IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (3) Part (1): 216-229

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521
University of Kirkuk Journal For Administrative and Economic Science

Mohammed Ahmed Husham. Enhancing the prediction of artificial neural network algorithms based on Principal Component Analysis. *University of Kirkuk Journal For Administrative and Economic Science* (2025) 15 (3) Part (1):216-229.

Enhancing the prediction of artificial neural network algorithms based on Principal Component Analysis

Ahmed Husham Mohammed [1]

[1] *Department of Statistics, College of Administration and Economics, University of Basrah, Basrah, Iraq*

ahmed.albasrai@uobasrah.edu.iq [1]

Abstract: Statistical methods and artificial intelligence algorithms are one of the development methodologies that contribute to addressing many of the problems facing the data. Therefore, the paper tried to find an approach that contributes to addressing the algorithm of artificial neural networks and improving its work by adopting the method of the principal component analysis. This method addresses the problem of linear multiplicity between variables and reduces the problem of higher dimensions in a mathematical manner that preserves by finding new linear compounds and then adopting these linear compounds as inputs to artificial neural networks. Therefore, the paper was able to apply this methodology to the financial banking sector of the United Arab Emirates and determine estimates of financial performance in adopting variables characterized by a non-financial logistical nature (environmental, social and governance). The paper concluded that this methodology was able to improve the results of artificial neural networks and increase the accuracy of estimates of financial performance in the banking sector while retaining all variables and their effects through the new linear compounds.

Keywords: artificial neural networks, Estimation, multidimensional, Principal Component Analysis.

تحسين تقديرات خوارزمية الشبكات العصبية الاصطناعية باعتماد تحليل المركبات الرئيسية

م.د. احمد هشام محمد [1]

[1] *جامعة البصرة ـكلية الإدارة والاقتصاد- قسم الاحصاء، البصرة، العراق*

المستخلص: تعتبر الاساليب الاحصائية وخوارزميات الذكاء الاصطناعي احد المنهجيات التطويرية التي تسهم في معالجة الكثير من المشكلات التي تواجه البيانات لذا حاولت الورقة الى ايجاد منهج يسهم في معالجة خوارزمية الشبكات العصبية الاصطناعية وتحسين عملها باعتماد اسلوب المكونات الرئيسية كون هذا الاسلوب يعالج مشكلة التعدد الخطي بين المتغيرات ويقلل مشكلة الابعاد العليا بأسلوب رياضي يحافظ على من خلال ايجاد مركبات خطية جديدة ومن ثم اعتماد هذه المركبات الخطية كمدخلات في الشبكات العصبية الاصطناعية، لذا تمكنت الورقة من تطبيق هذه المنهجية على القطاع المالي المصرفي لدولة الامارات العربية المتحدة وتحديد تقديرات الاداء المالي في باعتماد متغيرات تتسم بطابع لوجستي غير مالي وهي الابعاد (البيئية، والاجتماعية والحوكمة) وتوصلت الورقة الى ان هذه المنهجية تمكنت من تحسين نتائج الشبكات العصبية الاصطناعية وزيادة دقة تقديرات الاداء المالي في القطاع المصرفي مع الاحتفاظ بجميع المتغيرات وتأثيراتها من خلال المركبات الخطية الجديدة .

الكلمات المفتاحية: الشبكات العصبية، التقدير، متعدد الابعاد، المكونات الرئيسية.

Corresponding Author: E-mail: ahmed.albasrai@uobasrah.edu.iq

Introduction

Artificial intelligence algorithms are one of the most important outputs of the current technological development through which we are trying to improve the work of many mathematical and statistical methods. Therefore, we note that in recent years, there has been increasing interest in creating a hybrid environment between artificial intelligence methods and statistical methods that are concerned with processing data, whether large or small. Therefore, the aim of this paper was to improve the results of analyzing traditional main compounds by adopting the algorithm of artificial neural networks, which is one of the algorithms of artificial intelligence through which data can be trained by adopting an uncontrolled machine method to perform classifications, estimation, and prediction. Therefore, this paper presented an algorithm to improve the work of traditional main compounds by classifying variables by adopting artificial neural networks. Therefore, there were contributions made in this field. We are trying to show the importance and role of this paper in data processing, where he presented:

Davoudabadi et al(2020), included the evaluation of a series of suppliers by adopting a number of criteria, the most important of which is cost and time of supply, where two methodologies were adopted in the analysis process, namely PCA-DEA, with the aim of evaluating the performance of suppliers (Davoudabadi, Mousavi, & Sharifi, 2020). Ibrahim et al. (2020), a paper that addressed the problem of determining the quality of Euphrates water through the use of the neural network algorithm, where the model of multi-stage neural networks was adopted. The paper concluded that the neural network algorithm is an effective method in determining water quality (Ibrahim, Mohammed-Ridha, Hussein, & Faisal, 2020). Karami et al,2020, developed decision-making criteria as the primary factor influencing supplier performance, company reputation, and product quality. An integrated approach was proposed that relies on reducing criteria using principal component analysis, and then enhancing the efficiency measurement of supplier models using the DEA model (Karami, Yaghin, & Mousazadegan, Supplier selection and evaluation in the garment supply chain: an integrated DEA–PCA–VIKOR approach, 2020). Layeb et al, 2020 This paper presents a realistic study on how to evaluate the efficiency of distribution centers based on the performance of logistics service providers, the methodology of Principal component analysis was adopted to determine the appropriate factors, and then the data envelopment analysis method was adopted to measure the efficiency of storage and transportation activities for each center (Layeb, Omrane, Siala, & Chaabani, Toward a PCA-DEA based Decision Support System: A case study of a third-party logistics provider from Tunisia, 2020). Hasan et al. (2021) They designed a system capable of diagnosing human brain tumors, whether benign or malignant, using a curve transformation algorithm and artificial neural networks. The proposed methodology was able to improve the diagnostic results of cancerous tumors through magnetic resonance images (Hasan, Yousif, & Al-Talib, 2021).Wu et al,(2021) Proposed sought to use several multivariate methods with the aim of conducting a comprehensive and systematic evaluation of the overall performance of 125 for the period (1997-2017) by highlighting spatial differences , as well as changes and trends in energy security performance. The methods were the analysis of the main vehicles PCA and the analysis of the data envelope in the presence of the warranty area index and then the adoption of the cluster method by adopting the ESP classification results (Wu, Chung, & Huang, Evaluating Global Energy Security Performances Using an Integrated PCA/DEA-AR Technique, 2021). Ashour (2022) presented an approach that improves the results of time series prediction processes by analyzing the behavior of linear and non-linear data and determining the appropriate model for neural networks. Through the application, the paper reached the identification of the ideal neural network to solve time series problems, which are reverse propagation networks and repetitive neural networks, whether linear, semi-linear or non-linear. The radial exponent function method failed to address non-linear time series while demonstrating good efficiency in the case of linear or quasi-linear time series only (Ashour, 2022). The study of Gebisa, et al., (2022) included finding a developmental methodology in processing the higher dimensions of data by adopting the methodology of neural networks and key compounds in data analysis. This proposal contributed to enhancing the work of the model and preserving the characteristics of the basic data while improving the efficiency and accuracy of the results (Gebisa, Gebresenbet, Gopal , & Nallamothu, 2022).A study by Ali and Daher, (2024), presented a paper aimed at determining the possibility of artificial neural networks to predict prices and their behavior according to financial market indicators. The paper was able to conclude that artificial neural networks are highly predictive in analyzing price behavior according to financial market indicators (Ali & Daher, 2024).

Through the above, the contribution of this paper is to reduce errors and improve the prediction of artificial neural network algorithms by reducing the dimensions of financial data.

1st: Method and Material

1. Artificial Neural Network

Artificial Neural networks are unsupervised machine learning algorithms whose models have been used in recent years to process large and small data for classification, estimation and prediction purposes. The development of Ann models depends heavily on the quality and accuracy of the data. The importance of artificial neural networks comes from their remarkable characteristics in processing information that is mainly related to non-linearity **(Basheer & Hajmeer, 2020)**, fault tolerance and noise. The methodology of this algorithm depends on building it on three main layers: the input layer (Inputs Layer), the hidden layer (s) and the output layer (Outputs Layer) **(Ahmed & Mohammed, 2020)**.

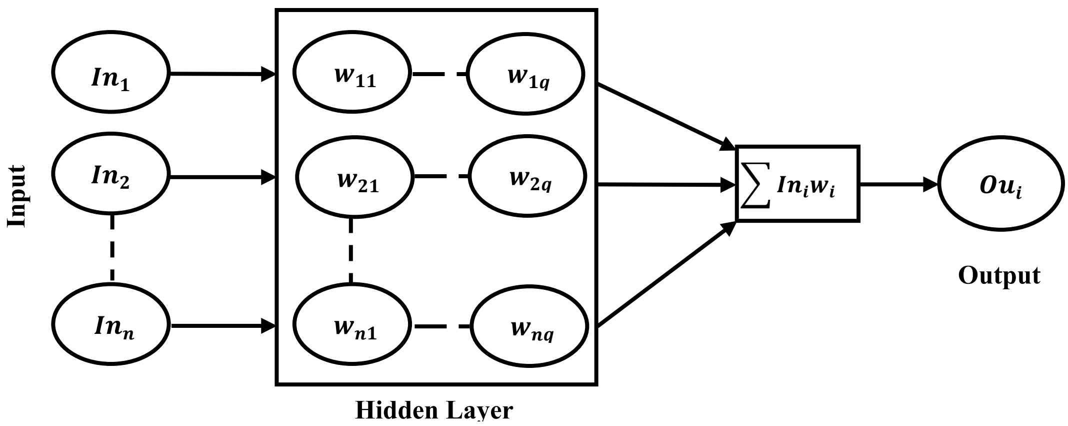


Figure (1): Artificial Neural Networks Algorithm

2. Neural Networks Environment

Neural network models operate in a variety of environments and are the basis for determining the quality of the neural network model. Therefore, the environments of neural networks have been determined based on the nature of the network feeder Feed Forword Network. Feed Backward Network Each of these networks transfers signals from the input units to the output units. The front feeder networks allow the transmission of signals in one direction, while the back feeder networks allow two-way transmission through the presence of loops with the possibility of modifying the paths until they reach a state of equilibrium.

3. Artificial neural networks

Artificial neural networks operate under a specific architecture. If we assume that we have an input $\boldsymbol{X} = [\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \dots, \boldsymbol{X}_{iq}]$ and output vector, the following steps $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n]$ can be built:

1. format data
2. Select Training Sample and Test Sample

3. Determining the Threshold Threshold
4. Select Activation Function Hyperbolic tangent

$$\boldsymbol{f(x) = \frac{\sinh(x)}{\cosh(x)}} \qquad (1)$$

5. Determining the number of neurons
6. Determine weights calculation model for hidden layers

$$\boldsymbol{Z_{j(t+1)} = Z_{ij(t)} + \eta(y^* - y)x_{ij} \;; i = 1, 2, \dots, n \;; j = 1, 2, \dots, q} \qquad (2)$$

Whereas:
$Z_{j(t)}$: Representing weights
y: Represents real output values
y^*Estimated output:
7. Calculate the value y_{ik} of the output from the entrance $x_{ij(k)}$ with a hidden layer H_k

$$y_{ik} = f(H_k);\ k = 1,2, \dots, h \qquad (3)$$

Where:
h: that represented numbers of hidden layers

$$H_k = \sum_{i=1}^{n}\sum_{j=1}^{q} Z_{ij(k)}x_{ij(k)} - u \quad \forall\, k = 1,2, \dots, h \qquad (4)$$

8. Calculate the expected output value $\hat{y}_i$ of a hidden layer H_k with the output of the hidden layer y_{ik}

$$\hat{y}_i = \sum_{i=1}^{n}\sum_{j=1}^{q} Z_{ij(k)}y_{ik} - u \qquad (5)$$

Where **(u)** represents the threshold through which the comparison is made between the network outputs and the original data outputs
9. Random error check

$$e_i = \hat{y}_i - y_i \qquad (6)$$

where e_i is the random error

$$if \begin{cases} e_i = 0\,; & stop \\ e/w\,; edite\ weighted & Z_{ij(k)} \end{cases}$$

If the condition is not met, the process is repeated from step (6) to step (9) until the stop condition is met (Mohammed & Qasim, 2013).

4. Principal Component Analysis Modelling

Suppose we have P of random variables $\underline{A}' = [A_1, A_2, \dots, A_p]$that have a multivariate normal distribution (MVN) (Johnson & Wichern, 2002, p. 149) as building a PCA model requires estimating the covariance matrix Σ, and calculating the roots and eigenvectors (Eigen Values and Eigen Vector), and accordingly, the information matrix that has the degree (n×p) can be written as follows:

$$\boldsymbol{A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}} \qquad (7)$$

The covariance matrix Σ can be obtained by applying the covariance and common covariance formulas.

$$\boldsymbol{S_i = \frac{1}{n-1}(a_{ij} - \bar{a}_i)^T(a_{ij} - \bar{a}_i)} \qquad (8)$$

$$\boldsymbol{S_{ik} = \frac{1}{n-1}(a_{ij} - \bar{a}_i)^T(a_{kj} - \bar{a}_k) \quad ; i \neq k} \qquad (9)$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1k} \\ S_{21} & S_{22} & \dots & S_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ S_{k1} & S_{k2} & \dots & S_{kk} \end{bmatrix} \qquad (10)$$

After determining the covariance matrix according to formula (10), the eigenvalues and eigenvectors will be calculated as below:

$$\left|\boldsymbol{\Sigma} - \boldsymbol{\lambda_j}\mathbf{I}\right| = \mathbf{0} \qquad (11)$$

$$\left|\boldsymbol{\Sigma} - \boldsymbol{\lambda_j}\mathbf{I}\right|\boldsymbol{v} = \mathbf{0} \qquad (12)$$

Through formulas (10, 11 and 12), the main compounds can be executed:

$$\mathbf{C}_{PC_j} = e_{1j}A_1 + e_{2j}A_2 + \dots + e_{ij}A_j \qquad (13)$$

Whereas:

i: represents the observations $i = 1,2, \dots, n$

j: represents the number of variables $j = 1,2, \dots, p$

$\mathbf{C}_{PC_j}$**:** represents the main compounds. In other words, they are the new realized variables, so that each new variable is a linear combination of the original variables.

And the constraints associated with the principal compounds achieved according to formula (13) are:

$$\boldsymbol{Var}\left(\mathbf{C}_{PC_j}\right) = \boldsymbol{max} \qquad (14)$$

$$\boldsymbol{e^T e} = \mathbf{1} \qquad (15)$$

Where formula (14) shows that the best component is the one that represents the highest variance, while formula (15) shows that the sum of the squares of the eigenvector values must be equal to one (Põldaru & Roots, 2014, p. 67), or by adopting the Kaiser Guttman (KG) test, proposed by Guttman in 1954. It disclosed that the identification of the principal important components is reliant on the eigenvalues verified from the matrix of correlation coefficients according to the following condition (Mohammed & Ashour, 2022)**:**

$$\mathbf{PC}_j = \begin{cases} \boldsymbol{\lambda_j} > \mathbf{1}\,; \boldsymbol{Retaind\ the\ PC} \\ \boldsymbol{Trivial} \qquad ; \boldsymbol{e/w} \end{cases} \qquad (16)$$

2nd: Practicality

In this paragraph, the theoretical methodology was applied through the data of the financial sector of the United Arab Emirates through the adoption of some environmental, social and governance factors as dimensions affecting the financial performance of banks, in addition to the trend towards analyzing financial sustainability in light of factors of various trends. Accordingly, the data of the banking sector were adopted for a sample of banks in the United Arab Emirates, which amounted to (16) banks for four years (2020-2023), where the achieved sample was (64) views. The data proven in the doctoral thesis were used **(Abdul Hamid, 2025)**, so the paper's orientation was how to address and reduce dimensions while maintaining the quality of data and increasing the accuracy of estimates by adopting the non-monitored education algorithm. Accordingly, the results were discussed:

1. Discussion of PCA findings

A. Testing the suitability of the analysis data

This test requires estimating the correlation coefficient matrix to determine the KMO and Bartlett Sphericity values

		X_{11}	X_{12}	X_{13}	X_{21}	X_{22}	X_{31}	X_{32}
Correlation	X_{11}	1.000	-0.075	-0.091	0.211	0.172	-0.112	-0.028
	X_{12}	-0.075	1.000	0.514	0.076	0.126	-0.188	-0.042
	X_{13}	-0.091	0.514	1.000	0.170	-0.166	-0.118	0.056
	X_{21}	0.211	0.076	0.170	1.000	-0.071	0.201	0.102
	X_{22}	0.172	0.126	-0.166	-0.071	1.000	0.140	0.246
	X_{31}	-0.112	-0.188	-0.118	0.201	0.140	1.000	0.369
	X_{32}	-0.028	-0.042	0.056	0.102	0.246	0.369	1.000

Figure (2): Matrix of correlation coefficients between independent variables

Table (1): KMO and Bartlett's test results to determine the suitability of the data

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.430
Bartlett's Test of Sphericity	Approx. Chi-Square	.060
	Df	21
	Sig.	0.000

Source: Prepared by the researcher according to the results of the SPSS V.26 program

We note from the results of the table (1) that the data was poorly appropriate, as the KMO test = 0.430 was achieved, and this is due to the quality of the data, but the value of the Bartlett test Shpericity was $\chi^2 = \mathbf{56.060}$ at the level of significance Sig=0.000, and this is evidence of the significance of the analysis and its compatibility with the hypothesis, which states that the matrix of correlation coefficients is no single.

B. Identification of components PC_j

After testing the significance of the data and verifying their suitability for analysis, we tend to identify the principal compounds that serve as a linear composition of all the target variables, relying on the subjective values resulting from the application of the formula (11) and reaching their variations, how much in the table (2):

Table (2): showing the intrinsic values and their variations for the target variables

Comp.	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Var.	Cum.%	Total	% of Var.	Cum. %	Total	% of Var.	Cum.%
1	1.594	23.906	23.906	1.594	23.906	23.906	1.514	22.705	22.705
2	1.426	21.390	45.296	1.426	21.390	45.296	1.424	21.353	44.058
3	1.137	17.055	62.351	1.137	17.055	62.351	1.153	17.291	61.349
4	1.074	16.105	78.456	1.074	16.105	78.456	1.140	17.107	78.456
5	0.630	9.445	87.902						
6	0.460	6.903	94.805						
7	0.346	5.195	100.000						

Source: Prepared by the researcher according to the results of SPSS V.26

It is clear from the table (2) that the number of basic compounds that meet the conditions according to the formula (14, 16) is (4) compounds (λ_1 λ_2 λ_3, ,λ_4), which had the advantage according to the achieved variations (23.906, 21.390, 17.055, 16.105) respectively, where we note that these compounds achieved an interpreted variation (78.456%). This is evidence that the four compounds were able to represent the data well and form new variables as linear compounds in terms of a number of variables.

In confirmation of this, the method of drawing Scree Plot was proven to meet the conditions, where we note that the number of components that passed one according to their intrinsic values is (4) basic compounds as in the figure (٣)**:**

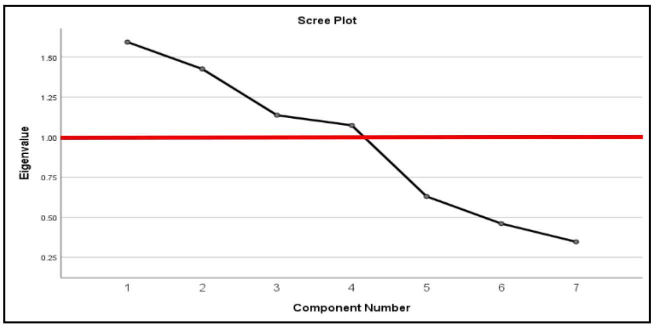


Figure (3): show the number of PC for logistics indicators.

The source is prepared by the researcher according to the results of the analysis of SPSS V.26

Table (3): Transactions of independent variables according to their importance in the corrected basic compounds

Rotated Component Matrix[a]				
Component Variables	**1**	**2**	**3**	**4**
Emissions	---	---	0.799	---
Energy Consumption	0.863	---	---	---
Water Consumption	0.855	---	---	---
Percentage of Female Employees	---	---	0.753	
Emiratization	---	---	---	0.906
Board Diversity	---	0.833	---	---
Independence	---	0.755	---	---
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.				
a. Rotation converged in 8 iterations.				

Source: Prepared by the researcher according to the results of SPSS V.26

It is clear from the table (3) that all the coefficients of the variables are and $\mathbf{a}_{ij} \geq \mathbf{0.70}$ therefore new linear variables were formed that summarized the important variables and took a specific classification as in the table (٤):

Table (4): Classification of P.C. according to the coefficients of the variables

Variable	**Acidity**	**Linear Component**
Fact.1	Environmental Component	C_{PC_1}
Fact.2	Governance Component	C_{PC_2}
Fact.3	Mix Component	C_{PC_3}
Fact.4	Localization Component	C_{PC_4}

Source: Prepared by the researcher according to the results of SPSS V.26

It is clear from the table (4) that the component $\mathbf{C}_{PC_4}$ is the weakest component, as it included a single variable, which is localization, which is considered one of the social dimensions, while the linear component $\mathbf{C}_{PC_1}$ represented the social dimension by adopting the two variables (energy consumption and water consumption), and then in the second place came the linear component for the governance$\mathbf{C}_{PC_2}$ dimension, which included two variables (diversity in the council and independence)

3rd: Results of Artificial Neural Networks

In this paragraph, the stages of building the artificial neural networks algorithm were described with the identification of the training and testing sample, which is one of the basics of the work of unattended education algorithms. The data was configured with a size of (64) views and then the training sample was determined, which was formed by (84%) of the original data ($n_{tr} = 54$), which is the sample responsible for building the neural networks model, while the test sample was formed by (28%) of the original data ($n_{te} = 10$), which is the sample responsible for testing the estimated model. Furthermore, the data were Rescaling Method using the Standardized.

1. Discussion of Traditional Artificial Neural Networks

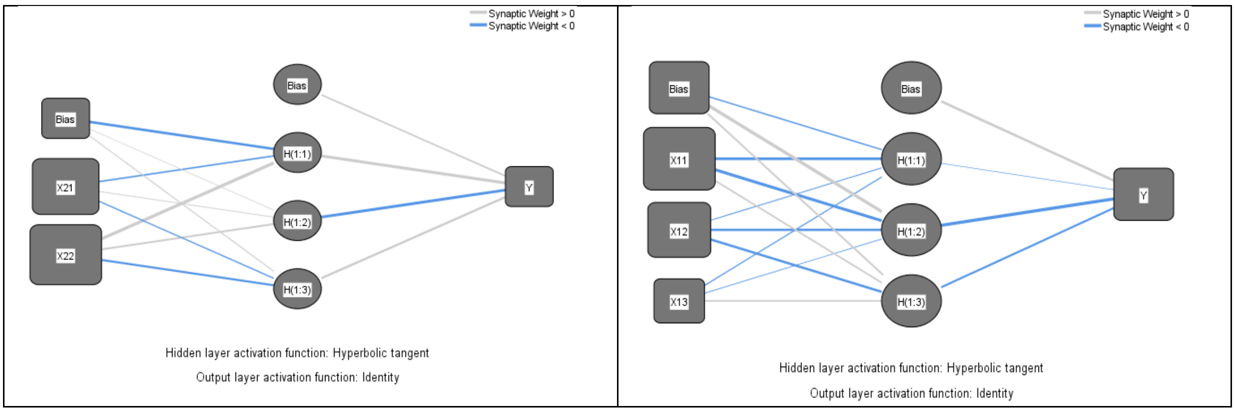
In this paragraph, the results of the modeling of traditional artificial neural networks of independent variables are confirmed and shown in the table (5):

Table (5): Traditional Artificial Neural Networks Results

	Model	ANN_Eni.	ANN_So.	ANN_Go	ANN_For all
	Number of Units in Hidden Layer	3	3	3	4
	Train. MSE	0.443	0.374	0.471	0.465
	Test MSE	0.259	0.174	0.272	0.213
Importance Var.	X_{11}	0.574	---	---	.136،
	X_{12}	0.362	---	---	.144،
	X_{13}	0.065	---	---	.048،
	X_{21}	---	0.473	---	.053،
	X_{22}	---	0.527	---	.406،
	X_{31}	---	---	0.313	.052،
	X_{32}	---	---	0.687	.162،
	Best Model	4	1	3	2

Source prepared by the researcher on and only the results of SPSS V.26

It is clear from the results of the table (5) that traditional neural networks have shown that the neural networks model of the social dimension is the best in determining financial performance, as it achieved the sum of the error squares MSE=0.174, which is the lowest compared to other models, while the neural networks model came in second place for all factors, which achieved MSE = 0.213. In addition, the models were able to identify the most important variables, as we note that the artificial neural networks model of the social dimension has shown the importance of the localization variable $(\boldsymbol{X_{22}})$**.** In addition, the localization variable has the highest priority in the neural networks model at all factors. In the end result, we note that the model has excluded most of the variables at each model**.**



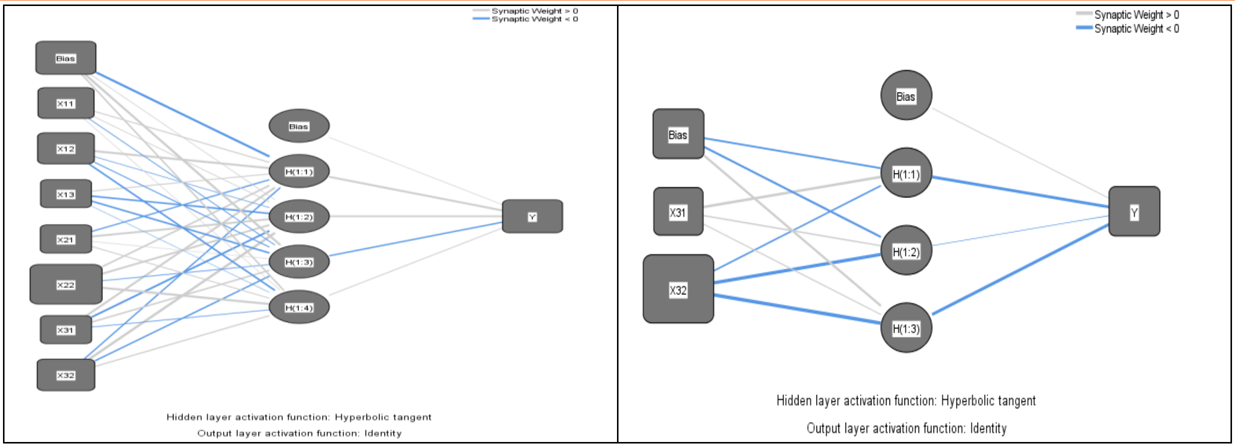


Figure (4): Traditional Artificial Neural Networks of Native Variants
Source: Prepared by the researcher according to the results of SPSS V.26

2. Discussion of Enhanced Artificial Neural Networks Results

In this paragraph, artificial neural networks were built based on the results of the main components, relying on the latter's ability to reduce factors and build new linear components as shown in Table (4), taking into account the training and testing of networks on the same pre-defined criteria, and accordingly the results were reached as in Table (6):

Table (6): Results of artificial neural networks based on the P.C.

Model		ANN_Eni. Pc1	ANN_Go. Pc2	ANN_Mix Pc3	ANN_Loca. Pc4	ANN_For all Pc
Number of Units in Hidden Layer		5	4	4	4	3
Train. MSE		0.485	0.458	0.459	0.397	0.326
Test MSE		0.279	0.173	0.252	0.153	0.150
Importance Var.	Fac1	---	---	---	---	0.233
	Fac2	---	---	---	---	0.211
	Fac3	---	---	---	---	0.261
	Fac4	---	---	---	---	0.295
Best Model		5	3	4	2	1

It is clear to us from the results of the table (6) that the neural networks improved using the P.C. have shown an improvement in their models in terms of the degree of error rates. In addition, most of the models have improved the number of hidden nodes. The greater their number, the greater the accuracy of the estimates. The results showed that the neural networks model for all linear basic components has achieved the best total error boxes for the test sample, which reached MSE=0.150. Moreover, the neural networks model for all factors was characterized by a similar amount of importance. Thus, the impact of any factor on banking financial performance was not neglected. The figure shows the improved artificial neural networks by adopting the main components, through which the number of components within the hidden layer is shown.

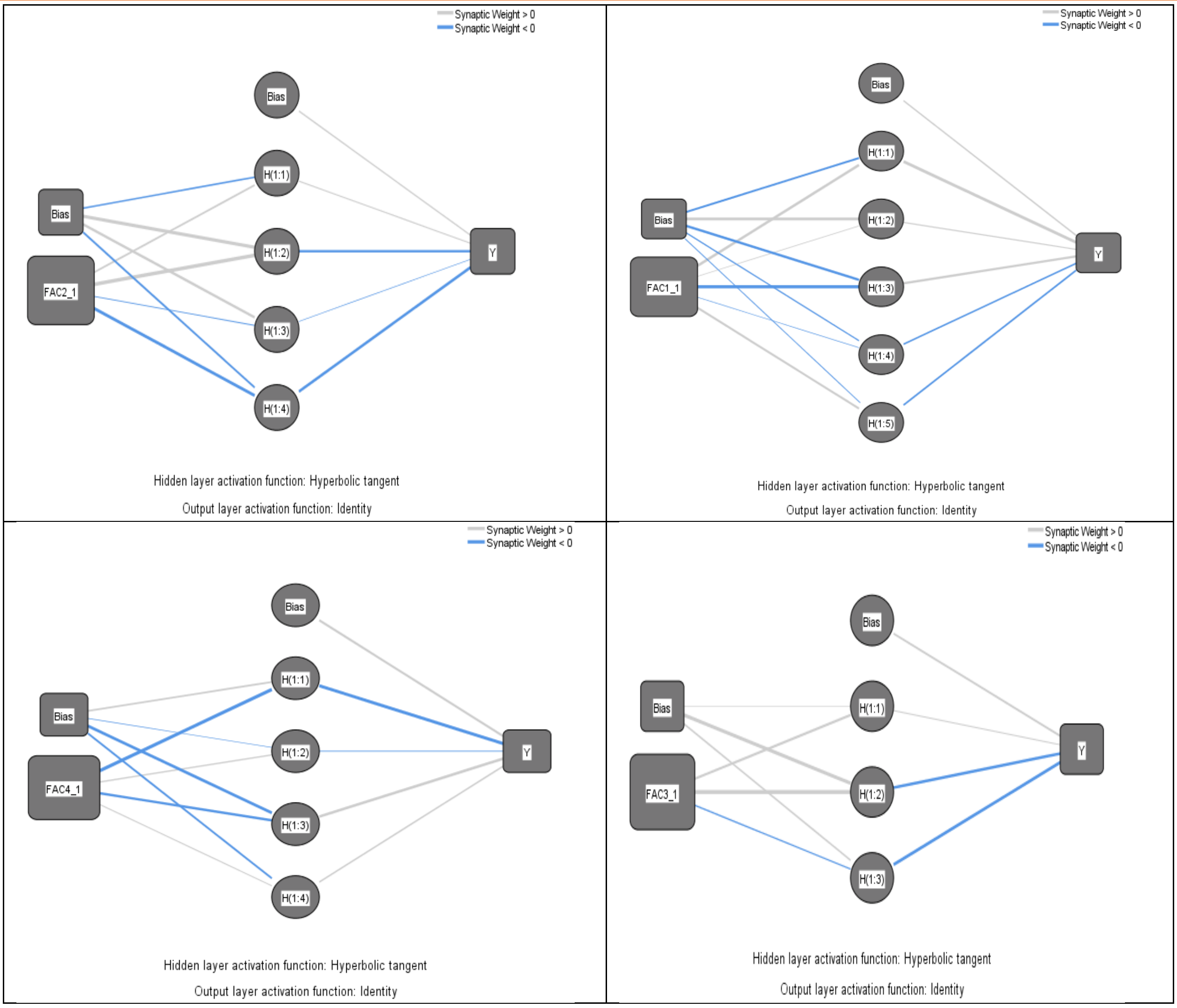


Figure (5): PCA Enhanced Artificial Neural Networks

Source: prepared by the researcher according to SPSS

From the above, we can compare similar models in terms of description to reach a final decision about the importance of improvement in the work of the neural network algorithm

Table (7): Comparison between traditional and enhanced neural network models by adopting the P.C.

Model	MSE
ANN_Eni	0.259
ANN_Eni_Fac1	0.279
Best Model	ANN_Eni
ANN_Go	0.272
ANN_Go_Fac2	0.173
Best Model	ANN_Go_Fac2
ANN_For all Var.	0.213
ANN_For all Fact.	0.150
Best Model	ANN_For all Fact.

Source: Prepared by the researcher on and only the results of the SPSS program v.26

It is clear from the results of the table (7), which represents the final comparison between the models of neural networks, that the preference was for the models of improved neural networks by adopting the principal components, where it achieved the lowest total error squares (0.150, 0.173) respectively at the typical ANN_For all Fact. and ANN_Go_Fac2, while the traditional neural networks model of the environmental dimension (ANN_Eni) was determined on the preference as it achieved the MSE (0.279) compared to the improved neural networks model.

3. Estimated values of the model ANN_For all Fact.

The improved model of artificial neural networks was adopted by adopting the principal components as a model for estimating financial performance and obtaining the estimated values of the model, as in the table (8):

Table (8): Estimated values of the financial performance variable

Y	MLP_Pred y_for all var.	MLP_Pred y_for all Fact.
0.01145	0.00939	0.00514
0.00197	0.00752	0.00503
-0.04561	0.00627	0.00477
-0.01390	0.00686	0.00573
0.01563	0.01487	0.01508
0.00926	0.01090	0.00575
-0.01844	0.00921	0.00553
0.00958	0.01624	0.01076
0.01255	0.01549	0.01804
0.00757	0.00643	0.01177
0.00998	0.01037	0.01621
-0.00760	0.00607	0.00598
-0.01181	0.00472	0.00509
0.01091	0.01926	0.01752
0.01150	0.00622	0.00505
-0.00068	0.00439	0.00500
0.01253	0.01520	0.01311
0.00628	0.00551	0.00520
0.00462	0.00357	0.00474
0.00309	0.01057	0.01025
0.02010	0.01868	0.02422
0.01192	0.01302	0.01476
0.00126	0.00576	0.00598
0.01317	0.00828	0.00801
0.01702	0.02000	0.02331
0.00935	0.00720	0.01317
0.01354	0.01240	0.01841
0.00608	0.00079	0.00728
0.01067	0.00142	0.00362
0.01483	0.01210	0.01854
0.01270	0.00944	0.00663
0.01268	0.01377	0.01770
0.01209	0.01322	0.01032
0.00711	0.01012	0.00649
0.01768	0.00218	0.00530
0.00715	0.00983	0.00995
0.02679	0.01783	0.02416
0.01292	0.01257	0.01204
-0.00424	0.00609	0.00590
0.01753	0.01036	0.00734
0.02148	0.01741	0.02167
0.01101	0.00565	0.01104
0.01753	0.01146	0.01796

0.01938	-0.00063	0.00654
-0.00609	0.00168	0.00378
0.01926	0.01187	0.01842
0.01571	0.01458	0.01443
0.01659	0.01256	0.01365
0.01413	0.01512	0.01339
0.00899	0.00667	0.00499
0.01653	0.00472	0.00488
0.01402	0.00944	0.00911
0.03401	0.01892	0.02419
0.01447	0.01290	0.01429
-0.00698	0.00937	0.00563
0.02412	0.01165	0.00753
0.02723	0.01779	0.02188
0.01293	0.00567	0.01207
0.02494	0.00521	0.01339
0.03615	-0.00087	0.00672
0.02369	0.00315	0.00423
0.02230	0.00957	0.01779
0.02055	0.01482	0.01569
0.02416	0.00967	0.00919

MLP_Pred y_for all var.	**Estimated values according to the traditional artificial neural network algorithm.**
MLP_Pred y_for all Fact.	**Estimated values according to the improved artificial neural network algorithm**

It is clear from the figure (6) that the estimated values according to the model of the improved neural networks by adopting the main components were the most accurate compared to other models

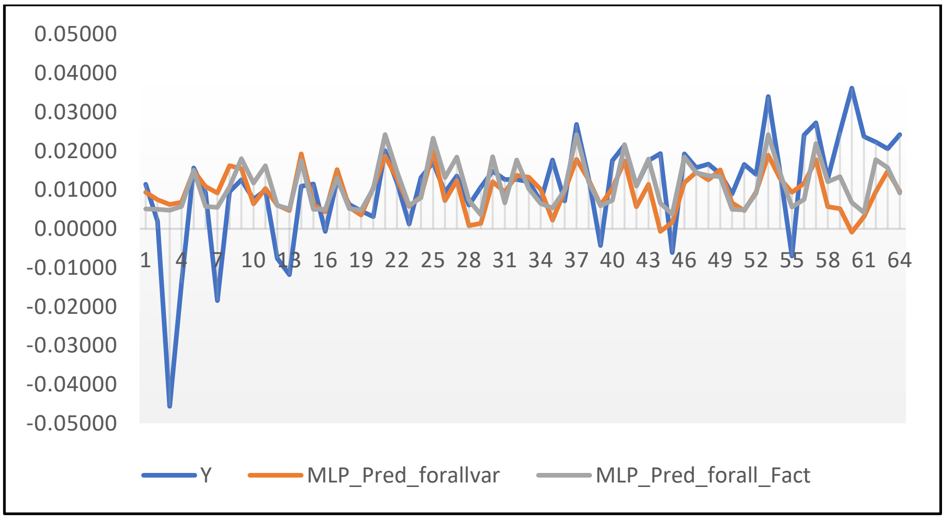


Figure (6): Comparing the estimated values of the two neural network models with the original values

4th: Conclusions and Recommendations

1. Conclusions

Through the results presented, a number of conclusions were reached
a. Adopting the method of principal components is important in processing artificial intelligence algorithms, especially when there is high-dimensional data.
b. The proposed algorithm contributed to determining the appropriate estimates of the financial performance of the banking sector under the algorithm of artificial neural networks, as the results showed that it is more balanced and volatile.
c. The paper was able to reach the conclusion that the estimated model by adopting the optimization algorithm for neural networks by adopting the principal components is more accurate when adopting the new factors, as it achieved the lowest total error squares, which amounted to MSE =1.496 compared to the traditional neural networks model.

2. Recommendations

a. The paper recommends the development of hybrid methods through the adoption of statistical and intelligent methods in improving and increasing the accuracy of the results.
b. Work to diversify the sources of financial data in such studies because financial variables have an unstable tendency during time periods.
c. Processing financial data by adopting developed statistical methods such as fuzzy methods and high-dimensional statistical methods that process non-linear data because data often suffer from this problem.

References

1- Abdul Hamid, M. T. (2025). The role of sustainable finance in enhancing the financial performance of the banking sector in the United Arab Emirates, with special reference to Iraq. Basrah: University of Basra, College of Administration and Economics, Department of Financial and Banking Sciences, Unpublished PhD Thesis.
2- Abonyi, J., Babuska, R., & Szeifert, F. (2002). Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 32(5), 612 - 621. doi:https://doi.org/10.1109/TSMCB.2002.1033180
3- Ahmed, M. S., & Mohammed, F. A. (2020). The Use of Genetic Algorithm to Train the Artificial Neural Network for the Purpose of Predicting Baghdad Bank Daily Closing Prices on the Iraqi Stock Exchange. Tikrit Journal of Administrative and Economic Sciences, 16(51), 482-498. Retrieved from https://www.iasj.net/iasj/article/187541
4- Al Aqad, M. H., & Cardoso, L. M. (2023). Ant Colony Optimization (ACO) Based Fuzzy C-Means (FCM) Clustering Approach for MRI Images Segmentation. Wasit Journal of Computer and Mathematics Science, 2(4), 115-125. doi:https://doi.org/10.31185/wjcms.230
5- Ali, A. H., & Daher, S. S. (2024). Testing the ability of artificial neural networks to predict the price behaviour of financial market indices " An empirical study using historical data and technical analysis indicators". Journal of Management Research, 42(2), 1-54. doi:https://doi.org/10.21608/jso.2024.273421.1264
6- Ali, N. J., & Yaba, S. P. (2023). Automated Thalassemia cell image segmentation using hybrid Fuzzy C-Means and K-Means. Zanco Journal of Pure and Applied Sciences, 35(4), 22-33. doi:http://dx.doi.org/10.21271/zjpas
7- Al-Janabee, O., & Al-Sarray, B. (2022). Evaluation Algorithms Based on Fuzzy C-means for the DataClustering of Cancer Gene Expression. raqi Journal for Computer Science and Mathematics, 3(2), 27-41. doi:https://doi.org/10.52866/ijcsm.2022.02.01.004
8- Anderson, T. (2003). An Introduction to Multivariate Statistical Analysis (3 ed.). (Anderson, Ed.) John Wiley & Sons, John Wiley & Suns, Inc. Huboken, New Jersey.
9- Ashour, M. A. (2022, 8 1). Optimized Artificial Neural network models to time serie. Baghdad Science Journal, 19(4), 899-904. doi:https://doi.org/10.21123/bsj.2022.19.4.0899
10- Azadeh, A., Nasirian, B., Salehi, V., & H. Kouzehchi, H. (2016). Integration of PCA and DEA for identifying and improving the impact of Six Sigma implementation on job characteristics in an automotive industry. Quality Engineering, Taylor & Francis, 29(2), 273-290.
11- Azadeh, A., Nasirian, B., Salehi, V., & H. Kouzehchi, H. (2016). Integration of PCA and DEA for identifying and improving the impact of Six Sigma implementation on job characteristics in an automotive industry. Quality Engineering, Taylor & Francis, 29(2), 273-290.
12- Basheer, L. A., & Hajmeer, M. (2020). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43(1), 3-31. doi:https://doi.org/10.1016/S0167-7012(00)00201-3

13- Davoudabadi, R., Mousavi, S., & Sharifi, E. (2020, 2). A new integrated weighting and ranking model based on entropy, DEA and PCA considering two aggregation approaches for resilient supplier selection problem. Journal of Computational Science, 40, 1-32.
14- Gath, I., & Geva, A. B. (1989). Unsupervised Optimal Fuzzy Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 773 - 780. doi:https://doi.org/10.1109/34.192473
15- Gebisa, A., Gebresenbet, G., Gopal , R., & Nallamothu, R. B. (2022). A Neural Network and Principal Component Analysis Approach to Develop a Real-Time Driving Cycle in an Urban Environment: The Case of Addis Ababa, Ethiopia. Sustainability, MDPI, 14(21), 1-37. doi:https://doi.org/10.3390/su142113772
16- Hasan, S., Yousif, M., & Al-Talib, T. M. (2021). MRI Probabilistic Neural Network Screening System: a benign and malignant recognition case study. Iraqi Journal of Science(Special Issue), 161-166. doi:10.24996/ijs.2021.SI.1.22
17- Hochreiter, S. (2013). Basic Methods of Data Analysis. Institute of Bioinformatics, Johannes Kepler University Linz, statistics. Austria: Johannes Kepler University Linz.
18- Ibrahim, M. A., Mohammed-Ridha, M. J., Hussein, H. A., & Faisal, A. A. (2020). Artificial Neural Network modeling of the water quality index for the Euphrates river in Iraq. Iraqi Journal of Agricultural Sciences, 51(6), 1572-1580. doi:https://doi.org/10.36103/ijas.v51i6.1184
19- Johnson, R. A., & Wichern, D. W. (2002). Applied Multivariate Statistical Analysis (5 ed.). Prentice Hall, upper saddle river, New Jersey.
20- Karami, S., Yaghin, G. R., & Mousazadegan, F. (2020, 5). Supplier selection and evaluation in the garment supply chain: an integrated DEA–PCA–VIKOR approach. The Journal of The Textile Institute, 112(4), 578-595.
21- Karami, S., Yaghin, G. R., & Mousazadegan, F. (2020, 5). Supplier selection and evaluation in the garment supply chain: an integrated DEA–PCA–VIKOR approach. The Journal of The Textile Institute, 112(4), 578-595.
22- Layeb, S. B., Omrane, N. A., Siala, J. C., & Chaabani, D. (2020). Toward a PCA-DEA based Decision Support System: A case study of a third-party logistics provider from Tunisia. 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET) (pp. 294-299). Hammamet, Tunisia: IEEE.
23- Layeb, S. B., Omrane, N. A., Siala, J. C., & Chaabani, D. (2020). Toward a PCA-DEA based Decision Support System: A case study of a third-party logistics provider from Tunisia. 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET) (pp. 294-299). Hammamet, Tunisia: IEEE.
24- Mohammed, A. H., & Ashour, M. A. (2022). Improving the efficiency measurement index using principal component analysis (PCA). Journal of Algebraic Statistics, 13(2), 2818 - 2831. Retrieved from https://www.publishoa.com/index.php/journal/article/view/518/451
25- Mohammed, I. R., & Qasim, O. S. (2013). A mathematical and analytical study of artificial neural network algorithms in model fitting for medical diagnosis. AL-Rafidain Journal of Computer Sciences and Mathematics, 10(1), 183-194. Retrieved from https://www.iasj.net/iasj/article/71737
26- Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2005, June). How many principal components? stopping rules for determining the number of non-trivial axes revisited. Computational Statistics & Data Analysis, 49(4), 974-997.
27- Põldaru, R., & Roots, ü. (2014, 3). A PCA-DEA approach to measure the quality of life in Estonian counties. Socio-Economic Planning Sciences, 48(1), 65-73.
28- Wang, C. (2024). Statistical method for clustering high-dimensional data based on fuzzy mathematical modeling. Applied Mathematics and Nonlinear Sciences, 9(1), 1-16. doi:https://doi.org/10.2478/amns.2023.2.01452
29- Wu, T.-H., Chung, Y.-F., & Huang, S.-W. (2021, 6). Evaluating Global Energy Security Performances Using an Integrated PCA/DEA-AR Technique. Sustainable Energy Technologies and Assessments, 45, 1-17.
30- Wu, T.-H., Chung, Y.-F., & Huang, S.-W. (2021, 6). Evaluating Global Energy Security Performances Using an Integrated PCA/DEA-AR Technique. Sustainable Energy Technologies and Assessments, 45, 1-17.
31- Zhang, C., & Fang, T. (2023). Research on Fault Diagnosis Method of Rolling Bearing Based on Variational Mode Decomposition and Gath-Geva Clustering. 2022 4th International Conference on Mechanical Engineering and Automation (MEA 2022) 09/12/2022 - 11/12/2022 Chongqing, China. 2528, pp. 1-6. China: Journal of Physics: Conference Series. doi:doi:10.1088/1742-6596/2528/1/012023