ISSN 2303-4521
**Periodicals of Engineering and Natural Sciences** *Original Research*
Vol. 10, No. 3, June 2022, pp.187-196

# Comparison of some artificial neural networks for graduate students

**Rabab Abdulrida Saleh[1], Manal Jabbar salman[2*]**
[1] Department of Statistics, University of Baghdad, Iraq
[2] Department of Postgraduate Affairs, University of Baghdad, Iraq

## ABSTRACT

Artificial Neural Networks (ANN) is one of the important statistical methods that are widely used in a range of applications in various fields, which simulates the work of the human brain in terms of receiving a signal, processing data in a human cell and sending to the next cell. It is a system consisting of a number of modules (layers) linked together (input, hidden, output). A comparison was made between three types of neural networks (Feed Forward Neural Network (FFNN), Back propagation network (BPL), Recurrent Neural Network (RNN). he study found that the lowest false prediction rate was for the recurrentt network architecture and using the Data on graduate students at the College of Administration and Economics, University of Baghdad for the period from 2014-2015 to The academic year 2017-2018. The variables are use in the research is (student's success, age, gender, job, type of study (higher diploma, master's, doctorate), specialization (statistics, economics, accounting, industry management, administrative management, and public administration) and channel acceptance). It became clear that the best variables that affect the success of graduate students are the type of study, age and job.

**Keywords**: Artificial Neural Network, Feed Forward Neural network, Back Propagation Learning Network, Recurrent Neural Networks, layers

***Corresponding Author:***

Manal Jabbar salman
Department of Postgraduate Affairs
University of Baghdad, Iraq
Baghdad, Iraq
E-mail: manal.Jabbar@uobaghdad.edu.iq

## 1. Introduction (11 pt, Sentence case)

Artificial neural networks are models that simulate biological neural networks, and they use a number of methods used in biological neural systems with the help of simulation programs in addition to the parallel processing method, that is, artificial neural networks process information in a manner similar to the way the human brain works. A neuron is a mathematical model used to model statistical data, as it is an interconnected group of neurons. Therefore, scientists sought to know the mechanism of the human brain in terms of its ability to process data and the ability to memorize, memory, distinguish between things and make decisions by studying the neural network that consists of billions of cells connected to each other by a nerve. Synapse, which is responsible for transmitting instructions between cells and hence the idea of building a network. A synthetic consists of a number of nodes (cells) with related relationships between them and on the basis of mathematical models for processing the data under study [1]. In 2006 the researcher Brockett & other presented a study on the effect of the chosen statistical/mathematical model and the studied set of variables on the ability to identify financially distressed life insurance companies. The models considered are two artificial neural network methods (backpropagation and vector learning quantization (LVQ)) and two more standard statistical methods (multiple differentiation analysis and logistic regression analysis). and a data set of twenty-two variables with the result that BP and LVQ outperformed conventional statistical methods for all groups across two different assessment criteria (total misclassification cost and re-substitution risk criteria [2]. In 2020, the researcher (Alamia and others) made a comparison between feedforward versus repetitive neural networks and human performance in learning synthetic grammar, in order to deepen our understanding of human cognition, and

explained that the structure of the neural network (feedback versus repetitive) matches human behavior in learning synthetic grammar, And that recurrent neural networks simulate the learning of technical human rules more closely than the feed-forward structures [3]. In 2022 the researcher (***Ashour)*** presented a study on better and more efficient diagnosis of ANN models (backpropagation, radial function neural networks (RBF), and recurrent neural networks) in solving linear and nonlinear time series behavior. To select the best neural networks, the mean squared error and the mean absolute squared error were used to measure the estimation accuracy of the methods used. The important result obtained in this research is that the optimal neural network was Backpropagation (BP) and Recurrent Neural Networks (RNN) for solving time series, whether linear or non-linear. And the inefficiency of RBF in solving nonlinear time series. The results showed improvements in Modern methods of time series forecasting [4]. In this paper, a comparison is made between some types of artificial neural networks (Feed Forward Neural Network (FFNN), Back propagation network (BPL), Recurrent Neural Network (RNN). The statistical criterion mean squares error (MSE) was used, and it turned out that the recurrent neural network (RNN) is the best among the networks because it gives the least possible error.

## 2. Artificial neural network

Artificial neural networks are considered one of the modern methods of data processing that have been widely applied recently due to their efficiency and distance from bias. This method appeared in 1943 when each of Warren Mecullen and Walter pitts Where they made the first model of artificial neural networks Where they made the first model of artificial neural networks [5,6]. The network consists of a number of layers (nodes) which are called neurons and are three main layers

1. Input layer: that includes the variables included in the study and their number is not restricted.
2. The hidden layer: in this layer contains mathematical operations and comparisons, and their number depends on the type of treatment to be studied.
3. The Output Layer: It is the last layer of the network and includes the study outputs, and it is one or more [1,5,7]. It can be illustrated in Figure (1)

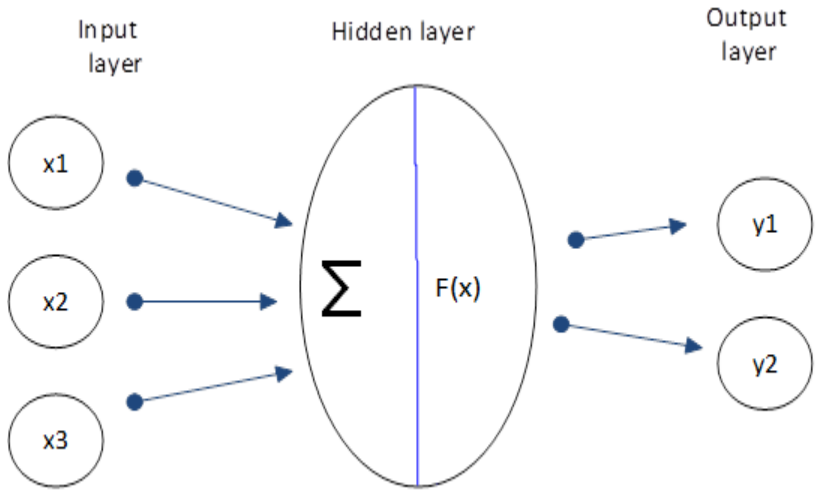


Figure 1. the mathematical form of the artificial neural network

The activation (transfer) function is included in the second (hidden) layer and is on Various types depending on the required treatment, including linear and non-linear and others [8]. The weights are updated (which is the main idea of the network), ie the possibility of reaching the best parameters of the model.

## 3. Methods of teaching an artificial neural network

To get more efficient results, it is trained in two ways:

1. Observed instruction (by teacher) Supervised Learning of ANNs by presenting the inputs and outputs in pairs (ie the presentation of the target input and output) and either by using memory (ie storing information) or by teaching it true and false.

2. Self-learning (without a teacher) unsupervised learning of ANNs, here the network is provided with inputs only, as the network remains on the basis of developing its ability to know shapes, properties and patterns without giving prior examples [9,10].
The main goal of networks is to teach them to update continuous weights to reach the best estimate of the model parameters. In this research, three types of networks have been relied on (forward propagation network, back propagation network, Recurrent network) as shown below.

## 4. Feed Forward Neural Network (FFNN)

It is one of the widely spread networks that was called the front of the algorithm for the process of the algorithm in one direction only (from input to output) so that the weights of the subsequent units are not updated, the number of layers in it is not restricted [11] Let's assume we have the input the variables ($X_1$, $X_2$,$X_3$, $X_4$).

$c_i = f_i(w_i x_i + b_i)$ i=1,2,3,.... .....(1)

Where wi represents the weights, bi: represents the amount of bias

$z_j = f_j \sum_{j=1}^{n} mjcj$ j=5,6 .....(2)

mj = weights input, f represents the update function

$Y = f\ [h_k \sum_{k=1}^{n} z_k]$ k=7 .....(3)

Which can be represented as follows (2)

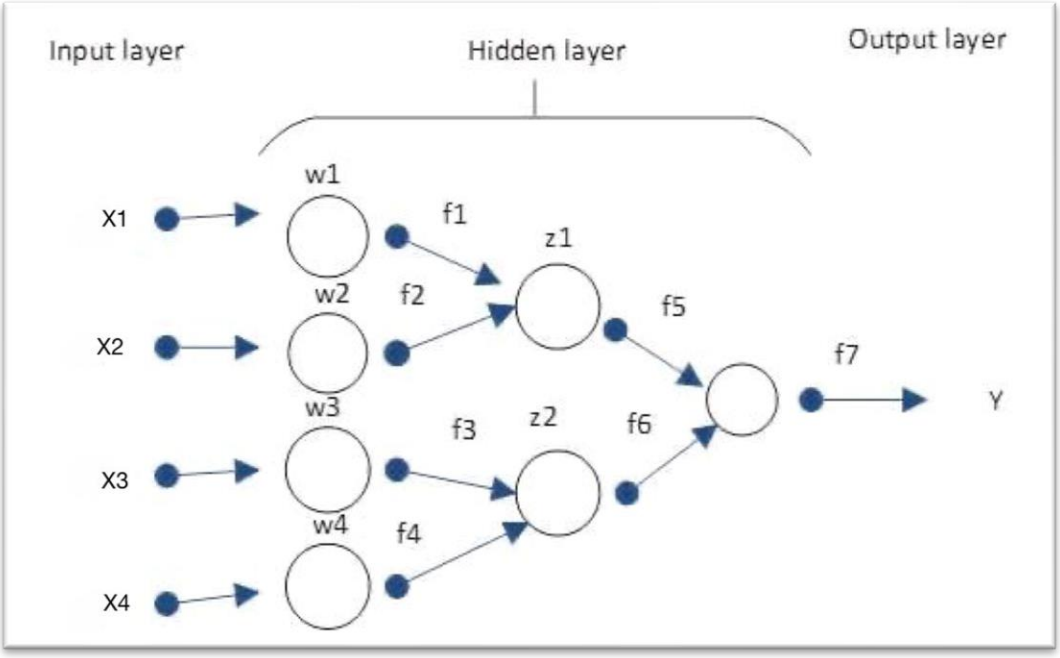


Figure 2. The structure of the forward propagation neural network [5]

## 5. Back Propagation Learning (BPL)

The Back Propagation neural network is one of the most widely used neural networks at present, the BP used in finance, accounting, management science, e-commerce and computer science, as it is also called multilayer. was proposed by Rumelhart, Hinton and Williams [12] BP neural network is a typical multi-layer forward network, the work of the network is to enter information from the input layer, then process it in the hidden layer, and then the output layer. If the final output is not reached, the network weight is modified, it uses forward propagation of information and backward propagation of errors to adjust communication weights between cells to reduce error [13] It can be explained by the following steps:

1- Initialization in which the incoming weights and the network are configured (imposed) with random value

2- Training the network, which is feeding the network with all possible examples of the outputs, taking into account the training periods to maintain mathematical accuracy and after applying all directions to the test, weights are adjusted in them and are saved at each weight during the training period and are estimated only once to fit with the least error The direction of results from inputs to outputs. Backward spread of weights adjustment based on error analysis between calculated output units and actual output units (Error(w))

Error (w) = $\frac{1}{2}\sum_{d\in D}\sum_{k\in K}(y_{kd}-\hat{y}_{kd})^2$ ..... (4)

Where D is the training group and K is the group of output units

$\Delta W_{ij}=-\eta\frac{\partial E}{\partial W_{ij}}$ ..... (5)

3- Re-repeat the steps, and the weights are approved in the different case in the next unit.

$net_j=\sum W_{ij}x_{ij}$ .....(6)

Where xij is the input in the unit j

To illustrate, we use the chain rule to change the weights as follows:

$\frac{\partial Ed}{\partial Wij}=\frac{\partial Ed}{\partial netj}.\frac{\partial netj}{\partial Wij}=\frac{\partial Ed}{\partial netj}x_j$ ..... (7)

$\frac{\partial Ed}{\partial netjj}=\frac{\partial Ed}{\partial \hat{y}j}.\frac{\partial \hat{y}j}{\partial netj}$ ..... (8)

$\because Ed=\frac{1}{2}\sum_{k\in K}(Y_k-\hat{y}_k)^2$ ..... (9)

$\frac{\partial Ed}{\partial Y_j}=\frac{\partial}{\partial \hat{y}_k}.\frac{1}{2}\sum_{k\in K}(Y_k-\hat{y}_k)^2$..... (10)

The derivative of $(\frac{\partial}{\partial Y_j}\sum_{k\in K}(Y_k-\hat{y}_k)^2$

It will be equal to zero for all output units the unit k is equal to k = j

$\frac{\partial Ed}{\partial Y_j}=\frac{\partial}{\partial \hat{y}_j}(y_j-\hat{y}_j)^2=\frac{1}{2}\sum_{k\in K}(Y_k-\hat{y}_k)^2\frac{\partial(y_j-\hat{y}_j)}{\partial \hat{y}_j}$ ..... (11)

$\frac{\partial \hat{y}_j}{\partial netj}=\hat{y}_j(1-\hat{y}_j)$ ..... (12)

$\frac{\partial Ed}{\partial netj}=-(y_j-\hat{y}_j)\hat{y}(1-\hat{y}_j)$ ..... (13)

$\Delta w_{ij}=-\acute{\eta}\frac{\partial Ed}{\partial w_{ij}}=\acute{\eta}(y_j-\hat{y}_j)\hat{y}(1-\hat{y}_j)x_{ij}$ ..... (14)

4- To adjust the weights we use [8,14].

$w_{ijnew}=w_{ijold}-\Delta w_{ij}$.....(15)

## 6. Recurrent Neural Network (RNN)

Recurrent neural networks (RNNs) is used as an extension to feedforward networks, that are naturally suited time series data prediction, processing control, and so on , in which the outputs from neurons are used as feedback to the neurons of the previous layer. In other words, the current output is considered as an input for . It similar to the forward propagation networks, but they differ from them in that they .the next output [15] work on modifying the outputs by returning to the input weights and this is similar to the work of back Based on the analysis of the error in the model and using the internal memory here .propagation networks [8]. to deal with the sequence of inputs as measured by Equation No. (9) by taking partial derivation to update the When the least achievable error is reached, weights are adopted as parameters (5) .weights as in Equation No to the next units, and so on [11,16]. The number of hidden nodes is not restricted here and depends on the quality of the calculated processing. This method is considered one of the strongest methods because it works to update the weights based on the previous stock in memory and connects them with the rest of the data [14, 17-19]

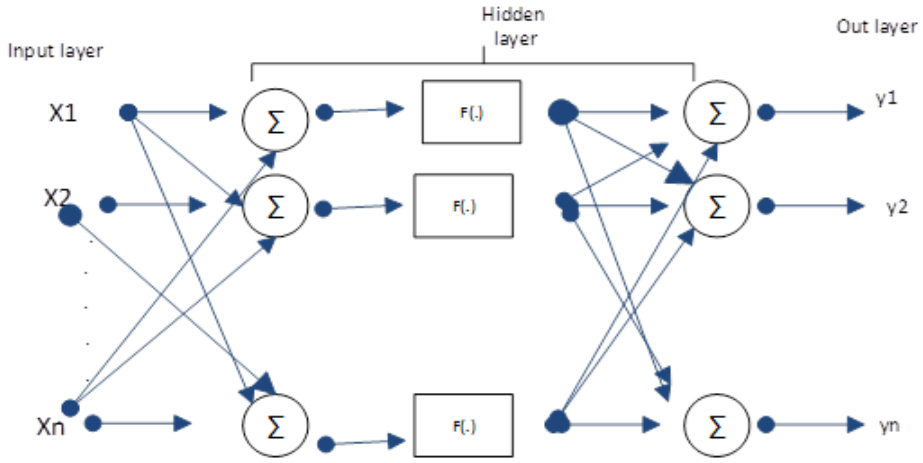


Figure 3. The structure of the iterative neural network

It is noticeable that reducing the continuous error leads to a state of breakdown. To overcome this problem, internal duplication of units has been relied upon, and this is what distinguishes this method from back propagation that does not include internal redundancy.

## 7. The practical side

In this aspect, a random sample of graduate students in the College of Administration and Economics at the University of Baghdad was withdrawn from the 2014-2015 academic year until the 2017-2018 academic year and for the sample size of 390 students. And to know the most important explanatory variables that affect the dependent variable (student success), which are, age X1, gender X2, Specialization (statistics, economics, accounting, industrial management, business administration, public administration) X3, type of study (higher diploma, master's, doctorate) X4, admission channel X5, job (employee Unemployed (X6) and using three types of grids for estimation (Forward propagation, back propagation and the Recurrent Neural Network) use the mean square error scale of the comparison.

## 8. Analysis of the results

The data for the three neural networks (FFNN, BPL, RNN) were analyzed using the 6 explanatory variables, which is equal to the number of input layer units in addition to one bias unit, and two units of output layer with different number of hidden layer units. And the ratios of the Training group is (90%, 80%, 75% and 60%) and the ratios of the test group 10%, 20%, 25% and 40%. by using the binary logistic activation function at the hidden layer and at the input layer to reach the best possible architecture for the three neural networks with the lowest possible error and the lowest prediction rate False.

### 8.1. The results of the forward propagation network application

Table No. (1), which shows the results of the forward propagation network, as the best architecture for the neural network used is (6,5,2), where the ratio for the training group was 60%, with an average error squares of 0.425, 40% of the test group, and the average error squares 0.925 and a false prediction rate 0. 037 and at a time of 751 ms.

Table 1. The results of the forward propagation network

| (FFNN) The first method Feed Forward Neural Network | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input, hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 1 | (6,1,2) | 60% | 0.441667 | 40% | 0.18 | 0.2125 | 367 |
| 2 | (6,2,2) | 60% | 0.358333 | 40% | 0.875 | 0.0375 | 504 |
| 3 | (6,3,2) | 60% | 0.408333 | 40% | 0.175 | 0.075 | 450 |
| 4 | (6,4,2) | 60% | 0.416667 | 40% | 0.9125 | 0.04375 | 665 |
| 5 | (6,5,2) | 60% | 0.425 | 40% | 0.925 | 0.037 | 751 |
| 6 | (6,1,2) | 75% | 0.293333 | 25% | 1.4 | 0.6 | 887 |
| 7 | (6,2,2) | 75% | 0.3 | 25% | 1.54 | 0.46 | 482 |
| 8 | (6,3,2) | 75% | 0.333333 | 25% | 1.64 | 0.36 | 525 |
| 9 | (6,4,2) | 75% | 0.335484 | 25% | 1.66 | 0.34 | 598 |

| (FFNN) The first method Feed Forward Neural Network | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input, hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 10 | (6,5,2) | 75% | 0.353333 | 25% | 1.68 | 0.32 | 643 |
| 11 | (6,1,2) | 80% | 0.275 | 20% | 1.225 | 0.775 | 436 |
| 12 | (6,2,2) | 80% | 0.28125 | 20% | 1.525 | 0.475 | 432 |
| 13 | (6,3,2) | 80% | 0.2125 | 20% | 1.675 | 0.325 | 528 |
| 14 | (6,4,2) | 80% | 0.325 | 20% | 1.55 | 0.45 | 588 |
| 15 | (6,5,2) | 80% | 0.33125 | 20% | 0.8 | 0.2 | 645 |
| 16 | (6,1,2) | 90% | 0.25 | 10% | 1.25 | 0.75 | 342 |
| 17 | (6,2,2) | 90% | 0.011111 | 10% | 1.1 | 0.9 | 412 |
| 18 | (6,3,2) | 90% | 0.283333 | 10% | 1.2 | 0.8 | 459 |
| 19 | (6,4,2) | 90% | 0.027778 | 10% | 1.3 | 0.7 | 501 |
| 20 | (6,5,2) | 90% | 0.294444 | 10% | 1.35 | 0.65 | 550 |

### 8.2. Results of the application of the back-propagation network

From Table ( 2) , which shows the results of the back-propagation network, as the best architecture for the neural network used is (6,2,2), where the ratio for the training group was 60% with an average error squares of 0.716, 40% of the test group, the error amount is 0.45, and a false prediction rate 0. 075 and at a time of 594 ms.

Table 2. Back-propagation network results

| (BPL) The second method Back propagation learning | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input, hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 1 | (6,1,2) | 60% | 1.4333333 | 40% | 1.4625 | 0.5375 | 469 |
| 2 | (6,2,2) | 60% | 0.7166667 | 40% | 0.4625 | 0.075 | 594 |
| 3 | (6,3,2) | 60% | 0.5666667 | 40% | 0.5375 | 0.5058824 | 147 |
| 4 | (6,4,2) | 60% | 1.4333333 | 40% | 0.5058824 | 0.5375 | 406 |
| 5 | (6,5,2) | 60% | 0.7166667 | 40% | 0.5375 | 0.5375 | 585 |
| 6 | (6,1,2) | 75% | 0.6 | 25% | 1.58 | 1.5 | 402 |

| (BPL) The second method Back propagation learning | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input, hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 7 | (6,2,2) | 75% | -0.3 | 25% | 1.9 | 1.08 | 499 |
| 8 | (6,3,2) | 75% | 0.6 | 25% | 0.3818182 | 1.5 | 910 |
| 9 | (6,4,2) | 75% | 0.5806452 | 25% | 1.58 | 1.37 | 809 |
| 10 | (6,5,2) | 75% | 0.6206897 | 25% | 0.42 | 0.5 | 1765 |
| 11 | (6,1,2) | 80% | 0.60625 | 20% | 0.8 | 0.4 | 372 |
| 12 | (6,2,2) | 80% | 0.5878788 | 20% | 1.95 | 0.05 | 475 |
| 13 | (6,3,2) | 80% | 0.303125 | 20% | 0.4 | 0.3555556 | 725 |
| 14 | (6,4,2) | 80% | 0.60625 | 20% | 0.4 | 0.6 | 977 |
| 15 | (6,5,2) | 80% | 0.60625 | 20% | 0.2 | 0.95 | 845 |
| 16 | (6,1,2) | 90% | 0.6388889 | 10% | 0.225 | 0.975 | 315 |
| 17 | (6,2,2) | 90% | 0.0944444 | 10% | 0.1 | 0.525 | 403 |
| 18 | (6,3,2) | 90% | 0.3486486 | 10% | 0.45 | 0.975 | 527 |
| 19 | (6,4,2) | 90% | 2 | 10% | 0.45 | 1.475 | 549 |
| 20 | (6,5,2) | 90% | 0.3611111 | 10% | 0.225 | 0.97 | 661 |

### 8.3. Results of the Recurrent Neural Network (RNN) Application

From Table ( 3), which shows the results of the iterative network, as the best architecture for the neural network used is (6,3,2), where the ratio for the training group was 80%, with an average of 0.103 squares of error, 20% of the test group, and the error amount was 0.002, with a false prediction rate of 0.011 and a time 353 milliseconds

Table 3. Recurrent Neural Network (RNN)

| The third method Recurrent Neural Networks (RNN) | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input,hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 1 | (6,1,2) | 60% | 0.07782 | 40% | 0.02375 | 0.2 | 972 |
| 2 | (6,2,2) | 60% | 0.07998 | 40% | 0.04496 | 0.81176 | 466 |
| 3 | (6,3,2) | 60% | 0.08 | 40% | 0.02749 | 0.83636 | 658 |
| 4 | (6,4,2) | 60% | 0.08 | 40% | 0.04583 | 0.1375 | 740 |

| The third method Recurrent Neural Networks (RNN) | | | | | | | |
|---|---|---|---|---|---|---|---|
| No | Network architecture | Training group | | Test group | | False prediction ratio | Time |
| | (Input,hidden, output) | The ratio | MSE | The ratio | MSE | | (in milliseconds) |
| 5 | (6,5,2) | 60% | 0.07782 | 40% | 0.027 | 0.8625 | 757 |
| 6 | (6,1,2) | 75% | 0.07778 | 25% | 0.018 | 1.13 | 344 |
| 7 | (6,2,2) | 75% | 0.09299 | 25% | 0.02523 | 0.23636 | 483 |
| 8 | (6,3,2) | 75% | 0.09332 | 25% | 0.02399 | 0.47619 | 684 |
| 9 | (6,4,2) | 75% | 0.09333 | 25% | 0.04 | 0.5 | 622 |
| 10 | (6,5,2) | 75% | 0.09333 | 25% | 0.024 | 0.45455 | 692 |
| 11 | (6,1,2) | 80% | 0.10417 | 20% | 0.03 | 0.09 | 246 |
| 12 | (6,2,2) | 80% | 0.10208 | 20% | 0.03313 | 0.8 | 229 |
| 13 | (6,3,2) | 80% | 0.103 | 20% | 0.002 | 0.011 | 353 |
| 14 | (6,4,2) | 80% | 0.1022 | 20% | 0.03333 | -1 | 396 |
| 15 | (6,5,2) | 80% | 0.10208 | 20% | 0.028 | 0.08889 | 473 |
| 16 | (6,1,2) | 90% | 0.05278 | 10% | 0.02115 | 0.9 | 250 |
| 17 | (6,2,2) | 90% | 0.0625 | 10% | 0.03257 | 0.8 | 224 |
| 18 | (6,3,2) | 90% | 0.10555 | 10% | 0.016 | 0.07273 | 433 |
| 19 | (6,4,2) | 90% | 0.06333 | 10% | 0.01686 | 1 | 403 |
| 20 | (6,5,2) | 90% | 0.10556 | 10% | 0.01725 | 0.8 | 536 |

From this we conclude that the best networks are the Recurrent Neural Network (RNN), as the best architecture for the iterative network is (6,3,2) and that is according to the least of error for the test group and the least false prediction at the same time To find out information about the chosen neural network, we include

the classification results (correct prediction) for the test and training groups, as in the following table.

Table 4. Classification for the test and training groups

| | | Predicted | | |
|---|---|---|---|---|
| Sample | Observed | 0 | 1 | Percent Correct |
| Training | 0 | 75 | 14 | 84.2% |
| | 1 | 10 | 157 | 94.0% |
| Testing | 0 | 30 | 7 | 81.3% |
| | 1 | 4 | 93 | 95.8% |

From Table No. 4, we find that the correct prediction rate for the two groups of training and testing was 90.6% and 91.7%, respectively, which is a good percentage. To know the Recurrent Neural Network (RNN) importance of the variables, we include the following table

Table 5. The significance of the explanatory variables

| | Importance | Normalized Importance |
|---|---|---|
| Age | .272 | 83.7% |
| Sex | .077 | 23.9% |
| Specialization | .073 | 22.6% |
| Studying | .324 | 100.0% |
| Acceptance | .087 | 26.8% |
| job | .166 | 51.3% |

Table (5) shows the most important factors that affect students 'success

Where we notice that the type of study has the greatest impact on the student's outcome, followed by age and position, and the least variable is the specialty.

## 9. Conclusions

Through a comparison between neural networks (forward spread, back propagation, and recurring network) and the influence of a number of factors (age, gender, specialization, study, acceptance channel, job) for a sample of (390) graduate students of the College of Administration and Economics for the period of the year Academic 2014-2015 through the academic year 2017-2018 University of Baghdad, it turns out:

- The best network is the recurrent neural network because it continuously adjusts the weights to fix the weights, the best recursive network architecture was (6,3,2) according to the least amount of error for the test group and the least false expectation at the same time,
- The correct prediction rate for the training and test sets was good for the recurrent neural network,
- The best variables that determine the student's position on the study (graduated or not graduated from postgraduate studies) is the type of study that came first. Then the age and job variables are among the variables that affect the student's attitude towards studies, and the least influential variables are specialization and gender.

**Declaration of competing interest**

The authors declare that they have no any known financial or non-financial competing interests in any material discussed in this paper.

**Funding information**

No funding was received from any financial organization to conduct this research

**References**

[1] K. Shiruru," Introduction to the Artificial Neural Networks", IJARIIE-ISSN (O), vol.1, pp. 2395-4396, 2016.
[2] P. Brockett,et," " A Comparison of Neural Network, Statistical Methods, and Variable Choice for Life Insurers' Financial Distress Prediction" Journal of Risk & Insurance,vol. 73(3),pp.397-419,2006.
[3] A. Alamia," Comparing feedforward and recurrent neural network architectures with human behavior in artificial grammar learning", Scientific Reports 10(1),2020.
[4] M.Ashour," Optimized Artificial Neural network models to time series" Baghdad Science Journal,2022.
[5] A. Krenker,J. Bešter,A. Kos," Introduction to the Artificial Neural Networks", Artificial Neural Networks - Methodological Advances and Biomedical Applications, Published in print April, 2011.
[6] D. Roth,C. Cheng,C. Cervantes," Introduction to Machine Learning'', Machine Learning Fall,2016.
[7] B. Windrow,A. Lehr," Artificial neural networks of the perceptron, Madaline, and Backpropagation Family", Elsevier science,1993.
[8] M. Smieja, et.al.,", Processing of missing data by neural networks", Machin learning, v3 Apr 2019.
[9] Y.Freund,D, Haussler, " Unsupervised Learning of distributions on binary vectors using two-layer networks", UCSU CRL 94-25 ,1994.
[10] M. Bocco,E. Willington ," Comparison of Regression and Neural Networks Models to Estimate Solar Radiation "CHIL. J. AGR. RES. - VOL. 70 (3),pp.428-435, 2010.
[11] H. Ullah, M. Bhuiyan," Performance Evaluation of Feed Forward Neural Network for Image Classification, Journal of Science and Technology, Vol. 10 No. 1, pp. 19-27,2018.

[12] J. Sun,Q.Chen," Optimization of Extraction Process Based on Neural Network" Asian Journal of Chemical Sciences,vol.11(2),pp.19-27,2022.
[13] H.Wang,R.Mo," . Neural network algorithm and its application in reactive distillation"procedia arxiv.org,16,2020.
[14] D. Brezak,et.al.," A Comparison of Feed-forward and Recurrent Neural Networks in Time Series Forecasting", Conference on Computational Intelligence for Financial Engineering and Economics (CIFE),2012.
[15] J. Gupta,K. Raza," Optimization of ANN Architecture: A Review on Nature-Inspired Techniques", Journals & Books, Pages 159-182,2019.
[16] T. Soares dos Santos1, D.Mendes,R.Torves, "Artificial neural networks and multiple linear regression model using principal components to estimate rainfall over South America", Nonlin. Processes Geophys., VOL.23, PP.13–20, 2016.
[17] H. ALRikabi, J. S. Qateef, A. Hadi., and R. M. Al_airaji, "Face Patterns Analysis and recognition System based on Quantum Neural Network QNN," International Journal of Interactive Mobile Technologies (iJIM), vol. 16, no. 9, 2022.
[18] I. A. Aljazaery, and R. M. Al_airaji, and A. M. Alaidi, "Automated Cheating Detection based on Video Surveillance in the Examination Classes," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 10, 2022
[19] J. Chung,C.Gulcehre,K.Cho,Y.Bengio ," Gated Feedback Recurrent Neural Networks", Computer Science Neural and Evolutionary Computing,v4 Jun 2015.