table 3: The profitability values of persons insurance companies in the training Phase

Time	Target	Output	Error	Time	Target	Output	Error	Time	Target	Output	Error
2009M01	50340	50341	-1	2012M05	172634	172645	-11	2015M09	746904	746898	5
2009M02	62620	62615	5	2012M06	174450	174461	-11	2015M10	758614	758604	10
2009M03	74649	74657	-8	2012M07	175984	175989	-5	2015M11	767631	767623	7
2009M04	86303	86298	5	2012M08	177210	177205	5	2015M12	773785	773787	-2
2009M05	97456	97446	10	2012M09	178102	178084	18	2016M01	777113	777124	-11
2009M06	107982	107981	1	2012M10	178633	178602	31	2016M02	778460	778472	-12
2009M07	117756	117761	-5	2012M11	178776	178737	39	2016M03	778874	778880	-6
2009M08	126654	126654	0	2012M12	178505	178475	30	2016M04	779404	779401	3
2009M09	134549	134544	5	2013M01	177849	177855	-7	2016M05	781097	781090	7
2009M10	141316	141315	2	2013M02	177054	177087	-33	2016M06	785002	784998	4
2009M11	146831	146835	-4	2013M03	176422	176453	-31	2016M07	792168	792169	-1
2009M12	150967	150971	-4	2013M04	176256	176264	-7	2016M08	803642	803644	-2
2010M01	153652	153658	-6	2013M05	176858	176840	18	2016M09	820473	820472	1
2010M02	155024	155024	0	2013M06	178529	178499	30	2016M10	843709	843705	4
2010M03	155270	155259	11	2013M07	181571	181549	23	2016M11	874398	874400	-1
2010M04	154582	154564	18	2013M08	186287	186284	3	2016M12	913589	913597	-8
2010M05	153148	153136	12	2013M09	192978	192994	-16	2017M01	961892	961887	5
2010M06	151158	151160	-2	2013M10	201947	201967	-20	2017M02	1018170	1018157	13
2010M07	148802	148815	-13	2013M11	213495	213500	-5	2017M03	1080846	1080841	5
2010M08	146268	146281	-14	2013M12	227924	227905	19	2017M04	1148345	1148352	-7
2010M09	143747	143747	0	2014M01	245428	245402	26	2017M05	1219092	1219102	-10
2010M10	141427	141410	18	2014M02	265765	265747	18	2017M06	1291511	1291514	-3
2010M11	139499	139476	23	2014M03	288584	288584	0	2017M07	1364026	1364020	6
2010M12	138152	138146	6	2014M04	313537	313556	-19	2017M08	1435062	1435053	9
2011M01	137531	137550	-19	2014M05	340272	340305	-33	2017M09	1503044	1503038	5
2011M02	137606	137633	-27	2014M06	368439	368474	-35	2017M10	1566395	1566395	0
2011M03	138303	138315	-13	2014M07	397689	397713	-24	2017M11	1623540	1623543	-3
2011M04	139546	139531	15	2014M08	427670	427676	-6	2017M12	1672904	1672911	-7
2011M05	141262	141221	41	2014M09	458034	458024	10	2018M01	1713319	1713331	-12
2011M06	143377	143324	53	2014M10	488429	488418	11	2018M02	1745246	1745253	-6
2011M07	145815	145770	45	2014M11	518505	518511	-6	2018M03	1769558	1769552	5
2011M08	148503	148483	20	2014M12	547913	547935	-22	2018M04	1787123	1787111	12
2011M09	151365	151374	-8	2015M01	576332	576344	-11	2018M05	1798812	1798804	8
2011M10	154329	154353	-24	2015M02	603565	603560	5	2018M06	1805495	1805498	-3
2011M11	157318	157334	-16	2015M03	629444	629433	11	2018M07	1808043	1808053	-10
2011M12	160259	160251	8	2015M04	653802	653796	6	2018M08	1807325	1807333	-8
2012M01	163085	163063	23	2015M05	676470	676475	-4	2018M09	1804213	1804208	4
2012M02	165763	165743	20	2015M06	697282	697293	-11	2018M10	1799575	1799560	16
2012M03	168264	168256	8	2015M07	716070	716081	-11				
2012M04	170564	170567	-4	2015M08	732667	732670	-4				

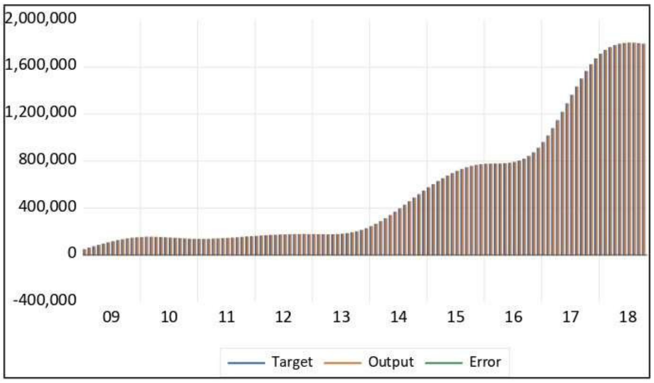

Figure 6. The profitability values of persons insurance companies in training phase

Table 4 and figure 7 show that profitability in the insurance companies of persons in the training phase . We note that the target values for the profitability of persons insurance companies in the training phase decreased from 1794283 in November 2018 and reached to 1788707 on May 2019.

Table 4: The profitability values of persons insurance companies in testing phase

Time	Target	Output	Error	Time	Target	Output	Error
2018M11	1794283	1794270	13.43364	2020M12	2589304	2589308	-4.38662
2018M12	1789207	1789211	-3.80569	2021M01	2628739	2628737	1.378739
2019M01	1785146	1785159	-12.749	2021M02	2662180	2662182	-2.08965
2019M02	1782618	1782625	-6.43134	2021M03	2689767	2689778	-10.3809
2019M03	1782071	1782066	5.050163	2021M04	2711640	2711651	-10.413
2019M04	1783951	1783942	9.683799	2021M05	2727938	2727939	-1.52808
2019M05	1788707	1788704	2.52217	2021M06	2738800	2738794	5.763361
2019M06	1796785	1796796	-11.3641	2021M07	2744366	2744363	2.796613
2019M07	1808632	1808653	-20.8815	2021M08	2744775	2744783	-7.60807
2019M08	1824696	1824713	-17.159	2021M09	2740167	2740180	-13.3718
2019M09	1845424	1845425	-1.26273	2021M10	2730681	2730687	-5.42002
2019M10	1871264	1871250	13.93279	2021M11	2716457	2716448	9.565355
2019M11	1902662	1902650	12.00481	2021M12	2697634	2697626	8.058029
2019M12	1940066	1940073	-6.93415	2022M01	2674390	2674404	-14.1639
2020M01	1983674	1983687	-13.2994	2022M02	2647059	2647074	-15.4054
2020M02	2032686	2032690	-3.60132	2022M03	2616011	2616010	0.554541
2020M03	2086054	2086047	7.567876	2022M04	2581618	2581605	13.07709
2020M04	2142729	2142719	9.295976	2022M05	2544253	2544242	11.49229
2020M05	2201661	2201660	1.200868	2022M06	2504286	2504287	-0.31832
2020M06	2261802	2261810	-7.66991	2022M07	2462090	2462100	-10.2274
2020M07	2322104	2322111	-6.74948	2022M08	2418036	2418043	-7.24006
2020M08	2381517	2381511	6.04905	2022M09	2372495	2372484	10.82225
2020M09	2438992	2438972	20.08586	2022M10	2325839	2325805	33.64944
2020M10	2493481	2493461	20.13361	2022M11	2278440	2278400	39.29058
2020M11	2543934	2543930	4.702927	2022M12	2230669	2230670	-1.1687

As well as the output values of these companies decreased from 1794270 in November 2018 and reached to 1788704 on May 2019 . Then the error was slight reached to (13.4) in November 2018 and (2.5) in March 2019.

Then the target values increased from 1796785 at June 2019 to 1824696 in August 2019. The output values of these companies also increased from 1796796 June 2019 to 1824713 in August 2019 , then the error between the target and the output values was slightly reached (-11.4 ,- 17.2) respectively. After that the target values increased another time from 1845424 in September 2019 to 230669 in August 2021. The output values of these companies' profitability also increased from 1845425 at September 2019 to 2744783 in August 2021 . The error between the target and output values was slightly where reached (-1.3,-7.61) in September 2019 and August 2021 respectively. Finally ,The target and output values decreased until November 2022 .The previous statistical models have been compared at the level of property insurance companies. The results showed that the ANN model is given better results when predicted than the following models: ANFIS, NARX, ANFIS-NARX, and NARX-ANN, because it shows lower values for the following predictive accuracy measures such as (RMSE, MAPE, MAE and Theil inequality).

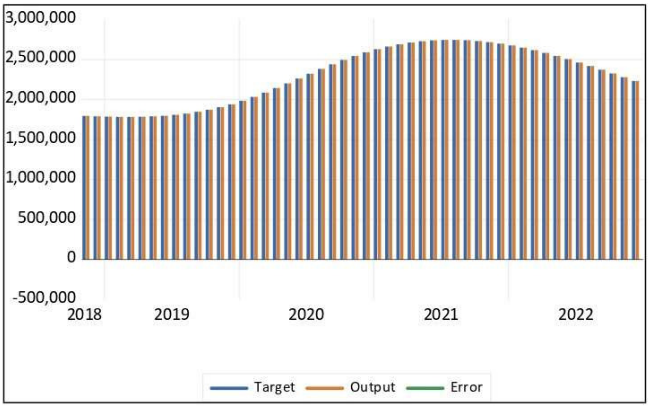

Figure 7. The profitability values of persons insurance companies in testing phase

4.4 Measures of prediction accuracy of property insurance companies

Table 5 shows that according to the RMSE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (56.73, 52.79),then the (NARX-ANN) model comes in the second rank where its value reached (143.12, 177.99),then the (NARX) model comes in the third rank where its value reached (212.31, 229.43), then the (ANFIS) comes in the fourth rank where its value reached (719.92, 1806225), the (ANFIS-NARX) model comes in the last rank and Its values reached (9942.083, 91309.0).Then according to the MAE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (38.01, 43.22),the (NARX-ANN) model comes in the second rank where its value reached (72.75, 90.49),then the (NARX) model comes in the third rank where its value reached (118.62, 105.68), then the (ANFIS) comes in the fourth rank where its value reached (511.07, 1680148), the (ANFIS-NARX) model comes in the last rank and Its values (1794.503, 13365.44). Then according to the MAPE criterion the (ANN) model comes in the first rank in the training and testing phases where it achieved (0.01998, 0.00156),then the (NARX) rank comes in the second rank where its value reached (0.0318, 3.4E-07),and the (NARX-ANN) model comes in the third rank where its value reached (0.04018, 0.82608), then the (ANFIS-NARX) model comes in the fourth rank and Its values (0.11325, 0.82608). Then the (ANFIS) comes in the last rank where its value reached (8.8E-06, 52.6835), The Theil Inequality criterion shows that the (ANN) model comes in the first rank in the training and testing phases where it achieved the lowest values reached (2.08E-05, 8.31E-06),the (NARX-ANN) model comes in the second rank where its value reached (5.17E-05, 2.79E-05),then the (NARX) rank comes in the third rank where its value reached (7.59E-05, 3.57E-05), then the (ANFIS) comes in the fourth rank where its value reached (0.00037, 0.391369), the (ANFIS-NARX) model comes in the last rank and Its values (0.003602, 0.031994).

Table 5: Measures of prediction accuracy of property insurance companies

		RMSE	MAE	MAPE	Theil Ineq.	R^2
ANFIS	Training	719.9153	511.0683	8.8E − 06	0.000264	0.8265
	Testing	1806225	1680148	52.6835	0.391369	0.374
NARX	Training	212.3035	118.6147	0.0318	7.59E − 05	0.826
	Testing	229.4334	105.6852	3.4E − 07	3.57E − 05	0.646
ANN	Training	56.7292	38.0119	0.01998	2.08E − 05	0.8265
	Testing	52.79063	43.21731	0.00156	8.31E − 06	0.6776
ANFIS-NARX	Training	9942.083	1794.503	0.11325	0.003602	0.8205
	Testing	91309.07	13365.44	0.82608	0.031994	0.4415
NARX-ANN	Training	143.1299	72.75424	0.04018	5.17E − 05	0.8268
	Testing	177.9952	90.49099	0.00342	2.79E − 05	0.6636

The (R^2) criterion measures the explanatory ability for the models. The model (ANN) comes in the first rank in the training and testing phases where it achieved the highest values reached (0.83, 0.618) ,the (NARX) model comes in the second rank reached (0.826 , 0.6636), then the (NARX-ANN) model comes in the third rank reached (0.8268 , 0.6636). The (ANFIS) model comes in fourth rank where its values reached (0.8265, 0.374), the (ANFIS-NARX) model comes in the last rank where its values reached (0.8205, 0.4415). It is clear from the previous presentation according to predictive ability standards that the best model is ANN and then (NARX), (NARX-ANN),(ANFIS) and (ANFIS-NARX) respectively.

4.5 Analysis of ANN model of property insurance companies

Figure 8 shows the architecture of the (ANN) developed in this study. This structure includes the input layer, hidden layers, and output layer. The networks were trained many times in an automatic way and stops when the best network is obtained. The best network obtained for property insurance companies was ANN (7-8-1). The researcher evaluated the accuracy of the model by using MSE and R-value to determine the degree of association between the predicted and actual values. The MSE and R values that were $(3218.2, 0.91)$ respectively in the training phase. In testing phase, the values of the MSE and R are $(2786.9, 0.82)$ respectively. It is clear from the previous data that there is a direct strong correlation relationship between the predicted and actual values.

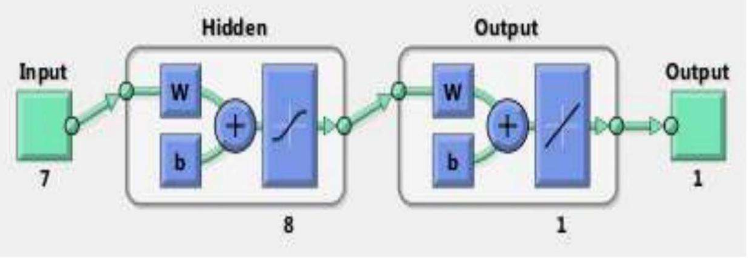

Figure 8 . Structure of the FFNN model for property insurance companies

Table 6 shows that the study depends on 168 periods that divided into 112 periods in the training phase and 56 periods in the testing phase. Also, the table 6 shows the impacts of independent variables on the dependent variable (Profit of insurance companies for persons) were as follows the impact of net premiums (NP) was 0.014, the impact of reinsurance commissions (RC) was 0.409, the impact of net income from earmarked investments (NIFINV) was 0.111, the impact of net compensation (NC) was 0.090, the impact of production cost commissions (CPC) was 0.133 and the impact of general and administrative expenses (GAE) was .243.

Table 6: Summary model and independent variables importance

Sample	N	Percent
Training	112	66.7%
Testing	56	33.3%
Valid	168	100%
Excluded	0	
Total	168	
Variables	Importance	Normalized Importance
NP	.014	3.4%
RC	.409	100%
NIFINV	.111	27.2%
NC	.090	22.1%
CPC	.133	32.5%
GAE	.243	59.5%

According to the importance relatively, reinsurance commissions (RC)achieved 100% so it came in the first rank, the general and administrative expenses variable (GAE) achieved 59.5% so it came in the second rank, then the production cost commission (CPC) variable achieved 32.5%, so it came in the third rank ,the net income from investments (NIFINV) variable achieved 27.2%, so it came in the fourth rank, the net compensation (NC) and net instalments (NP) came in fifth and sixth rank and achieved ((22.1%, 3.4%) respectively. Figure 9 shows the performance graph of neural network model that created during its training. The training phase stopped after 1000 epochs because the validation error increased. It is a useful diagnostic tool to plot the training, validation, and test errors to check the progress of training. The result shows a good network performance because the test set error and the validation set error have similar characteristics, and it does not appear any significance over fitting has occurred. After initial training phase of neural network model, it was retrained for 1000 epochs and the performance MSE was achieved 8616.06 in the training phase.

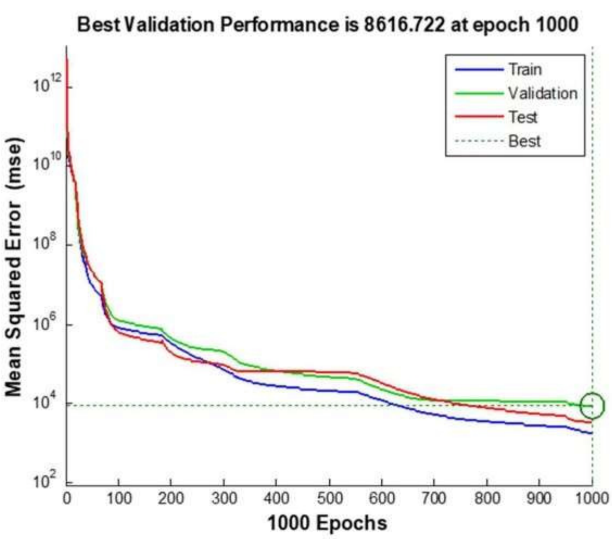

Figure 9. The performance plot of property insurance companies

There are six input parameters into the network and only one output parameter. The different networks with different numbers of hidden neurons were used; the number of neurons varied from 5 to 30. In the training of networks, the input vectors and target vectors have been randomly divided into three sets as follows: 70% used for training, 15% used to validate that the network is generalizing and to stop training before overfitting, and the remaining 15% used as a completely independent test of network generalization.

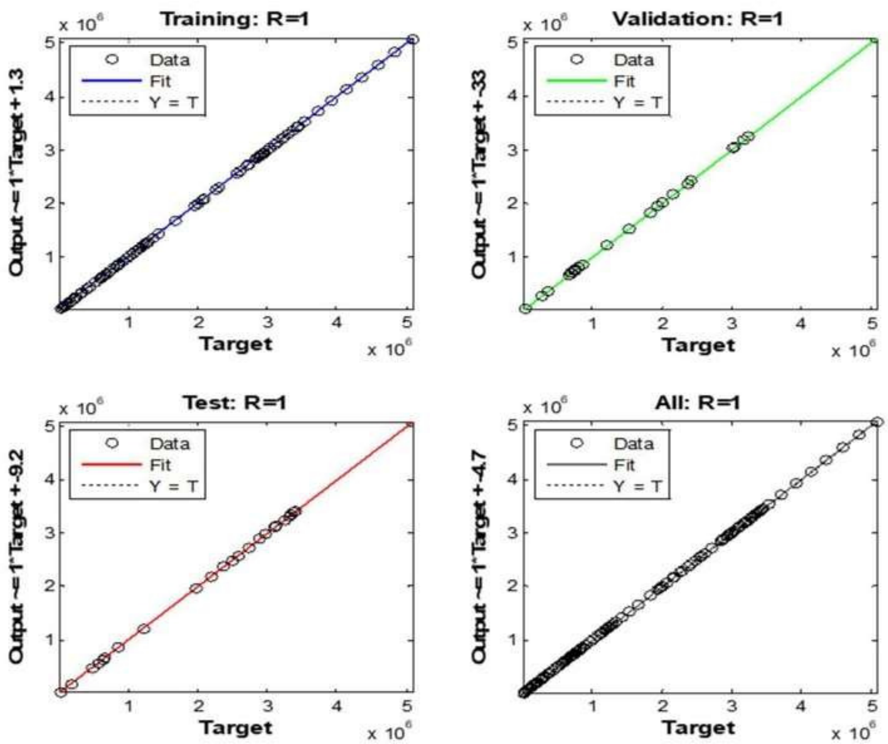

Figure 10. Regression plot of property insurance companies

Based on Figure 10, the value for R is equal to 1.This shows that the output produced by the network is closely similar to the target and that the model is satisfactory.

4.6 The profitability values of property insurance companies in training and testing phases

Table 7 and figure 11 show that the profitability in the insurance companies of property in the training phase. We note that the target values for the profitability decreased from 24701 in January 2009 and reached to 15662 in July 2009. As well as the output values of these companies decreased from 24674 to 15642 in July 2009, then the error was slight reached to (26) in January 2009 and (21) in July 2009. The target values increased from 15991 in August 2009 to 877822 in December 2012. The output values of these companies increased from 16008 in August 2009 to 877822 in December 2012, then the error between the target and the output values was slightly reached (-16, 19) respectively.

Table 7: The profitability values of property insurance companies in training phase

Time	Target	Output	Error	Time	Target	Output	Error	Time	Target	Output	Error
2009M01	24701	24674	26	2012M05	702891	702905	-14	2015M09	1289082	1289074	9
2009M02	22375	22408	-33	2012M06	745536	745539	-3	2015M10	1288241	1288233	8
2009M03	20246	20277	-32	2012M07	784334	784335	-1	2015M11	1282363	1282395	-32
2009M04	18411	18411	1	2012M08	818063	818070	-6	2015M12	1271427	1271480	-53
2009M05	16970	16938	32	2012M09	845499	845512	-12	2016M01	1255873	1255836	37
2009M06	16021	15981	41	2012M10	865420	865430	-10	2016M02	1237990	1237922	68
2009M07	15662	15642	21	2012M11	876602	876599	3	2016M03	1220525	1220519	6
2009M08	15991	16008	-16	2012M12	877822	877803	19	2016M04	1206229	1206281	-52
2009M09	17107	17154	-47	2013M01	868386	868361	25	2016M05	1197851	1197886	-35
2009M10	19108	19159	-52	2013M02	849712	849693	19	2016M06	1198140	1198105	34
2009M11	22092	22116	-23	2013M03	823747	823743	4	2016M07	1209845	1209772	73
2009M12	26158	26147	11	2013M04	792439	792453	-14	2016M08	1235716	1235682	34
2010M01	31382	31369	13	2013M05	757733	757763	-30	2016M09	1278502	1278538	-37
2010M02	37752	37749	3	2013M06	721577	721615	-39	2016M10	1340952	1340987	-35
2010M03	45236	45238	-2	2013M07	685917	685952	-35	2016M11	1425816	1425746	70
2010M04	53798	53801	-3	2013M08	652702	652718	-16	2016M12	1535843	1535739	104
2010M05	63406	63409	-3	2013M09	623877	623861	16	2017M01	1672290	1672354	-64
2010M06	74026	74029	-3	2013M10	601389	601335	54	2017M02	1830444	1830510	-66
2010M07	85625	85626	-1	2013M11	587186	587106	80	2017M03	2004101	2004051	50
2010M08	98168	98165	3	2013M12	583214	583152	62	2017M04	2187055	2186965	90
2010M09	111623	111617	6	2014M01	590860	590878	-19	2017M05	2373102	2373079	23
2010M10	125955	125955	0	2014M02	609267	609356	-89	2017M06	2556036	2556078	-41
2010M11	141131	141149	-18	2014M03	637019	637120	-101	2017M07	2729654	2729680	-26
2010M12	157118	157150	-32	2014M04	672699	672748	-48	2017M08	2887749	2887718	31
2011M01	173925	173934	-9	2014M05	714891	714844	47	2017M09	3024117	3024069	47
2011M02	191735	191719	16	2014M06	762178	762028	150	2017M10	3132553	3132543	10
2011M03	210774	210752	23	2014M07	813144	812919	225	2017M11	3206852	3206871	-19
2011M04	231269	231253	16	2014M08	866371	866129	242	2017M12	3240809	3240817	-8
2011M05	253445	253441	4	2014M09	920443	920255	188	2018M01	3230391	3230391	0
2011M06	277530	277539	-9	2014M10	973944	973872	72	2018M02	3180248	3180235	12
2011M07	303749	303772	-23	2014M11	1025456	1025521	-65	2018M03	3097202	3097185	17
2011M08	332328	332363	-35	2014M12	1073564	1073707	-143	2018M04	2988075	2988083	-7
2011M09	363493	363530	-37	2015M01	1117083	1117175	-92	2018M05	2859690	2859719	-29
2011M10	397472	397488	-16	2015M02	1155759	1155761	-2	2018M06	2718869	2718860	8
2011M11	434489	434462	28	2015M03	1189570	1189516	54	2018M07	2572433	2572344	89
2011M12	474772	474709	63	2015M04	1218496	1218437	59	2018M08	2427204	2427093	111
2012M01	518305	518269	36	2015M05	1242514	1242486	28	2018M09	2290005	2289993	12
2012M02	564106	564113	-7	2015M06	1261603	1261612	-8	2018M10	2167658	2167765	-107
2012M03	610953	610982	-29	2015M07	1275742	1275767	-24				
2012M04	657622	657650	-28	2015M08	1284909	1284922	-13				

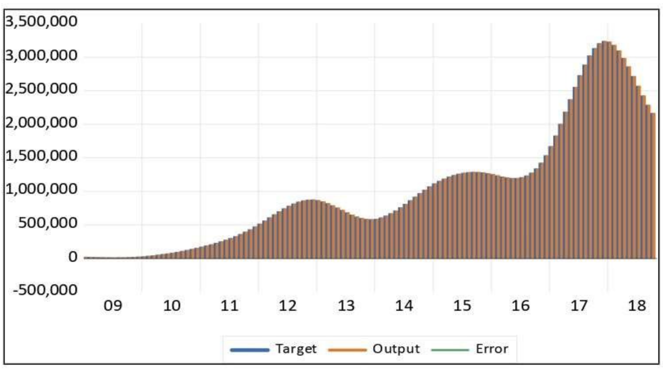

Figure 11. The profitability of property insurance companies in training phase

After that the target values decreased another time from 868386 in January 2013 to 583214 in November 2013. The output values of these companies' profitability also decreased from 868361 in January 2013 and reached 583152 in November 2013.

We note that the error between the target and output values also was slightly at (25 , 62) in January 2013 and November 2013.Also, we found that the target values increased from 590860 in January 2014 to 3097202 in March 2018. The output values of these companies' profitability also decreased from 590878 at January 2014 and reached 3097185 in March 2018. We note that the error between the target and output values also was slightly (-19 , 17) in January 2014 and March 2018 respectively. Finally, we noted that the target and output values decreased from April 2018 to October 2018 and the difference between them was slightly.

Table 8: The profitability values of property insurance companies in testing phase

Time	Target	Output	Error	Time	Target	Output	Error
2018M11	2066984	2067048	-63.8306	2020M12	3410300	3410357	-57.339
2018M12	1994806	1994731	75.43774	2021M01	3366984	3367035	-50.9166
2019M01	1956214	1956248	-34.1876	2021M02	3310757	3310765	-7.57158
2019M02	1949372	1949511	-139.309	2021M03	3245673	3245638	34.65982
2019M03	1970711	1970844	-133.144	2021M04	3175784	3175721	62.58615
2019M04	2016664	2016739	-75.0218	2021M05	3105143	3105073	69.99031
2019M05	2083663	2083686	-23.1224	2021M06	3037804	3037750	54.26243
2019M06	2168140	2168132	8.254359	2021M07	2977819	2977799	19.73997
2019M07	2266526	2266497	28.84469	2021M08	2929240	2929260	-19.4654
2019M08	2375254	2375211	43.39636	2021M09	2896121	2896163	-42.0193
2019M09	2490755	2490714	41.88893	2021M10	2882515	2882545	-29.2407
2019M10	2609463	2609450	12.68839	2021M11	2892475	2892457	18.21808
2019M11	2727808	2727841	-33.5763	2021M12	2930053	2929994	58.75551
2019M12	2842222	2842268	-46.0034	2022M01	2998187	2998154	33.06636
2020M01	2949601	2949582	19.51732	2022M02	3095353	3095362	-9.14833
2020M02	3048693	3048619	74.1819	2022M03	3218912	3218945	-32.8173
2020M03	3138710	3138642	68.23527	2022M04	3366224	3366258	-33.2481
2020M04	3218861	3218851	10.42135	2022M05	3534651	3534668	-16.9051
2020M05	3288360	3288422	-61.8404	2022M06	3721552	3721546	6.617773
2020M06	3346416	3346523	-107.334	2022M07	3924290	3924263	26.33046
2020M07	3392241	3392343	-101.243	2022M08	4140223	4140193	30.50158
2020M08	3425047	3425095	-47.9135	2022M09	4366714	4366703	10.6959
2020M09	3444045	3444027	18.18313	2022M10	4601122	4601150	-27.9699
2020M10	3448445	3448398	47.55544	2022M11	4840810	4840859	-49.6681
2020M11	3437460	3437451	9.341982	2022M12	5083136	5083101	34.65878

Table 8 and figure 12 show that profitability in the insurance companies of property in the testing phase.

We note that the target values for the profitability decreased from 2066984 in November 2018 and reached to 1970711 in March 2019. As well as the output values of these companies decreased from 2067048 in November 2018 and reached to 190844 in March 2019 . Then the error was slight reached (-64) at November 2018 and (-133) in March 2019. Then the target values increased from 2016664 in April 2019 to 3410300 in December 2020.

The output values of these companies also increased from 2016739 in April 2019 to 3410357 in December 2020 , then the error between the target and the output values was slightly (-75 , -57) in April 2019 and December 2020 respectively. Finally, we noted that the target and output values decreased from January 2021 to December 2022, after that they increased from January 2022 to December 2022 and their errors were slightly in the same period.

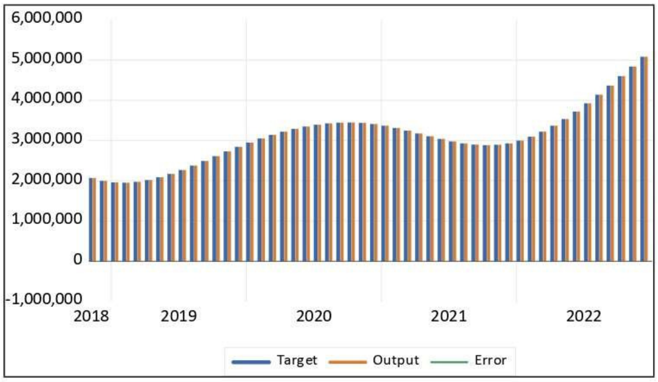

Figure 12. The profitability of property insurance companies in testing phase

5. Conclusion

This study aimed to predict the total profits of insurance companies for people and property from 1 January 2009 to 31 December 2022 per month. The study used the following models (ANN), (ANFIS), (NARX), (ANFIS-NARX) and (NARX-ANN). It relied on the following variables (net premiums (NP), reinsurance commissions (RC), net investment income (NIFINV), net compensation (NC), production cost commissions (CPC) and general and administrative expenses (GAE). The following is the results of this study:

• The results of this study showed the higher of the relative importance of the following variables (general and administrative expenditure (GAE) that achieved 100%, net investment income (NIFINV) that achieved 70.1%, production cost commissions (CPC) that achieved 59%, but the remaining variables such as reinsurance commissions (RC), net compensation (NC), net premiums (NP)) their relative importance were lower than 50%. So we will Reject the first hypothesis that there is a significant effect between net premiums (NP), net compensation (NC), net investment income (NIFINV) and the profits of insurance companies on persons, and there is an insignificant effect between commissions of production cost (CPC), general and administrative expenses (GAE), reinsurance commissions (RC) the profits of insurance companies on persons.

• The results of this study showed also the higher of the relative importance of the following variables (reinsurance commissions (RC) that achieved 100%, general and administrative expenditure (GAE) that achieved 59.5% but the remaining variables such as net investment income (NIFINV), commissions of production cost (CPC), net compensation (NC), net premiums (NP)) their relative importance were lower than 50%. So we will reject the second hypothesis that there is a significant effect between net premiums (NP), net compensation (NC), net investment income (NIFINV) and the profits of insurance companies on property, and there is an insignificant effect between commissions of production cost (CPC), general and administrative expenses (GAE), reinsurance commissions (RC) the profits of insurance companies on property.

• The predictive and interpretive capability criteria showed that the best model is ANN then (NARX-ANN), (NARX), (ANFIS) and (ANFIS-NARX) respectively in these companies. So we will reject of the third hypothesis that the predictive and explanatory ability of hybrid models such as (ANFIS, ANFIS-NARX, and NARX-ANN) will be expected to be better than non-hybrid models such as (ANN), (NARX) in insurance companies of persons and property.

• The results of this study showed that when using the neural network model and comparing target and output values that there were fluctuations between the rise and the decline in profits in insurance companies for persons and property, between January 2019 and December 2022. These may be due not only to the internal factors of the organization but also to local economic changes and global crises such as the COVID-19 crisis and the Ukrainian-Russian crisis which affected on the Egyptian economy and the economies of many developed and developing countries.

• Through the results of the first and second hypotheses , the researcher recommends the insurance companies of persons to reduce their general and administrative expenses and increase its investments so that net income increases, and it takes care of production cost commissions. Also, the researcher recommends the insurance companies of property to increase reinsurance commissions, and reduces its general and administrative expenses, in addition to increasing its investments so as to increase the net income of these investments.

• The study also recommends that based on fluctuations found in the financial profits of these companies use GARCH family models that can study these fluctuations and their causes more accuracy.

• The researcher recommends encouraging researchers to make further studies to discover new factors that affect the profitability of insurance companies and interest to local and global economic changes with the use of the dummy variables.

•The researcher also recommends presenting and using the results of this study to all companies in the insurance sector in the Egyptian market and the Financial Regulatory Authority (FRA) to benefit from them.

Future works

From the results of the current study, the researcher proposes some of the following future studies: (1) Using GARCH family models to study the fluctuations in the profitability of Egyptian insurance companies. (2) Using the fuzzy robust regression model to measure the impact of internal and global economic variables on the profitability of Egyptian insurance companies. (3) Predicting insurance companies' financial crises using Recurrent neural network and Long Term short memory models. (4) Using non-linear fuzzy models to predict the rate of loss in insurance companies.

Acknowledgement

I would like to thank my mother, my father, my brother and professor Hayam Wahba, professor of finance and investment, Vice dean for postgraduate studies and researches, Faculty of business, Ain Shams University for their constant support. I would also like to thank the editor and the referees for their suggestions and corrections to improve the quality of this paper.

REFERENCES

1. Abbasi, D., Ashrafzadeh, A., Asheghi, R., and Merufinia, E. *A survey on the performance of fuzzy-neural network at predicting the average monthly discharge of catchment basin areas having snow regimes*. Journal of Civil Engineering and Urbanism, 4, pp480–484, http://www.ojceu.ir/main,2014.
2. Ahmeti, Y., and Iseni, E. *Factors Affecting Profitability of Insurance Companies:Evidence from Kosovo*. Academicus International Scientific Journal,25(25):122-142,https://doi.org/10.7336/academicus.2022.25.08,2022.
3. Al-Dwiry, M., Al Shaher,T., Al-dweeri, M.,and Al-Hoi,A. *Major factors influencing insurance companies in the Middle East*. Academy of Accounting and Financial Studies Journal,Vol.24,Iss.2,pp1-9,2020.
4. Anam, M., Putra, R., Fitriah, Z., Yanti, I., Hidayat, N., and Mahanani, D. *Health claim insurance prediction using support vector machine with particle swarm optimization*. BAREKENG: Journal of Mathematics and Its Applications,Vol.17, Issue 2, https://doi.org/10.30598/ barekengvol17iss2, pp0797-0806,2023.
5. Basheer, I.,and Hajmeer, M. *Artificial neural networks: fundamentals, computing, design, and application*. Journal of Microbiol Methods,43,3–31, https://dx.doi.org/10.1016S0167-7012(00) 00201-3,2000.
6. Bisht, D., and Jangid, A. *Discharge Modelling using Adaptive Neuro - Fuzzy Inference System*. Journal of Advanced Science and Technology ,31 (1), 99-114, 2011.
7. Dhiab,L. *Determinants of Insurance Firms' Profitability: An Empirical Study of Saudi Insurance Market*. Journal of Asian Finance, Economics and Business, Vol. 8, No. 6, 0235–0243. https://doi.org/ 10.13106/ jafeb.2021.vol8.no6.0235,2021.
8. Esmaeelzadeh, S., Adib, A., and Alahdin, S. *Long-term streamflow forecasts by adaptive neuro-fuzzy inference system using satellite images and K-fold cross validation (case study: Dez, Iran)*. KSCE Journal of Civil Engineering, 1–9, https://doi.org/10.1007/s12205-014-0105-2,2014.

9. Faoziyyah, A., and Laila, N. *Internal factors and macroeconomic factors affecting the profitability of sharia insurance in Indonesia for 2016-2018 period*. Jurnal Ekonomi Syariah Teori dan Terapan, p-ISSN: 2407-1935, e-ISSN: 2502-1508,Vol. 7 No.6 , https://doi.org/10.20473/vol7,iss20206, pp1146-1163,2020.
10. Gkountakou, F. and Papadopoulos, B. *The use of fuzzy linear regression and ANFIS methods to predict the compressive strength of cement* . Symmetry, 12, 1295; https://doi.org/10.3390/sym12081295,MDPI,2020.
11. Haykin, S., and Lippmann, R. *Neural networks, A Comprehensive Foundation*. International Journal of Neural Systems,5, pp363–364, https://doi.org/10.1142/S0129065794000372,1994.
12. Jang, J. *ANFIS : Adaptive-Network-Based Fuzzy Inference System*. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665-685, https://doi.org/10.1109/21.256541,1993.
13. Jang, J., Sun, C., and Mizutani , E. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. First edition, New Jersey, USA, Prentice Hall,1997.
14. Kaur, J. and Bassi, P. *Predictive Performance of Indian Insurance Industry Using Artificial Neural Network (ANN) and Support Vector Machine (SVM): A Comparative Study, Prelims", Sood, K., Dhanaraj, R.K., Balusamy, B., Grima, S. and Uma Maheshwari, R. (Ed.) Big Data: A Game Changer for Insurance Industry*. (Emerald Studies in Finance, Insurance, and Risk Management), Emerald Publishing Limited, Leeds, pp.65-79, https://doi.org/10.1108/978-1-80262-605-620221005,2022.
15. Khadka, S. *Effect of firm specific and macroeconomic factors on profitability of Nepalese insurance companies*. The Lumbini Journal of Business and Economics,Vol.11, No.1, pp 232-251, https://doi.org/10.3126/ljbe.v11i1.54331,2023.
16. Khalil, A. , Liu, Z. and Ali, A. *Using an adaptive network-based fuzzy inference system model to predict the loss ratio of petroleum insurance in Egypt*. Risk Management and Insurance Review. Wiley, 1–14. https://doi.org/10.1111/rmir.12200,© 2022. The American Risk and Insurance Association,2022.
17. Klir, G., and Yuan, B. *Fuzzy sets and fuzzy logic: theory and applications*. Institutional Respostiory,2008 .
18. Lalon, R., and Das, J. *Investigating the dynamic impact of firm-specific and macroeconomic drivers on profitability of general insurance companies in Bangladesh*. International Journal of Research in Business and Social Science, Istanbul, Turkey, Vol.11, Iss.5, PP304-313, https://doi.org/10.20525/IJRBS.V11I5.1869, 2022.
19. Lin, J., Cheng, C., and Chau, K. *Using support vector machines for long-term discharge prediction*. Hydrological Sciences Journal, 51(4),599-612,https://doi.org/10.1623/hysj.51.4.599,2006.
20. Lupačov, J., and Stanković, Z. *Artificial intelligence in insurance companies*. Journal TTTP-Traffic and Transport: Theory and Practice, https://doi.org/10.7251/JTTTP2202077L,UDC: 005.334: [007.52:004. 85, TTTP (2022)7(2)77-81, 2022.
21. Mohammed, M., and Darus, I. *Pid controller for NARX and ANFIS models of marine pipe cylinder undergoes vortex induced vibration*. Journal of Theoretical and Applied Information Technology, Vol.66, No.1, pp 359-367, 2014.
22. Nayak, P., Sudheer, K., Rangan, D. , and Ramasastri, K. *A neuro-fuzzy computing technique for modeling hydrological time series*. Journal of Hydrology,291(1–2):pp52–66, https://doi.org/10.1016/j.jhydrol.2003.12.010, 2004.
23. Olarewaju, O., and Msomi, T. *Factors affecting the profitability of reinsurance companies in sub-Saharan Africa: Evidence from dynamic panel analysis*. Cogent Business and Management,9: 2093485, https:/doi.org/10.1080/23311975.2022. 2093485, pp1-17, 2022.
24. Oroian, M. *Influence of temperature , frequency and moisture content on honey viscoelastic parameters–Neural networks and adaptive neuro-fuzzy inference system prediction*.LWT- Food Science and Technology,63(2), pp1309–1316, https://doi.org/10.1016/j.lwt.2015.04.051,2015.
25. Peng, H., Ozaki, T., Toyoda, Y., Shioya, H., Nakano , K., Ozaki, V., and Mori, M. *RBF-ARX model based nonlinear system modeling and predictive control with application to a NO decomposition process*. Control Engineering Practice, Vol. 12, No. 2, pp.191–203, https://doi.org/10.1016/S0967-0661(03)00050-9,2004.
26. Qazi, A., Fayaz, H., Wadi, A., Raj, R., Rahim, N., and Khan, W. *The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review*. Journal of Cleaner Production, 104, 1–12. https://doi.org/10.1016/j.jclepro.2015.04.041, 2015.
27. Quej, V., Almorox, J., Arnaldo, J., and Saito., L. Manju, J., and Debaleena, C. *ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment*. Journal of Atmospheric and Solar-Terrestrial Physics, Elsevier,Vol. 155, pp 62-70, https://doi.org/10.1016/j.jastp.2017.02.002, 2017.
28. Rezrazi, A., Hanini, S., and Laidi, M. *An optimisation methodology of artificial neural network models for predicting solar radiation: a case study*. Theoretical and Applied Climatology,123,pp1–15, https://doi.org /10.1007/s00704-015-1398-x, 2015.
29. Sitharthan, R., and Rajesh, M. *Retracted Article: Application of machine learning (ML) and internet of things (IoT) in healthcare to predict and tackle pandemic situation*. Distributed Parallel Databases 40, 887 (2022). https://doi.org/10.1007/s10619-021- 07358-7, 2022.
30. Sonmez, A., Kale, S., Ozdemir, R., and Kadak , A. *An Adaptive Neuro-Fuzzy Inference System (ANFIS) to Predict of Cadmium (Cd) Concentrations in the Filyos River*. Turkish Journal of Fisheries and Aquatic Sciences, Turkey,18:1333-1343 (2018),www.trjfas. org, ISSN 1303-2712, https://doi.org/ 10.4194/1303-2712-v18-12-01, www.trjfas.org , 2018.
31. Stefenon, S., Freire, R., Coelho, L., Meyer, L., Grebogi, R., Buratto, W., and Nied, A. *Electrical Insulator Fault Forecasting Based on a Wavelet Neuro-Fuzzy System. Energies, MDPI, Energies 2020, 13, 484; https://doi.org/10.3390/en13020484, 2020*.
32. Vojinovi´c, Z., Milutinovi´c, S., Serti´c, D., and Lekovi´c, B. *Determinants of Sustainable Profitability of the Serbian Insurance Industry: Panel Data Investigation*. Sustainability, 14, 5190. https://doi.org /10.3390/su14095190, pp1-22, 2022.
33. Worku, A., Bayleyegne, Y., and Tafere, B. *Determinants of profitability of insurance companies in Ethiopia: evidence from insurance companies from 2011 to 2020 years*. Journal of Innovation and Entre-preneurship, Vol.13, No.4, Springer, https://doi.org/10.1186/s13731-023-00357-1, pp1-19, 2024.
34. Yadav, A., and Chandel, S. *Solar radiation prediction using artificial neural network techniques: a review, Renewable Sustainabilty*. Energy Review, 33, 772–781, https://doi.org/10.1016/j.rser.2013.08.055, 2014.