IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2023) 13 (2): 208-220

UKJAES
University of Kirkuk Journal For Administrative and Economic Science

ISSN 2222-2995
University of Kirkuk Journal For Administrative and Economic Science

Mahmood Saman Hussein & Khider Hoshamin Haji. Comparison of Linear Discriminant Analysis and Artificial Neural Networks for Stroke patients Classification. *University of Kirkuk Journal For Administrative and Economic Science* (2023) 13 (2): 208-220.

Comparison of Linear Discriminant Analysis and Artificial Neural Networks for Stroke patients Classification

Saman Hussein Mahmood [1], Hoshamin Haji Khider [2]

[(1,2)] *Statistics and Informatics Department/Administration and Economics College/ Salahaddin University-Erbil, Erbil, Iraq*

Saman.mahmood@su.edu.krd [1]
Hoshamin.haji@yahoo.com [2]

Abstract : Stroke is a chronic disease that affects the blood vessels in the brain and leads a blockage the blood flow to a part of the brain is cut off or reduced, and thus deprived brain tissue of oxygen and nutrients and brain cells begin to die. Due to the importance of this subject in our daily life, the research aims to study the subject by building three predictive models for two mutually exclusive groups (dead and live) using Linear Discriminant Analysis (DA) and Artificial Neural Networks (ANN) and compare the classification results of the two methods. Therefore, the data sample of 134 cases collected from stroke patients in Rzgari Hospital and Rojhalat Emergency Hospital in the capital of Kurdistan region Erbil. The results, build three models each of DA and ANN for males, females, and both genders together and then compared the models by correct classification rate, we found that the DA and ANN models to be equal by results for males and females separately (98.7% and 100%). But the DA model for both genders is better than the ANN model. Finally, we identified the most important independent variables influencing the patient's condition three months after strokes.

Keywords: Neural Networks Analysis, Discriminant analysis, Classification, Brain Stroke.

مقارنة التحليل التمييزي الخطي والشبكات العصبية الاصطناعية لتصنيف مرضى السكتة الدماغية

أ.م.د. سامان حسين محمود [١] ، م.م. هوشمين حاجي خدر [٢]

قسم الاحصاء والمعلوماتية/ كلية الادارة والاقتصاد/ جامعة صلاح الدين/اربيل،اربيل، العراق [(١،٢)]

المستخلص: السكتة الدماغية هي مرض مزمن يصيب الأوعية الدموية في الدماغ ويؤدي إلى انسداد تدفق الدم إلى جزء من الدماغ أو ينخفض، وبالتالي تحرم أنسجة المخ من الأكسجين والمغذيات وتبدأ خلايا الدماغ في الموت. نظرًا لأهمية هذا الموضوع في حياتنا اليومية. يهدف البحث إلى دراسة الموضوع من خلال بناء ثلاثة نماذج تنبؤية

لمجموعتين متنافيتين (ميتة وحية) باستخدام التحليل الخطي التمييزي (DA) والشبكات العصبية الاصطناعية (ANN) والمقارنة نتائج التصنيف للطريقتين. لذلك ، تم جمع عينة بيانات من ١٣٤ حالة من مرضى السكتة الدماغية في مستشفى رزكاري ومستشفى روزهلات للطوارئ في عاصمة إقليم كردستان أربيل. النتائج ، تم بناء ثلاثة نماذج لكل من DA و ANN للذكور والإناث وكلا الجنسين معًا ثم مقارنة النماذج حسب معدل التصنيف الصحيح ، وجدنا أن نموذجي DA و ANN متساويان بالنتائج للذكور والإناث بشكل منفصل (٩٨.٧ ٪ و ١٠٠ ٪). لكن نموذج DA لكلا الجنسين أفضل من نموذج ANN. أخيرًا ، حددنا أهم المتغيرات المستقلة التي تؤثر على حالة المريض بعد ثلاثة أشهر من السكتات الدماغية.

الكلمات المفتاحية: تحليل الشبكات العصبية، التحليل التمييزي، التصنيف، السكتة الدماغ.

Corresponding Author: E-mail: Saman.mahmood@su.edu.krd

Introduction

Classification is a technique used to predict group membership and identify instances of data. Classification involves building a function that maps the input feature space to an output space of two classes. There are several methods used in the classification process, including Artificial Neural Networks and Discriminant Function Analysis. Neural Networks are an effective tool in the field of classification, using to training feed forward neural network and Back propagation Algorithm is a technique to build a model. However, the success of the networks is highly dependent on the performance of the training process and hence the training algorithm. The application of classification of artificial neural networks and discriminatory function analysis to stroke data due to the increase in the number of deaths and disabilities in the Kurdistan region and the worldwide. According to WHO report, 15 million people worldwide suffer from a stroke each year. Of these, 5 million die and another 5 million are permanently disabled.

In this research the dependent variable is status of patients after three months from brain stroke. The categories are died and alive. Discriminant analysis is another technique used to analyze study data when the dependent variable is categorical and the independent variable is a set of interrelated variables. The term categorical variable means that the dependent variable is divided into a number of categories.

When following up the status of patients after three months from brain stroke, we observed that 63.4% of the patients survive from this disease and 36.6% of the patients died due to the same medical emergency. In this research we built 6 functions by using (DA) and (ANN). Classification results from discriminant analysis for both genders together is able to significantly improve status of patient after three months from brain stroke prediction accuracy when compared to artificial neural networks. The best efficiency is 87.8% and 85.5%, respectively. But correct classification rate for male and female separately in artificial neural networks and discriminant analysis are not unequal.

Methodology

1. Discriminant Analysis:

Discriminant analysis (DA) is a multivariate statistical tool used to classify a set of objects or individuals into one of the predetermined groups. DA is also a tool to solve the problem of separation where a sample of n objects will be separated into predefined groups. We wish to develop a rule to discriminate between the objects belonging to one group and the other. The result of the classification rule then becomes a random variable and some misclassifications may occur. The objective therefore is to arrive at a rule that would minimize the percentage of misclassification (McLachlan, 1992).

Linear Discriminant Analysis (LDA) is a statistical procedure proposed by R.A. Fisher (1936) in which multiple linear regression is used to relate the outcome variable (Y) with several explanatory variables, each of which is a possible marker to determine Y. When the outcome is dichotomous (taking only two values) the classification will be binary. This is not the case with the usual multiple linear regression where Y is taken as continuous. The regression model connects Y with the explanatory variables, which can be either continuous or categorical. In LDA the dichotomous

variable (Y) are regressed on to the explanatory variables as done in the context of multiple linear regressions (Sarma & Vardhan,2019).
We confine our attention to Fisher's Linear Discriminant Function only. Let μ1 and μ2 be the means vectors of the two multivariate populations with covariance matrices Σ1 and Σ2 respectively. LDA is useful only when H0: μ1 = μ2 is rejected by MANOVA. In other words, we need the mean vectors in the two groups shall differ significantly in the two populations.

2.Assumption:

The following conditions are assumed in the context of LDA(Sarma & Vardhan,2019).

a. The two groups have the same covariance matrix ($\Sigma_1 = \Sigma_2 = \Sigma$). In the case of unequal covariance matrices, we use another tool called quadratic discriminant analysis or we can use LDA.
b. A normality condition is not required. When this condition is also true then the Fisher's method gives optimal classification.
c. Data should not have outliers. (If there are a few they should be handled suitably).
d. The size of the smaller group (number of records in the group) must be larger than the number of predictor variables to be used in the model.

3. Some basic concepts of Discriminant analysis:

a: Linear discrimination function:

This type of discrimination is one of the simplest cases of discrimination and requires the following conditions (Afifi, May,& Clark, 2012):

(1) It is assumed that the explanatory variables are distributed in a multivariate natural distribution.

(2) The variances for all the studied totals (Matrices of Contrast and Contrast) are equal, ie, acceptance of the null hypothesis when testing the hypothesis:

$$\mathbf{H_0 = \Sigma_1 = \Sigma_2 = \cdots = \Sigma_{\mathit{k}}}$$
$$\mathbf{H_1 = \Sigma_1 \neq \Sigma_2 \neq \cdots \neq \Sigma_{\mathit{k}}}$$

Where: **Σ:** Matrix of Contrast and Contrast , **k:** Number of totals

(3) Categorize existing views to totals ..., **n2, n1** accurately.

b: The discriminant function for two groups or several groups:

We assume that the two populations to be compared have the same covariance matrix **Σ** but distinct mean vectors μ_1 and μ_2. We work with samples **y11, y12, ...,y1n1** and **y21, y22, …, y2n2** from the two populations. As usual, each vector yij consists of measurements on p variables that maximizes the distance between the two 'transformed' group mean vectors. A linear combination $\mathbf{z = \acute{a}y}$ transforms each observation vector to a scalar (RENCHER,2002).

In discriminant analysis for several groups, we are concerned with finding linear combinations of variables that best separate the k groups of multivariate observations. Discriminant analysis for several groups may serve any one of various purposes: (RENCHER,2002)

(1) Examine group separation in a two-dimensional plot. When there are more than two groups, it requires more than one discriminant function to describe group separation. If the points in the **p**-dimensional space are projected onto a two-dimensional space represented by the first two discriminant functions, we obtain the best possible view of how the groups are separated.

(2) Find a subset of the original variables that separates the groups almost as well as the original set.

(3) Rank the variables in terms of their relative contribution to group separation.

(4) Interpret the new dimensions represented by the discriminant functions.

(5) Follow up to fixed-effects **MANOVA**.

c: Wilks's lambda:

The Wilk’s Lambda is the statistic measures discrimination via multivariate group differences. Lambda values range between 0 and +1.0. The smaller and closer to zero the lambda value is, the more separation there is between groups relative to within groups and vice versa. Wilks’s lambda can be interpreted as a test of inference using the chi-square distribution. Statistically significant Λ values indicate the discriminant function is statistically significant, and this can be interpreted as evidence of discrimination between groups. When we have only two groups in the dependent variable, only one Λ value will be computed. (Hahs-Vaughn, 2017)

d: Stepwise Selection of Variables:

When in the LDA phenomenon research we have a lot of predictors, the stepwise method can be useful by automatically selecting the "best" variables to use in the model (SPSS,2020). The stepwise method for selecting variables in discriminant analysis is rather like doing a stepwise regression and is especially useful in similar circumstances, namely when we have rather a long list of possible classification variables and it is unlikely that all will make a useful contribution to a set of discriminant functions. We would like to find the best subset, or else something close to that. The single variable that gives the significant classification into our groups is chosen first, after that we look at the remaining variables and add the one that gives the biggest improvement. We check the two variables now, and make sure that each makes a significant contribution in the presence of the other. At each step we see whether another variable can be added that will make a significant improvement, and whether any previous ones can be removed. The process stops when no more variables can be added or removed at the level of significance we are using.

4. Artificial Neural Networks (ANN)

Scientists have found that the brain (cerebral cortex) which located in the head contains approximately 14 to 16 billion neurons and the estimated number of neurons in the cerebellum is 55 to 70 billion. Each neuron is connected by synapses to several thousand other neurons. This neuron that manages the mechanism of neuronal action drives neuroscientists, psychologists and computer engineers try to simulate the human mind so that they can ultimately build a structure for giant computers that simulates the work of the human mind [6].
A neural network is characterized by its pattern of connections between the neurons (called network architecture) and its method of determining the weights on the connections (called training or learning algorithm). The weights are adjusted on the basis of data. In other words, neural networks

learn from examples and exhibit some capability for generalization beyond the training data (**Haykin,1999)**.
Artificial neural networks are viable computational models for a wide variety of problems. Already, useful applications have been designed, built, and commercialized for various areas in engineering, business and biology(**Dayhoff,1990)(Murray,1995)**.
Although they may have been inspired by neuroscience, the majority of the networks have close relevance or counterparts to traditional statistical methods(**Sarle,1994)(Bishop,1995),** such as non-parametric pattern classifiers, clustering algorithm, nonlinear filters, and statistical regression models(**Wu & McLarty,2000)** .

5. Artificial Neural Network Parameters:

There are a number of dissimilar parameters that must be decided upon when designing a neural network. Amid these parameters are the number of layers, the number of neurons per layer, the number of training iterations, et cetera. Some of the more important parameters in terms of training and network capacity are the number of hidden neurons, the learning rate and the momentum parameter (Tesfamariam & Najjaran,2007).

The algorithm for adjusting the weights of different types to train the NN model and assess its performance by obtaining the lowest error rate of the training and testing models in the tiniest number of epochs. Also influence of the number of hidden layers, number of neurons in the input layer, the hidden layers, and the usage of various activation functions in the hidden layer and the output layer, also the pre-processing of the network training models by converting them into numerical values adequate with the nature of the input of the network. This was talked over and worked on in the practical part of present thesis.

6. Use of Neural Networks in Prediction:

Lately researchers have begun to use the methods of neural networks in processors Statistical, and predictive use of neural networks, where neural networks are used sometimes for short-term forecasting especially in electronic industries and this modern technology provides a flexible function which can fit nonlinear models, including neural network models which are in fact flexible nonlinear functions and the general shape of the neural network function used.

We use neural networks in the prediction follows, the following formula (Floyd et al,1994):

$$Y = F[H_1x, H_2x, \dots \dots H_nx] + \mu$$

Where:

Y: Represents the dependent variable also called output.
x: Represents the explanatory variable or independent variables.
$H,\ F$: Represent the functions of neural networks
μ: The error limit in the function
(H) Activation functions for hidden layers, and (F) outputs of the hidden activation function.

The difference in the use of neural networks model to predict other models such as Linear regression model the regression function is linear, and there are no functions for the hidden layers in the regression model, while neural networks need to input units also ANNs are choose a few hidden nodes when you start training and then increase the hidden nodes until we get to a less square error. There are several steps to predict (Yegnanarayana,2009):

- **a.** Test variables.
- **b.** Data processing by performing some operations on the data used.
- **c.** Split the data into totals (Training set, Test set & Validation set).
- **d.** Identify the neural network model.
- **e.** Determine a standard for evaluation.
- **f.** Network training.
- **g.** Network implementation.

7. Neural Networks as a Classification Model

Neural networks return continuous value at the exit; hence they are excellent for estimating and classification. Such networks may analyze many variables at a time. It is possible to create a model even if the solution is very complex. The drawbacks of neural networks are the difficulties in setting architectural parameters, falling into local minima, long learning process and lack of clear interpretation(Krawiec,2004) (Rutkowska,Piliski,&Rutkowski,1997)

Application

We depended on the real data of applied analysis; the data set for this study about stroke was collected in Rzgari Hospital and Rojhalat Emergency Hospital in the capital of Kurdistan region Erbil/Iraq. The data contains 134 patients or cases 78 of the cases are male and 56 of the case are female. Out of those patients 49 died during the study and 85 survived (alive). The data collected randomly by questioner of the research consisted of 29 variables. The dependent variable of the study is the status of those patients which faced brain stoke after three months. Twenty eight variables (Parameters) of the study are independent.
We applied two methods of classification. In the Discriminant analysis we used stepwise method to determine the most important independent variables. It is important to know that in the Artificial Neural Networks we depend on those independent variables we have in the discriminant analysis model. For building the best models, we split the data for male, female and both genders together, and then make 6 models to evaluate the classification accuracy of the models, Each model is classified according to the patient's condition (Died or Alive)). Finally, from the final models, we determine the most important independent variables affecting the status of patient after three months from brain strokes either the patient died or survive.

1: Applied Discriminant Analysis (DA) Method

In the applied of DA we should investigate the data to get the best model. Also in the apply discriminant analysis stepwise method used to get the best models for classification:

Table (1): Eigenvalues and Wilks' Lambda based on Gender

Test of Function	Wilks' Lambda	Chi-square	df	Sig.
Male function	**.100**	**147.211**	**14**	**.000**
Female function	**.168**	**84.705**	**7**	**.000**
both genders together function	**.455**	**95.747**	**5**	**.000**

Table (2) Variables in the analysis by stepwise method

Table (2) 14 Variables in the Analysis for Male			
Variable	**Tolerance**	**F to Remove**	**Wilks' Lambda**
Quadriplegia	**0.55**	**11.1**	**0.12**
Size (mm)	**0.4**	**19.7**	**0.13**
facial or perioral numbness	**0.11**	**80.4**	**0.24**
Intra- Ventricular extension	**0.35**	**53.4**	**0.19**
Lt Side weakness	**0.26**	**67.3**	**0.22**
Chief complaint	**0.47**	**24.9**	**0.14**
Disturbed consciousness	**0.41**	**26.8**	**0.15**
duration spent for reaching hospital	**0.33**	**36.7**	**0.16**
Diplopia or visual disturbance	**0.26**	**11.1**	**0.12**
Speech difficulty	**0.54**	**17.9**	**0.13**
Seizure	**0.64**	**11.7**	**0.12**
midline shift	**0.34**	**12.5**	**0.12**
Rate Blood Suger	**0.48**	**10.9**	**0.12**
Systolic Blood Pressure	**0.48**	**5.41**	**0.11**

Table (3) 7 Variables in the Analysis for Female			
Variable	**Tolerance**	**F to Remove**	**Wilks' Lambda**
midline shift	**0.48**	**97.43**	**0.532**
Age	**0.755**	**10.94**	**0.209**
Smoking	**0.67**	**24.278**	**0.259**
Diastolic Blood Pressure	**0.536**	**26.835**	**0.268**
Systolic Blood Pressure	**0.36**	**31.895**	**0.287**
CT scan for detecting bleeding	**0.694**	**18.462**	**0.237**
Address	**0.719**	**8.895**	**0.201**

Table (4) 14 Variables in the Analysis for Male and Female			
Variable	**Tolerance**	**F to Remove**	**Wilks' Lambda**
Intra- Ventricular extension	0.88	9.765	0.492
Lt Side weakness	0.992	20.105	0.531
midline shift	0.912	20.404	0.532
Age	0.903	10.715	0.495
Vertigo	0.97	7.907	0.485

Table (2,3,4) shows Wilk's lambda, F, and significance values contribute information about difference means for each variable based on variable gender. It is immediately clear that the stepwise rule here was to minimize Wilks' Lambda at each step. It also we should to know that tolerance is the proportion of a variable's variance not accounted for by other independent variables in the equations. A variable with very low tolerance contributes little information to a model and can cause estimation problems. F to remove values is useful for describing what takes place if a variable is removed from the current models (given that the other variables stay). F to remove for the entering variable is the same as F to Enter.

2: Classification Results for Discriminant analysis

Table 5: Classification Results According to Gender

Gender		Predicted Group Membership				Total
		Alive	%	Died	%	
Male	Alive	52	100%	0	0%	52
	Died	1	4.30%	22	95.70%	23
98.7% of original grouped cases for male correctly classified.						
		Predicted Group Membership				Total
		Alive	%	Died	%	
Female	Alive	31	100%	0	0%	31
	Died	0	0%	25	100%	25
100% of original grouped cases for female correctly classified.						
Male and Female		Predicted Group Membership				Total
		Alive	%	Died	%	
	Alive	76	91.60%	7	8.40%	83
	Died	9	18.80%	39	81.30%	48
87.8% of original grouped cases correctly classified.						

Table (5) illustrated a simple summary of number and percent of patients classified correctly and incorrectly based on variable gender. It is immediately clear that the classification results for male 100%, and 95.7% of dead and alive patients (cases) were classified correctly. Only 4.3% of dead and alive patients were misclassified. It is also clear that the classification results for female 100%, and 100% of dead and alive patients were classified correctly. It means that we do not have misclassified. It must be pointed out the classification results for both genders 91.6%, and 81.3% of dead and alive patients (cases) were classified correctly. Only 18.8% of dead and alive for both genders and 8.4 of dead and alive for both genders patients were misclassified. We should note that the percentages for both genders are less than the percentages of male and female separately. The main aim of doing analysis for male, female and both gender separately were to get the best model for well forecasting. In other words 98.7% of original grouped cases for male, 100% of original grouped cases for female and 87.8% of original grouped cases for both gender classified correctly.

3: Applying Artificial Neural Networks (ANN) Method

In the (ANN) application, a multilayer perceptron (MLP) was used. ANN analysis we relied on the results of the discriminant analysis in the previous section, and selected those variables in those models on the basis of the variable gender.

4: Structure of the MLP ANN according to Variable Gender

Figure (1,2,3) shows the structures of the MLP ANN for male, female and both genders together. Each one of structure is known as feed-forward architecture because the links in the network flow forward from the input layer to the output layer without any feedback loops. The input layer depends on the variables mentioned in tables (2,3,4). The hidden layer of the structures contains unobservable units or nodes. The output layer contains the status of patient after three months from brain stroke (alive or died).

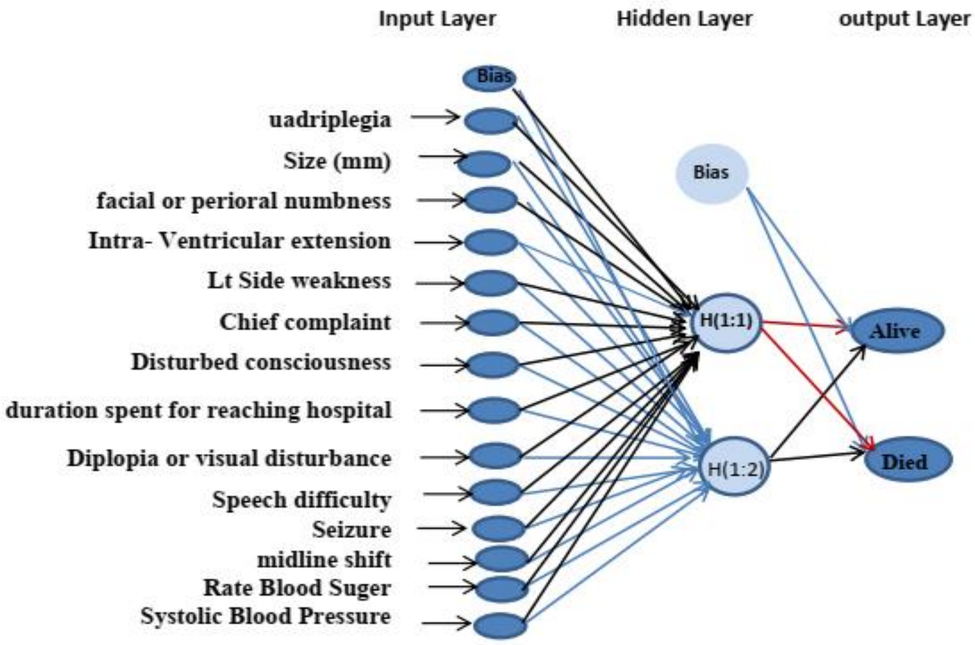

Figure 1 Structure of the MLP ANN for Male

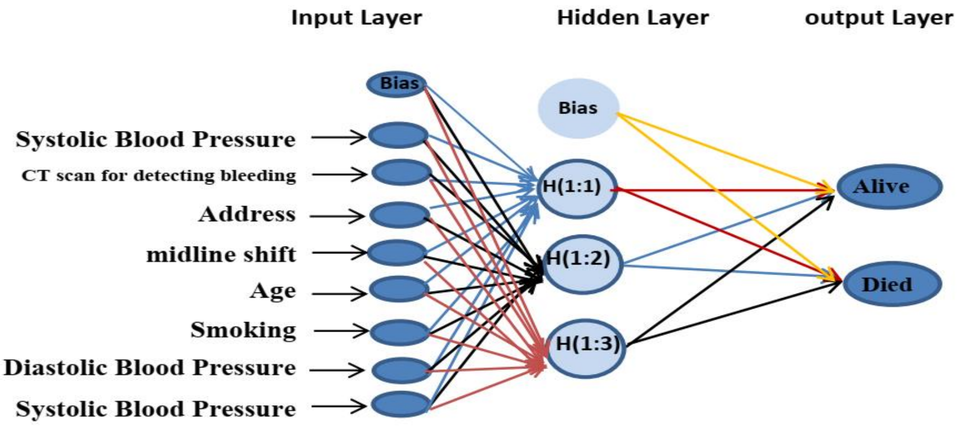

Figure 2 Structure of the MLP ANN for Male

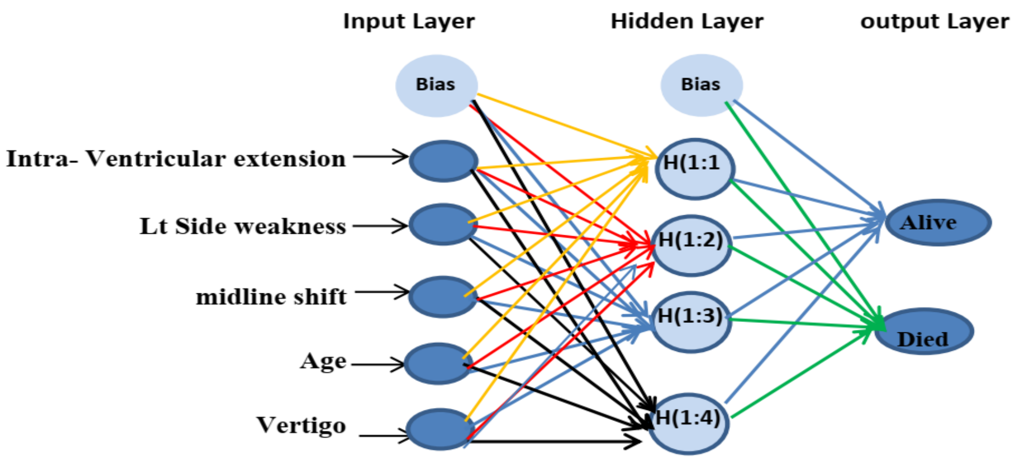

Figure 3 Structure of the MLP ANN for both genders together

5: Classification

Table (6) displays the cross classification according to performance of the MLP ANN for male, female and both genders together. It is immediately clear that The percent of correct classification for the training set for the sample of male is found to be 100 where for the testing group it found to be 95.5 We can similarly interpret the percent of correct classification for the training and testing set of the sample female and sample of both genders together the same as interpretation of sample male.

Table 6 : Cross Classification of used cases with respect to Categories of Status based on Variable Gender

	ANN	*Status*	Predicted			
			Alive	Died	Percent Correct	
Male	Training	Alive	40	0	100%	100%
		Died	0	13	100%	
	Testing	Alive	12	0	100%	95.50 %
		Died	1	9	90.00%	
Overall Percent		$\frac{74}{75} =$ 98.7% of original grouped cases correctly classified.				
	ANN	*Status*	Predicted			
			Alive	Died	Percent Correct	
Female	Training	Alive	23	0	100.0%	100%
		Died	0	20	100.0%	
	Testing	Alive	8	0	100.0%	100%
		Died	0	5	100.0%	
Overall Percent		$\frac{56}{56} =$ 100% of original grouped cases correctly classified.				
	ANN	*Status*	Predicted			
			Alive	Died	Percent Correct	
Male and Female	Training	Alive	52	5	91.20%	86.95%
		Died	7	28	80.00%	
	Testing	Alive	23	3	88.50%	82%
		Died	4	9	69.20%	
Overall Percent		$\frac{112}{131} =$ 85.5% of original grouped cases correctly classified.				

Selecting the most important independent variables

Figure (4) contained three Independent variable importance charts or figures for male, female and both genders together. Each one of these charts shows the normalized importance of the input variables sorted in an ascending manner. In the chart of male it is immediately clear that [size] (size of haemorrhage) accounted for the highest normalized importance which indicates its role in the judgment of the cases to either group or category of the dependent variable status of patient after three months from brain stroke. The variables LTWK (left side weakness, Quad (Quadriplegia), CCcode (Chief complaint), timecode (duration spent for reaching hospital), facialnum (facial or perioral numbness), IVE (Intra- Ventricular extension) and SPDIF (Speech difficulty) have a remarkable normalized importance but less than that of the [size]. We can interpret female and gender schemata together, and note that (age, midline shift (mid), Systolic Blood Pressure(SBPC) and Vertige) has a important effect on stroke but in different proportions.

Figure (4): The normalized importance for the input variables used to produce the MLP ANN based on variable gender

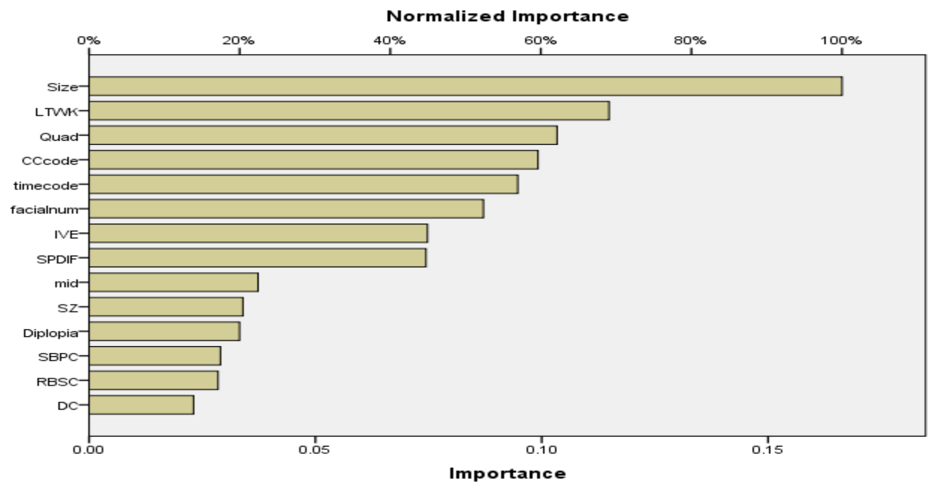

Figure A. The normalized importance for the input variables for Male

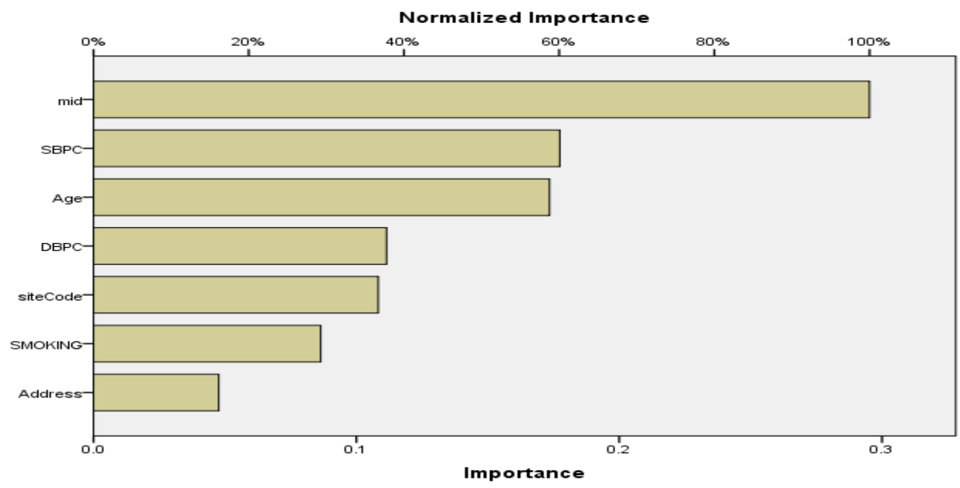

Figure B. The normalized importance for the input variables for Female

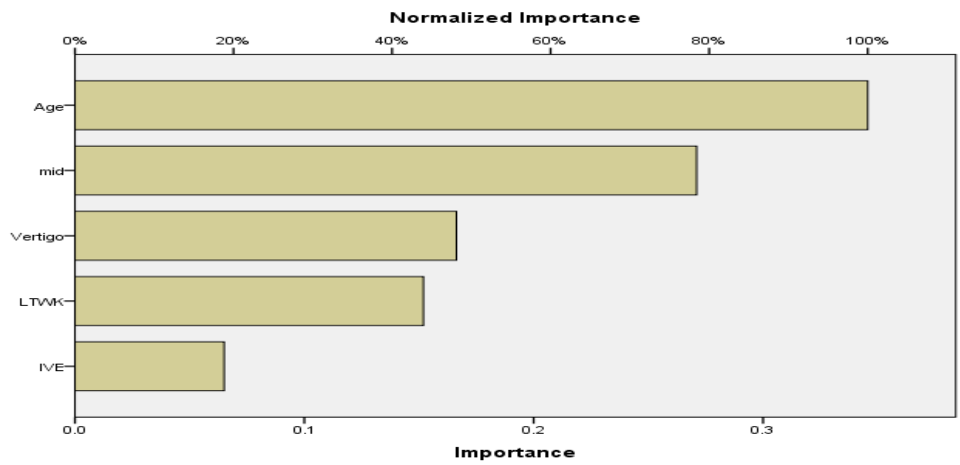

Figure C. The normalized importance for for Male and Female

6: Comparison of various models: DA vs. ANN

To compare the performance of the NN approach with the DA technique, the Mean Correct Classification Rate of predictive accuracies in neural network and discriminant analysis models are shown in Table (7). Clearly, the discriminant analysis for both genders together, demonstrates a superior ability to predict the status of patient after three months from brain stroke.

Table (7): Classification results from Discriminant Analysis and ANN

	Correct Classification Rate % for		
Model	**Male**	**Female**	**Both Genders Together**
Discriminant Analysis	98.7%	100%	87.8%
ANN	98.7%	100%	85.5%

Table (7) shows the classification results for the discriminant analysis model and the neural network model based on variable gender described in this thesis. The average classification correct rate for the Discriminant Analysis for both genders together was higher than the Artificial Neural Networks for both Genders together. The best efficiency is 87.8% and 85.5%, respectively. It is also noticeable that the average classification correct rate for the Discriminant Analysis for male and the Artificial Neural Networks for male are equal. It is also clear that the average classification correct rate for the Discriminant Analysis for female and the Artificial Neural Networks for female are not difference.

Conclusions

In the light of the study results by using Discriminant analysis and ANN, the following points are the main conclusions:

1. During a follow-up of patients three months after their stroke, we observed that 63.4% of patients survived and 36.6% died due to the same medical emergency.
2. After using stepwise method in the discriminant analysis according to variables gender, it indicated that the number of independent variables in the discriminant function model for male and the model for female are more than the number of independent variables in the model for both genders.
3. In this research we built 6 functions by using discriminant analysis and neural network analysis for classification status of patient after three months from brain stroke. This paved the way to the correctly classifications for male and female are higher than from the correctly classifications for both genders.
4. The classification results from discriminant analysis and ANN prove that both sexes discriminant analysis together is better efficient compared to artificial neural networks, but the correct classification rate for males and females separately in artificial neural networks and discriminant analysis are equivalent.
5. A set of important independent variables were selected from the models ANN, the most important of which are (hemorrhage volume, left side weakness, quadriplegia, main complaint, time taken to reach hospital, age, midline shift, systolic and pulmonary blood pressure)

References

1- Afifi A, May S, Clark VA " Practical Multivariate Analysis" . 5th ed. Boca Raton: CRC Press , 2012.
2- Bishop C. M. " Neural Networks for Pattern Recognition", Clarendon Press, Oxford, Birmingham, UK , 1995.
3- Dayhoff, J. E "Neural network architectures: an introduction" Van Nostrand Reinhold, New York, United States ,1990
4- Floyd, C. E., et al. "Prediction of breast cancer malignancy using an artificial neural network". Cancer, 74(11), PP(2944-2948), 1994.
5- Hahs-Vaughn,D.L. APPLIED MULTIVARIATE STATISTICAL CONCEPTS. New York and London: Routledge, 2017.
6- Haykin, S. "Neural Network A Comprehensive Foundation",2nd edition. London: Prentice Hall ,1999 .
7- Krawiec, K., & Stefanowski, J. "Machine Learning and Neural Networks" . Publishing House of Poznan university of technology. Poznan , 2004.
8- McLachlan, G. J. "Discriminant Analysis and Statistical Pattern Recognition". New Jersey, USA: John Wiley & Sons , 1992.
9- Murray,A. F "Applications Of Neural Networks" Eds. Kluwer Academic Publisher, Dordrecht,Netherlands 1995.
10- RENCHER, A. C. " Methods of Multivariate Analysis. Brigham" A JOHN WILEY & SONS, INC ,2002.
11- Rutkowska, D., Piliski, M., & Rutkowski, L. "Neural Networks, Genetic Algorithms and Fussy Logic". Warsaw: PWN. 1997.
12- Sarle , Warren S:" Neural Networks and Statistical Models", Proceeding of The Nineteenth Annual SAS Vsers Group International Conference, SAS Institute Inc., Cary, NC, USA, pp1538-1550, 1994.
13- Sarma, K. V., & Vardhan, R. V. "Multivariate Statistics Made Simple". New Yourk: Taylor & Francis Group. 2019.
14- SPSS."Stepwise Discriminant Analysis" .Retrieved from IBM:https://www.ibm.com/docs/en/spss-statistics/25.0.0?topic= analysis-discriminant-stepwise-method (2020)
15- Tesfamariam, S., & Najjaran, H. "Adaptive Network-Fuzzy Inferencing to Estimate Concrete Strength Using Mix design". Journal of Materials in Civil Engineering, 550-560, 2007.
16- Wu , C. H., & McLarty , J. W., "Neural Networks and Genome Informatics.USA: Elsevier Science,2000.
17- Yegnanarayana, B . "Artificial Neural Networks". Prentic Hall of India, New Delhi, 2009