ANN-Based Prediction of Kidney Dysfunction Using Clinical Laboratory Data

Ali Hussein Ali Al-Timemy
Dept. of Biomedical Engineering, Al-Khawarizimi College of Eng., Baghdad University, Baghdad, Iraq.
Email: alialtimemy2006@yahoo.com

Iyden Kamel
Dept. of Biomedical Engineering, Al-Khawarizimi College of Eng., Baghdad University, Baghdad, Iraq.
Email: idenel_1969@yahoo.com

Hussam K. Abdul Ameer
Dept. of Biomedical Engineering, Al-Khawarizimi College of Eng., Baghdad University, Baghdad, Iraq.

Abstract

This paper presents the prediction of Kidney dysfunction using probabilistic neural network (PNN). Six hundred and sixty (660) sets of analytical laboratory test have been collected from one of the private Clinical laboratories in Baghdad. For each subject, Serum urea and Serum creatinin levels have been analyzed and tested by using clinical laboratory measurements. The collected Urea and cretinine levels are then used as inputs to the Artificial Neural network model in which the training process is done by PNN which is a class of radial basis function (RBF) network is used as a classifier to predict whether Kidney is normal or it will have a dysfunction. The accuracy of Prediction, sensitivity and Specificity were found to be equal to 99%, 98% and 99% respectively for this proposed network .We conclude that the proposed model gives faster and more accurate prediction of Kidney dysfunction and it works as promising tool for predicting of routine kidney dysfunction from the clinical laboratory data.

1.introduction

Renal failure is a serious medical condition affecting the kidneys. When a person suffers from renal failure, their kidneys are not functioning properly or no longer work at all. Renal failure can be a progressive disease or a temporary one depending on the cause and available treatment options [1].

The kidneys are glands that are located in the abdominal region just above the pelvis on either side of the body. When functioning normally, the kidneys separate and filter excess water and waste from the blood stream. The kidneys are responsible for producing urine, which is used to flush away the toxins. The kidneys also maintain a healthy balance of fluids and electrolytes, or salt compounds, in the body.

In renal failure the kidneys undergo cellular death and are unable to filter wastes, produce urine and maintain fluid balances. This dysfunction causes a build up of toxins in the body which can affect the blood, brain and heart, as well as other complications. Renal failure is very serious and even deadly if left untreated. There are two types of renal failure: acute and chronic. Acute renal failure occurs suddenly and is usually initiated by underlying causes, for example dehydration, infection, serious injury to the kidney or the chronic use of over the counter pain medications like Tylenol (acetaminophen) or Advil (ibuprofen). Acute renal failure is often reversible with no lasting damage.

Chronic renal failure is more serious than acute renal failure because symptoms may not appear until the kidneys are extremely damaged. Chronic renal failure can be caused by other long term diseases, such as diabetes and high blood pressure. Chronic renal failure can worsen over time, especially when the problem has gone undiagnosed and treatment is delayed [2].

Recent changes in health care have motivated attempts to improve measures of illness severity and predict outcomes for several diseases like kidney disease. Adjustments for illness severity may have an important role in evaluating quality of care. Computerized scoring systems may be useful if they have a high prognostic accuracy.

The neural networks derive their power due to their massively parallel structure, and an ability to learn from experience. They can be used for fairly accurate classification of input data into categories, provided they are previously trained to do so. The accuracy of the classification depends on the efficiency of training. The

knowledge gained by the learning experience is stored in the form of connection weights, which are used to make decisions on fresh input[3].

One computer technique under investigation is the artificial neural network [4]. Neural networks are tools for multivariate analysis that can be used to estimate disease risk. They are able to model complex nonlinear systems with significant variable interactions. Theoretical work suggests that neural networks may be able to consistently match or exceed the performance of traditional statistical methods [5]. Neural networks have been used effectively in several clinical studies, in areas including the evaluation of radiological studies [6], the diagnosis of acute illness [7], and the prediction of intensive- care-unit length of stay [8].

The purpose of this study is to develop a probabilistic neural network as predictor for the kidney dysfunction using a number of different admission laboratory and clinical variables.

2. Probabilistic neural network theory

PNN which is a class of radial basis function (RBF) network is useful for automatic pattern recognition, nonlinear

mapping and estimation of probabilities of class membership and likelihood ratios . It is a direct [9] continuation of the work on Bayes classifiers in [10] which it is interpreted as a function that approximates the probability density of the underlying example distribution. The PNN consists of nodes with four layers namely input, pattern, summation and output layers as shown in Fig. 1. The input layer consists of merely distribution units that give similar values to the entire pattern layer.

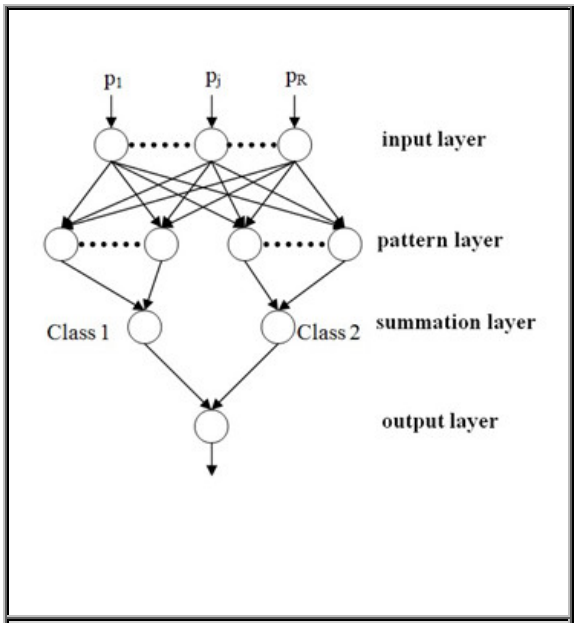

Figure (1) PNN Architecture

For this work, RBF is used as the activation function in the pattern layer. Figure 2 shows the pattern layer of the PNN.

The ||dist|| box shown in Fig. 2 subtracts the input weights, IW1,1 , from the input vector, p, and sums the squares of the differences to find the Euclidean distance. The differences indicate how close the input is to the vectors of the training set. These elements are multiplied element by element, with the bias, b, using the dot product (.*) function and sent to the radial basis transfer function. The output a is given as,

$$a = radbas(\| IW_{1,1} - p \| b) \quad (1)$$

where radbas is the radial basis activation function which can be written in general form as,

$$radbas(n) = e^{n^2} \quad (2)$$

The training algorithm used to train the RBF is the orthogonal least squares method which provides a systematic approach to the selection of RBF centers [11]. The summation layer shown in Fig.1 simply sums the inputs from the pattern layer which correspond to the category from which the training patterns are selected as either class 1 or class 2.

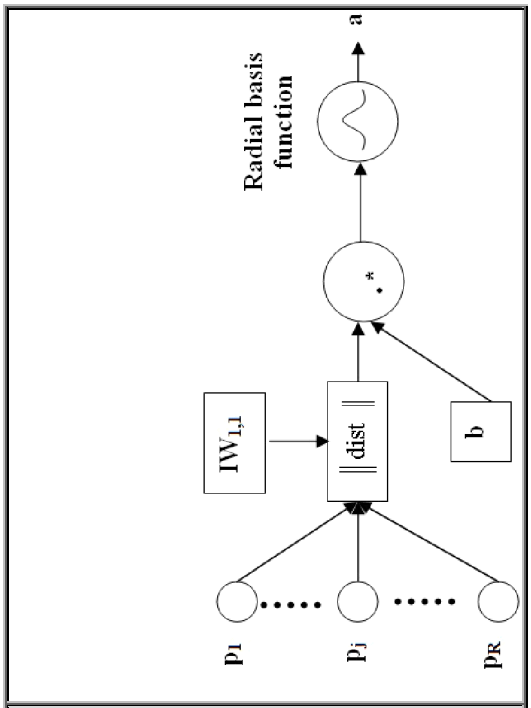

Figure (2) PNN pattern layer

Finally, the output layer of the PNN is a binary neuron that produces the classification decision. As for this work, the classification is either class 1 for stable cases or class 2 for unstable cases.
Performance of the developed PNN can be gauged by calculating the error of the actual and desired test data. Firstly, error is defined as,

Error, E_n =|Desired Output, DO_n- Actual Output, AO_n|	3

where, n is the test data number. The desired output is the known output data used for testing the neural network. Meanwhile, the actual output (AO) is the output obtained from testing on the trained network.
From equation (3), the percentage mean error, ME (%), can be obtained as:

$\text{Percentage Mean Error, ME}(\%) = \sum_{n=1}^{N} \frac{E_n}{N} \times 100$	4

where N is the total number of test data.

3. Testing data and tools used

In this work, data were collected from one of the private hospitals in Baghdad. Urea and Creatinine levels for six hundred and sixty three subjects have been analyzed by clinical laboratory methods. The total amount of cases for all subjects have been divided into two groups, one for training (602 cases) and the other for testing of the algorithm (61 cases).
MATLAB software package version 7 is used to implement the software for the current work. A sample of the testing data for twenty five cases is shown in Table .1. The Urea and Creatinine levels are used as an input to the PNN classifier. Then the PNN will predict whether the kidney of normal (output of the PNN is 1) or Abnormal state (the output of the PNN is 2).

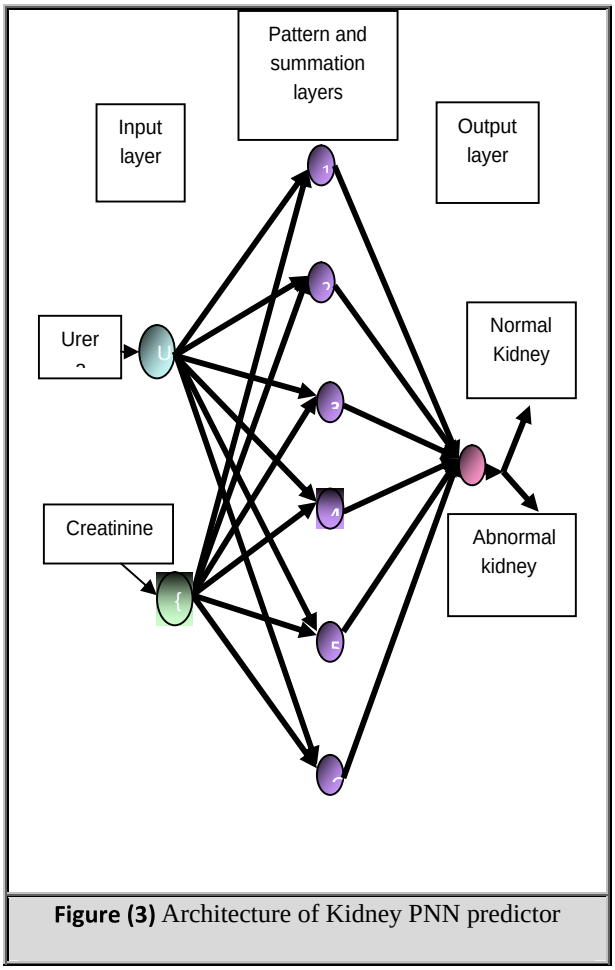

Figure (3) Architecture of Kidney PNN predictor

The neural networks derive their power due to their massively parallel structure, and an ability to learn from experience. They can be used for fairly accurate classification of input data into categories, provided they are previously trained

to do so. The accuracy of the classification depends on the efficiency of training. The knowledge gained by the learning experience is stored in the form of connection weights, which are used to make decisions on fresh input. The general architecture for the proposed system of prediction of kidney dysfunction is shown in fig. 3.

Table (1) Sample of the testing data

No	Urea	Cretin.	Diagnosis	Output of PNN
1	35	0.8	Normal	1
2	34	0.8	Normal	1
3	57	1.4	Abnormal	2
4	65	1.4	Abnormal	2
5	38	0.9	Normal	1
6	34	0.8	Normal	1
7	53	1.3	Abnormal	2
8	38	0.9	Abnormal	2
9	185	4.6	Abnormal	2
10	43	1.1	Normal	1
11	36	0.8	Normal	1
12	48	1.1	Normal	1
13	48	1.2	Normal	1
14	47	1.1	Normal	1
15	142	3.7	Abnormal	2
16	27	0.8	Abnormal	2
17	32	0.8	Normal	1
18	39	0.9	Abnormal	2
19	39	0.9	Normal	1
20	36	0.9	Normal	1
21	50	1.2	Normal	1
22	39	0.8	Normal	1
23	38	0.9	Abnormal	2
24	39	0.9	Normal	1
25	45	1.1	Normal	1
26	39	0.9	Normal	1
27	50	1.2	Normal	1
28	44	1.1	Normal	1
29	32	0.8	Normal	1
30	72	1.9	Abnormal	2
31	39	0.9	Normal	1
32	32	0.8	Normal	1
33	46	1.2	Normal	1
34	38	0.9	Normal	1
35	147	3.9	Abnormal	2
36	39	0.9	Normal	1
37	50	1.2	Normal	1

4. Results and discussion

The performance of the algorithm was evaluated by computing the percentages of Sensitivity (SE), Specificity (SP) and Accuracy of Prediction (AP), the respective definitions are as follows [12]:

Sensitivity: is the fraction of real events that are correctly detected among all real events.

[SE = 100xTP/ (TP+FN)]

Specificity: is the fraction of nonevents that has been correctly rejected.

[SP = 100xTN/ (TN+FP)]

Accuracy of Prediction: is the prediction rate.

[CP=100x (TP+TN)/ (TN+TP+FN+FP)]

where TP was the number of true positives, TN was the number of true negatives, FN was the number of false negatives, and FP was the number of false positives.
Since it is interesting to estimate the performance of predictors based on the prediction of normal and abnormal kidney, the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are defined appropriately as shown below:
FP: Predicts normal as abnormal.
TP: Predicts abnormal as abnormal.
FN: Predicts abnormal as normal.
TN: Predicts normal as normal.

In our study, the unique neuron of the output layer corresponds to the normal and abnormal kidney. In practice, the number of neurons in the hidden layer varies according to the specific recognition task and is determined by the complexity and amount of training data available. If too many neurons are used in the hidden layer, the network will tend to memorize the data instead of discovering the features. This

will result in failing to classify new input data. The resulted accuracy and sensitivity for PNN are shown in Table 2.

Table (2) The results after training of the network

	No. of cases	Sensitivity	Specificity	Accuracy of Prediction
PNN	**61**	**98%**	**99%**	**99%**

5. Conclusions

The use of PNN has been proposed for prediction of kidney dysfunction by means of classifying the kidney into either normal or abnormal kidney. Urea and Creatinine levels were first measured in the clinical laboratory. These data were carried out to generate training data for the PNN and to predict the kidney failure. The accuracy, sensitivity and Specificity were calculated to evaluate its effectiveness. We conclude that that the proposed model gives faster and more accurate prediction of Kidney dysfunction and it works as promising neural network technique for predicting of routine kidney dysfunction from the clinical laboratory data.

6. References

1. S, Miller S. "Acute oliguria". N Engl J Med338(10),671–5, 1998.

2. Meyer TW and Hostetter, TH "Uremia". N Engl J Med 357(13):1316, 2007.

3. Ali Hussain Ali Al-Timemy, F. M Al-Naima and Safa S. Mahdi, "Data Acquisition System for Myocardical Infarction Classification Based on Wavelets and Neural Networks", Porceedings of the Fifth International Multi-Conference on Systems, Signals and Devices (IEEE SSD'08), Jordan, 2008.

4. Chester M., " Neural networks: a tutorial", Englewood Cliffs, NJ: Prentice and Hall, 1993.

5. Hornik, K. " Multilayer feed forward networks are universal approximators Neural Networks", 1989;2:359-366.

6. Scott JA and Palmer EL, " Neural network analysis of ventilation-perfusion lung scans", Radiology;186:661-664, 1993.

7. Baxt WG, " Use of an artificial neural network for the diagnosis of myocardial infarction", ANN Intern. Med.;115:843-848., 1991.

8. Tu JV and Guerrier MRJ, "Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery", Computer Biomed Res;26:220-229, 1993.

9. Specht, D.F., "Enhancements to Probabilistic Neural Networks", International Joint Conference on Neural Networks, (1): 525 – 532, 1992.

10. Burrascano, P., "Learning Vector Quantization For The Probabilistic Neural Network", IEEE Transactions on Neural Networks, 2(4): 458-461, 1991.

11. Chen, S., C.F.N. Cowan, P.M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks", IEEE Transactions on Neural Networks, 2(2): 302-309, 1991.

12. N. Belgacem, M. A. Chikh, A. Chikh, F. B. Reguig , Applications of Neural Nets to Detect Atrial Premature Beat, Laboratory Report of Biomedical Engineering, Department of Electronics, Faculty of Science and Engineering, University Abou Bekr Belkaïd, Tlemcen, Algeria, 2002.

الذكاء الاصطناعي للتنبؤ بالفشل الكلوي من خلال تحاليل المختبر السريرية

علي حسين علي التميمي، د. ايدن كامل و د.حسام كاظم عبد الامير
قسم هندسة الطب الحياتي-كلية هندسة الخوارزمي-جامعة بغداد

الخلاصة

يقدم البحث التنبؤ بفشل الكلية بأستعمال الشبكة العصيبية الاحتمالية. تم جمع و قياس كمية اليوريا والكرياتنين في الدم لعدد كبير من المرضى والاصحاء في احد المختبرات الاهلية في بغداد . استعملت هذه الفحوصات كمدخلات للشبكة العصيبية الاحتمالية. لاختبار كفاءة الخوارزمية، تم حساب الدقة للشبكة ووجد انها ٩٩% ومن هذا البحث نستنتج ان النظام المقترح يعطي تنبؤ سريع ودقيق للفشل الكلوي.