Sapporo Medical Journal
Volume 55, Issue 01, January, 2021

Prediction the best predictor for urea reading among diabetic patients using artificial neural networks (ANNS) models

Nor Farid Mohd Noor[1*], Wan Muhamad Amir W Ahmad[1], Mohamad Arif Awang Nawi[1], Farah Muna Mohamad Ghazali[1], Nor Azlida Aleng[2], Ramizu Shaari[1], Ahmad Mukifza Harun[3], Razif Abas[4]

[1]School of Dental Sciences, Health Campus, UniversitiSains Malaysia,Kota Bharu, Kelantan, Malaysia
[2]Faculty of Ocean Engineering Technology and Informatics,
Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia
[3]Faculty of Engineering, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
[4]Department of Human Anatomy, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia,Serdang, Selangor

Abstract— This study aims to develop the Artificial Neural Network (ANN) through Multilayer Perceptron Neural Network (MLP) by considering the bootstrapping methodology. Applying the bootstrapping approach in MLP methodology improves the precision of the related urea level determination factor. This model developed to determine urea reading among diabetic patients. Three blood parameter Fasting Blood Glucose (X1), HbA1c (X2), and Sodium Reading (X3) were selected according to their clinical importance. All these parameters will be used as input for urea determination. Using The ANN-MLP Model the performance of analysis will be determined through the Predicted Mean Square Error (PMSE) obtained from (MSE-forecasts the Network). In this research paper, all possible combinations of input will be evaluated one by one. The performance of MLP was evaluated through the PMSE of the neural network for the (MSE-forecasts the Network) and special attention will be given for the smallest value of PMSE reading while running the analysis. In this study, PMSE is used as a measurement for the goodness of fit test of the obtained model. It can be used as a tool to measure how far the prediction value from the actual value. The smallest PMSE will indicate the excellent performance of the model. In conclusion, a combination of these three variables which were Fasting Blood Glucose (X1), HbA1c (X2), and Sodium Reading (X3) contributed significantly to the area level through the developed methodology.

Keywords—Multilayer perceptron neural network (MLP), predicted mean square error (PMSE), urea reading

1. Introduction

Serum urea is one of the common factors assessed in a patient, especially for diabetes mellitus. Certain conditions are commonly increased the serum urea level such as dehydration state. Besides, it can be used to measure pathological progression. For example, serum urea levels can be used to predict the complication or advancement of diabetes mellitus. Here, the risk of lower limb amputation in the diabetic patient is suggested related to their serum urea in a clinical study [1]. The low level of urea may be used to predict other conditions. For example, the reduction of urea level is used as one of the indicators in assessing the slow progression of diabetic nephropathy in rats receiving Asprosin [2]. The serum urea can be used in addition to other tests. Together with fasting blood sugar, serum urea is one of the good predictors for major foot amputation in the diabetic patient [3]. The progression of the pathology which is related to its signs of the complications can be predicted by using serum levels of a certain test. Highly correlation of HbA1c and serum urea is diabetic nephropathy in diabetes patients which can be used in managing the progression of the disease [4]. In contrast, serum urea is found to contradict the serum atrium in chronic renal failure patients [5].

Efron 1979 introduced the resampling technique known as the bootstrap method. The concept behind bootstrap is to use a sample in hand as a population to take a sample (by replacement) of the sample in hand and then make a huge number of "case re-sampling samples" known as bootstrap samples. The bootstrap process starts with the original sample obtained from a particular population under consideration. The next move is to repeat the original sample several times to establish a new population, taking into account the existing population. In any case, the bootstrap draws many samples to be replaced by a random sampling method and, as a result, produces a new sample from the origin. These technique stores new data sets and create new distributions for further analysis [9]. The advantage of using bootstrap is its ability to extend the sample to the same size as the original, which may include some observations while eliminating other observations.

Artificial neural networks (ANN), also known as neural networks (NNs),is a computational model based on the structure and functions of biological neural networks. A multilayer perceptron (MLP) is a class of feedforward artificial neural network (ANN) with one or more layers between the input, hidden, and output layer. In the MLP model, the analysis of the output node of this analysis is fixed at one since there is only one dependent variable. Figure 1gives the MLP with N input nodes, H hidden nodes, and one output node. The values of $\hat{y}$ are given as follows $\hat{Y} = g_i\left(\sum_{j=1}^{H} w_j h_j + w_0\right)$ where w_j an output weight from hidden node j to the output node, w_0 the bias for the output node, and g is an activation function. The values of the hidden node h_j, j =1…H are given by $h_j = g_i\left(\sum_{j=1}^{H} v_{ji} x_i + v_{j0}\right)$ where v_{ji} the output *weight* from input node i to hidden node j, v_{j0} is the bias for hidden node j where j =1, …, H and x_i are the independent variables where i =1, …, N and k is an activation function [6, 7, 8].The general architecture of the MLP is shown in Figure 1.

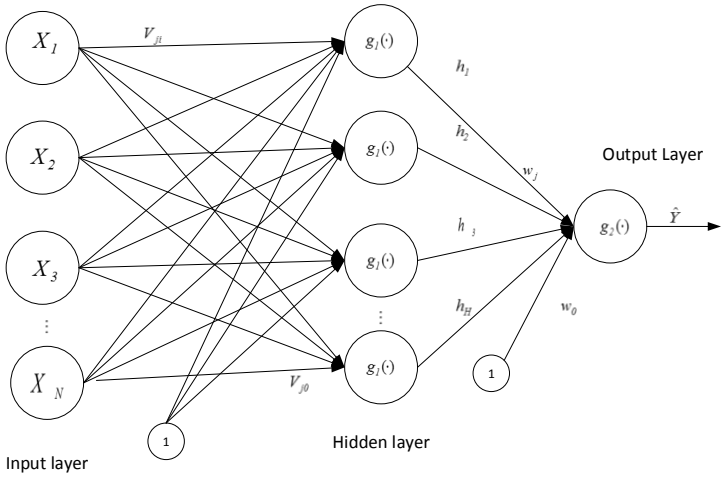

Fig. 1: The general architecture of the MLP with one hidden layer, N input nodes, H hidden nodes, and one output node

2. Data and the R syntax

Data from the medical unit record, Hospital USM were collected, reviewed and the related information was extracted. The sampling frame was the list of patients, which diagnosed with hypertension and diabetes. All related variable is being collected and summarized in Table1.Figure 2 shows the flow chart of the multilayer perceptron neural network (MLPNN).

Table 1: Data description

Variable	Code	Description
Urea	Y	Urea Reading
FBS	X1	Fasting Blood Glucose
HbA1c	X2	HbA1c
Sodium	X3	Sodium Reading

2.1 Below is the R syntax for the artificial neural network models

```
#/Complete Dataset for a Diabetic Patient/

Input =("
FBS HbA1c Urea Sod
5.300 7.20 5.70 142.00
11.700 9.70 2.90 133.00
7.400 7.50 5.70 142.00
4.600 6.20 3.90 139.00
6.400 8.20 5.20 139.00
9.800 6.90 5.80 134.00
4.400 5.90 4.50 137.00
9.900 9.40 6.60 140.00
6.100 5.80 4.70 136.00
14.700 10.80 13.90 138.00
4.800 8.60 7.50 139.00
5.800 5.90 27.00 139.00
6.000 8.20 6.30 140.00
6.000 8.20 6.30 140.00
5.200 6.20 5.20 142.00
10.200 8.00 7.50 137.00
8.000 9.00 4.30 141.00
7.600 7.50 3.20 134.00
8.700 8.20 2.30 140.00
8.700 8.20 2.30 140.00
6.800 9.40 6.30 137.00
7.500 7.10 5.20 137.00
12.100 11.20 5.60 136.00
11.200 9.20 4.70 140.00
7.100 6.80 4.60 137.00
8.100 6.10 68.00 144.00
6.400 7.30 4.20 138.00
7.500 7.90 5.30 144.00
5.800 7.51 4.10 143.00
5.800 7.51 4.10 143.00
8.200 6.60 4.90 144.00
8.000 7.50 5.60 138.00
5.800 6.20 5.70 136.00
8.600 6.90 12.10 139.00
17.300 12.50 4.60 136.00
3.800 5.80 6.60 139.00
")
data = read.table(textConnection(Input),header=TRUE)
```

```
###################################Bootstrap#####################################

#/Performing Bootstrap for 3000
mydata <- rbind.data.frame(data, stringsAsFactors = FALSE)
iboot <- sample(1:nrow(mydata),size=3000, replace = TRUE)
bootdata <- mydata[iboot,]

################################### MLP#########################################

#/Install the Neuralnet Package/
if(!require(neuralnet)){install.packages("neuralnet")}
library("neuralnet")

#/Checking for the Missing Values/.
apply(data, 2, function(x) sum(is.na(x)))
#/Scaling the data for Normalization
# Method (usually Called Feature Scaling) to get all The Scaled Data
# in the Range [0,1]/
max_data <- apply(data, 2, max)
min_data <- apply(data, 2, min)
data_scaled <- scale(data,center = min_data, scale = max_data - min_data)

#/Randomly Split the Data Into 70:30
#70 Percent of the Data at Our Disposal to Train The Network
#30 Percent to Test the Network/

index = sample(1:nrow(data),round(0.70*nrow(data)))
train_data <- as.data.frame(data_scaled[index,])
test_data <- as.data.frame(data_scaled[-index,])

#Print Data
print(train_data)
print(test_data )

#/Build the Network
#There 3 Hidden Layers Have 3 and 2 Neurons Respectfully
#Input Layer = 2
#Output Layer = 1/
nn <- neuralnet(Urea ~FBS+HbA1c+Sod, data=train_data, hidden=c(2,2),
linear.output = F, stepmax = 1000000)
plot(nn)
options(warn=-1)

#/30 Percent of The Available Data to Do This:
#Using Only the First 2 Columns Representing The Input Variables
#of The Network and 1 is The Output For NN/
predicted <- compute(nn,test_data[,1:3])

#/Use the Mean Squared Error NN (MSE-forecasts the Network) as a Measure of How Far
#Away Our Predictions Are from The Real Data/
MSE.net <- sum((test_data$Urea - predicted$net.result)^2)/nrow(test_data)

# Print the value of MSE-forecasts the Network
MSE.net
```

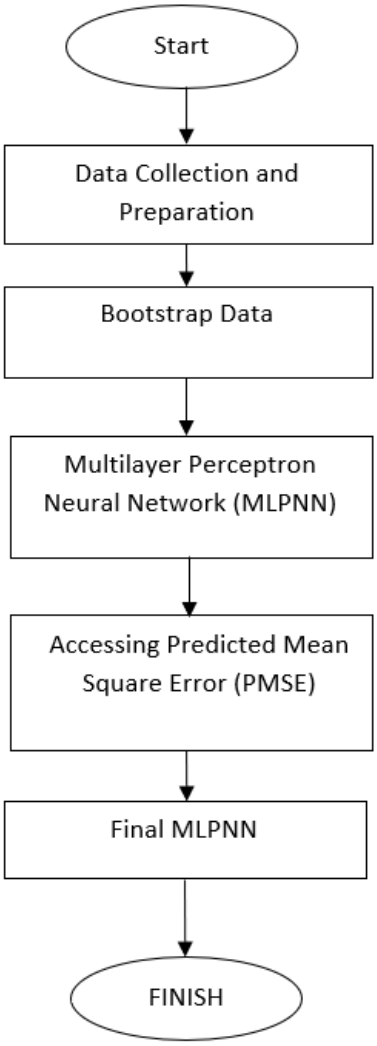

Fig.2: Flow chart of multilayer perceptron neural network (MLPNN)

3. Results and discussion

3.1 The result from multilayer perceptron neural network (using R syntax)

There are three variables in this study that are being focused which Fasting Blood Glucose, HbA1c, and Sodium. The value of mean square error (MSE) is being collected for every factor combination. The combination of these three variables gives the smallest value which is 0.00641612. That means, these three factors contributing significantly toward the urea level.The architecture of the MLP with one hidden layer, three input nodes, two hidden nodes, and one output node, as shown in Figure 3.

Table 2: Possible combination of input variable into MLP model

Input Variables	MEAN SQUARE ERROR
X1,X2	0.01030575
X1,X3	0.08028795
X2,X3	0.08229567
X1,X2,X3	0.00641612

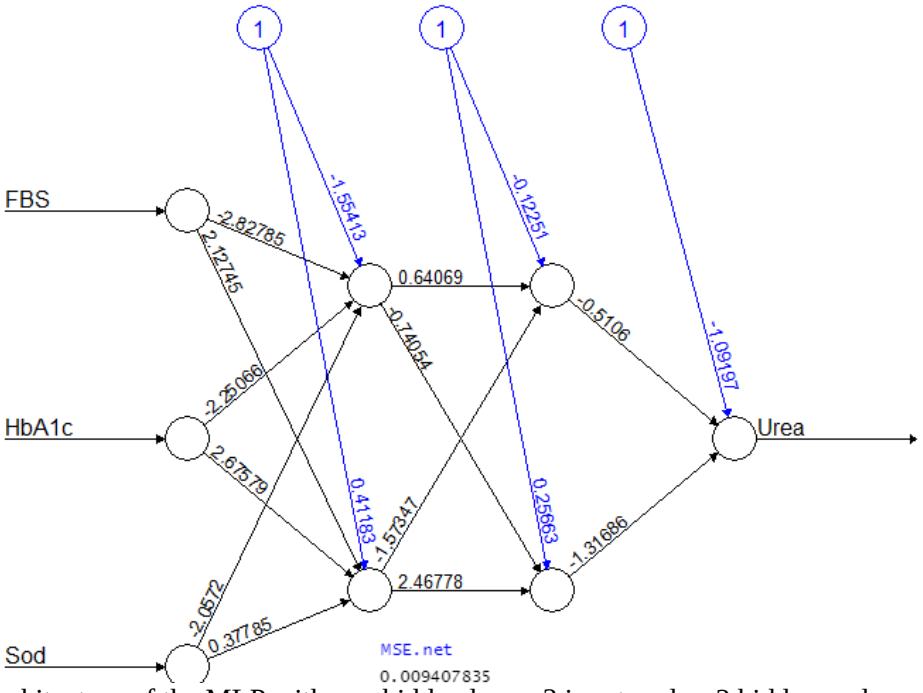

Fig. 3: The architecture of the MLP with one hidden layer, 3 input nodes, 2 hidden nodes, and one output node

4. Summary and conclusion

First, this paper is emphasizing to the development of MLP methodology and its application to the medical sciences field. Second, this research paper examines and determines the factors that influence the level of urea reading among diabetics. The methodology used in this study consists of two main components namely bootstrap methodology and MLP methodology. Ideas based on methodologies are coordinated with the R syntax algorithm. Through the MLP analysis, it was found that Fasting Blood Glucose, HbA1c, and Sodium Reading significantly contributing to the area level.

From this analysis, the PMSE was found to be the lowest value when all these three factors were combined. (Table 2). The smallest value of PMSE, the better result achieved. Therefore, for the modeling purpose, these three variables can be used together as input for the modeling procedure. This technique had led to successful research and give the best results for decision making, especially for the decision-maker.

Serum urea monitoring can be used to properly predict and help to manage the patient properly before any complication such as to the lower limb changes occurs [1]. The management needs a proper assessment to prevent a waste of time and money for both patient and clinical staff. Properly managing the complication such as diabetic nephropathy can be assessed in the diabetic patient using serum urea [2]. A certain major decision such as major foot amputation needs serum urea assessment as well as other serum levels such as fasting blood sugar [3]. Assessing the diabetic nephropathy in diabetes is achieved via the correlation of certain parameters such as HbA1c and the urea level [4]. In contrast however, in certain cases, serum urea is non-correlated with serum sodium such as in chronic renal failures [5].

5. Acknowledgments

The authors would like to express their gratitude to Universiti Sains Malaysia (USM) for providing the research funding (No.304/PPSG/6315410, School of Dental Sciences, Health Campus, UniversitiSains Malaysia, Kelantan, Malaysia).

6. References

[1] Atosona, A.and Christopher Larbie. Prevalence and Determinants of Diabetic Foot Ulcers and Lower Extremity Amputations in Three Selected Tertiary Hospitals in Ghana. Journal of Diabetes Research, vol. 2019, Article ID 7132861, 9 pages, 2019. https://doi.org/10.1155/2019/7132861

[2] H. A. E. Kattawy and W. M. R. Ashour. Anti-Asprosin: A Potential Protective Role Against the Progression of Diabetic Nephropathy in Type 2 Diabetic Rats. American Journal of Biomedical Sciences, 2019; 11(3): pp. 183–199, Jul. 2019.

[3] Mahdi, M.M. and I.O. Yasir, Risk Factors for Major Amputation in Diabetic Foot Patients. University of Thi-Qar Journal of Medicine, 2019; 17(1): p. 40-53.

[4] Vaishnavi, M.S., et al., Study of Correlation between HbA1c and Renal Dysfunction in Type-2 Diabetes Mellitus. Indian Journal of Public Health Research & Development, 2019; 10(6): p. 194-199.

[5] Aruna, G., Evaluation of minerals, urea, and creatinine in chronic renal failure. International Journal of Clinical Biochemistry and Research, 2019; 6(4): p. 447-451.

[6] Norizan M., Maizah H. A., Suhartono, and W. M. A. W. Ahmad. Forecasting Short Term Load Demand using Multilayer Feed-forward (MLFF) Neural Network Model. Applied Mathematical Sciences. 2012; Vol. 6, No. 108: 5359-5368.

[7] Norizan M., Nor Azlida A., W. M. A. W. Ahmad, and Maizah H. A. Multilayer Feed-Forward Neural Network Approach to Lymphoma Cancer Data. Int. J. Contemp. Math. Sciences. 2012; Vol. 7, No. 35: 1749-1756.

[8] Norizan, M., W. M. A. W. Ahmad, Nor Azlida A., and Maiza H. A. Modeling Multi-Layer Feed-forward Neural Network Model on the Influence of Hypertension and Diabetes Mellitus on Family History of Heart Attack in Male Patients. Applied Mathematical Sciences. 2013; Vol. 7, No. 41: 2047-2053.

[9] Efron, Bradley and Robert J. Tibshirani. An Introduction to the Bootstrap. New York, NY: Chapman and Hall. 1993.