International Journal of Engineering & Technology, 7 (4) (2018) 3367-3371

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.13686
Research paper

Diagnosis urine disease based on KNN algorithm and ANN

Maytham Salman Azeez [1] *, Zinah Muneer Maki [1]

[1] Univ. Al- Kafeel College, province of Najaf, Rep .of Iraq
**Corresponding author E-mail: Maytham.aljebory@alkafeeluc.edu.iq*

Abstract

The Artificial Neural Networks (ANN) are commonly applied in several medical fields for undertaking diagnosis of diseases. ANN can be used for diagnosing the urine bladder as well as nephritis inflammation. This research paper mainly focuses on undertaking the diagnosis of urine disease on the basis of K-Nearest Neighbor Algorithm (KNN) and Artificial Neural Network (ANN). The Acute Inflammation Data Set was employed in the research methodology. The data was collected from the UCI Machine Learning Respiratory which would enhance the successful carrying out of the diagnosis. The collected data is distinguished into inputs as well as targets. The systems will represent the inputs to a neural network. The neural network targets will be recognized as 1's for infected and as 0's for non-infected. It is evident from the results that the artificial neural network could be significant for recognizing the infected person. The results which can be obtained from the application of ANN methodology on the basis of the selected signs and symptoms clearly indicates the network ability to comprehend the specific patterns which correspond to the person's Symptoms-Nearest Neighbor algorithm normally unveils the satisfactory rate at which the diagnosis is done to ascertain the distinction between the infected as well as non-infected urinary system.

Keywords: *Artificial Neural Networks; K-Nearest Neighbor Algorithms; Urinary System Disease Diagnosis; UCI Machine Learning Respiratory*.

1. Introduction

The neural networks are more advantageous than conventional programming especially on the basis of providing the viable remedies to the problems that may have difficulties in getting the algorithm solutions. Neural networks are identified to be very crucial for handling different problems while allowing people to carry out the image analysis, prediction, pattern recognition, interpretation as well as the clinical diagnosis [1].

Machine learning involves scientific discipline that is mainly concerned with stage development, analysis as well as automated methods implementation which enable the machine to undergo certain evolution through a process of learning and therefore, fulfilling different tasks which may not be easy with other convectional algorithm measures. The learning is the major factor for determining different categories of learning algorithms. Such methods include K-Nearest Neighbor, Support Vector Machine and Neuron Network.

Many researchers have carried out the research study on the development of expert systems which are necessary for solving more complex problems associated with the medical diagnosis such as cancer as well as urinary system diagnosis [2]. The K-Nearest Machine has been commonly used for carrying out the diagnosis in the medical field. The study conducted by Ozyilmaz et al., [3] verifies three major neural networks strategies. The k-Nearest Neighbor algorithm has been adequately analyzed and its performance evaluated in the context of carrying out the diagnosis in the urinary systems. According to the study carried out by Heckerling et al., [4],it can been verified that Artificial Neural Networks (ANN) have been recognized to be helpful in the carrying out the predictions on the urinary tracts in order to determine the infection. The urodynamic stress has been diagnosed with the computer-aided device [5]. The application of the artificial neural networks has adequately investigated by Monadjemi and Moallen [6] for typical diagnosis of the disease. The possible infection which can be diagnosed is interstitial cystitis. This diagnosis is done by the specialists who have adequate knowledge on the relevant medical fields. The experimental results as well as advantages of utilizing the fuzzy methodology were well discussed. The review presented by Lisboa, [7] precisely verifies the importance of using the artificial neural networks (ANN) for conducting the diagnosis clinical functions, survival analysis as well as prognosis. The concepts of utilizing the artificial neural networks have become useful for diagnostic advancements in the medical fields.

The prediction and diagnosis of the urinary systems have been imposed great challenge. In this paper, a new technique of undertaking the diagnosis of urinary diseases such as interstitial cystitis has been selected in the order provided the effective solutions to challenges currently experienced when undertaking the diagnosis. K-Nearest Neighbor can greatly help in enhancing the administration of the urinary diagnosis. The possible urinary infections can further be recognized with Artificial Neuron Networks (ANN). This is an important technique, which can be used to make the viable patterns as well as predictions on the urinary diseases.

Recently there are lots of researches with in the field of diagnostic urinary as in [8] and [9] in which the problems concerning a heterogeneity of tumor and anticipatory diagnosis need a deep intellect. Therefore, challenges remain not changed.

2. Artificial neural networks

The artificial neural network (ANN) involves a computational model, which tries to provide the inherent parallel of the human brain. It comprises several elements that are usually facilitated by biological systems. The interconnections of the elements majorly determine the function of the network. Normally, the first layer and the final layer represent the input and the output respectively. The neural network is well illustrated by the figure given below;

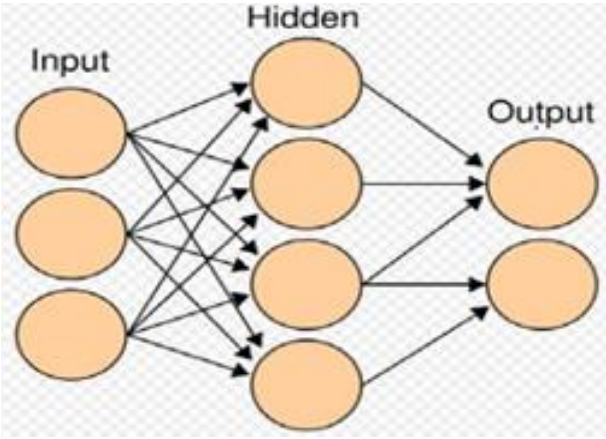

Fig. 1: Typical Neural Network [10].

The medical research carried out with the artificial neural network has been widely used and become prominent in the biomedical systems. In this case, the solution is not constrained in the linear sense.
The viable patterns and prediction of the urinary infections can be enhanced through the utilization of the Artificial Neuron Networks (ANN). This technique can easily convert different inputs into more helpful outputs by applying the following equation.

$$O1 = 1/ (1+expo (-F)) \quad (1)$$

Where:

F= W1 x X1 + W2 X2 + W3 X3.

Alternatively, the simple programming can be utilized to substitute the artificial neuron network but Artificial Neuron Networks technique has been selected instead of simple programming due to the fact that it utilizes the training sets to generate the effective decision-making process. The effective decision-making process is relatively accurate in comparison to simple programming algorithms. The real predictions can be made with the ANN. Artificial neuron networks contain inputs as well as outputs. The inputs of artificial neuron networks are usually represented by complex mathematical analysis while outputs are the useful results of the predictions [11].

3. K-nearest neighbor (KNN)

It involves the popular as well as simplest algorithms [12]. In this case, the generation cannot be done through the use of training data points. The interconnected elements can be classified on the basis of the theory of distance which can be calculated between the two adjacent elements. The KNN algorithm is usually achieved through careful selection of the key parameters which include distance function as well as parameter K. The KNN algorithm is well illustrated in figure 2.

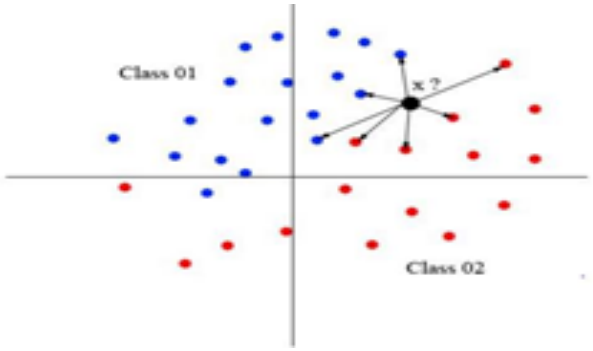

Fig. 2: K-NN Algorithm [13].

KNN algorithms allow for ease predictions of the urinary diseases such as interstitial cystitis. Upon making the predictions on the basis of KNN, the k-nearest outcomes can be achieved. This is mathematically illustrated by the equation given below.

$$y = \frac{1}{K}\sum_{i=1}^{x} y_i \quad (2)$$

Where
y_i = ith case and y= prediction outcomes.
The two techniques play critical roles in making the viable predictions concerning the diagnosis of urinary systems infections. They have a link to each other in that ANN could be utilized for making the anticipations on the medical field in order to provide better performance to given k-Nearest Neighbour algorithms. This clearly indicates that both can work in conjunction with each other.

4. The proposed diagnosis model

4.1. Diagnosis model

The diagnosis model entails three major parts which include; Image segmentation, extract features, and matching results. These parts are well illustrated in the flowchart given below in figure 3.

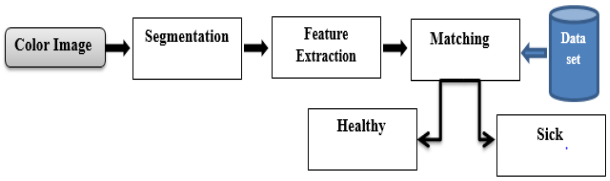

Fig. 3: Diagnosis Model.

The steps followed for the research;

i) The image segmentation was carried out
ii) The opened image was converted into binary form.
iii) The segmentation conversion of the binary image.
iv) For the case of the feature extraction, the feature extraction for LBP was used.

The flowchart methodology is illustrated in figure 4.

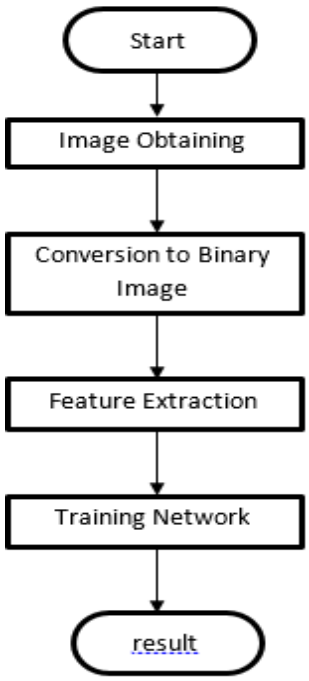

Fig. 4: The Methodology.

Segmentation was done on the image in order to provide the effective distinctions between the infected urinary systems and uninfected urinary systems. The results obtained from image segmen-

tation would enhance good analysis concerning the health conditions of the patient. During the analysis, image segmentation was done through making label allocations in terms of attribute possessions. Consequently, allowing easy conclusion.
Feature extractions are very crucial for checking the results of the data obtained from the analysis and differentiate them from the standard values. During the feature extraction, the data acquired from the image created is transformed and make labels into a certain set of features, which would facilitate further classifications. The common kind of features obtained from the research study was the color feature which would help differentiate between the standard colors with the observed color.
Matching technique is a very important tool which enables the possibility of occurrence of bias in the analyzed data during the research study. In other words, the technique reduces the potential risks that could have been encountered during analysis of data. Matching was done by the certified experience medical technologies in the medical labs. It was done electronically with the effective database. Therefore, the results obtained are relatively accurate.

4.2. Data analysis

The data used for carrying out the diagnosis of Urinary disease included images, symptoms as well as signals. UCI Machine Learning Repository was used in this research study for obtaining the data set.
The medical experts created data set in order to carry out the diagnosis of urinary disease. The data set was created through the experimental analysis. It has been adequately depicted by the statistical table. This diagnosis aims at interstitial cystitis. The concept of the created data set was to construct the model of Artificial Neural Networks as well as to apply K-Nearest Neighbour which were really required during the research study for performing the diagnosis of the urinary system. It is a standard database required for urinary diseases.
The Sigor Mission Hospital was selected to be used for conducting the research study where various pictures of different samples from the selected hospital were taken. The diagnosis was carried out with 10 samples representing the infected people as well as the healthy ones. The drones with special mobile cameras of Samsung type were used for taking the photographs during the research study.

5. Results and discussions

The results from the research study indicate that the data set used comprised 10 images where [5] images taken represented healthy persons while the other 5 images taken were for sick persons. The results achieved by KNN algorithm are presented in the table 1.

Table 1: KNN Results

Sample Number	Expected Answer	KNN Answer
1.	Healthy	Healthy
2.	Healthy	Healthy
3.	Healthy	Healthy
4.	Healthy	Healthy
5.	Healthy	Healthy
6.	Sick	Sick
7.	Sick	Sick
8	Sick	Sick
9.	Sick	Sick
10.	Sick	Sick

From the results achieved by the KNN algorithm, it can be observed that the KNN algorithm is proven 100% accuracy. The previous approach has been presenting great challenges.
The systems used to classify the sick /Healthy persons with images are verified in this section.

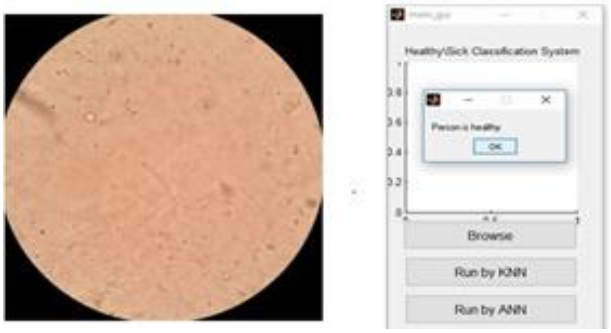

Fig. 5: Group (1), Image Number 1; Classify Healthy.

Fig. 6: Group (1), Image Number 2; Classify Healthy.

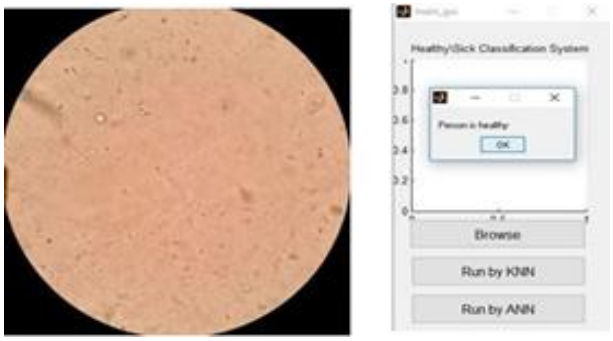

Fig. 7: Group (1), Image Number 3; Classify Healthy.

Fig. 8: Group (1), Image Number 4; Classify Healthy.

The group (1) of the images in the above figures 5, 6, 7, and 8 showed the purposed system that used KNN algorithm in diagnoses of the samples taken from normal persons (without UTI disease). While Group (2) of images has been shown in the figures 9, 10, 11, and 12 is for the patients suffering from UTI; The proposed system diagnoses in these cases using KNN algorithm. These samples taken from patient visiting The Sigor Mission Hospital was selected to be used for conducting the research study.

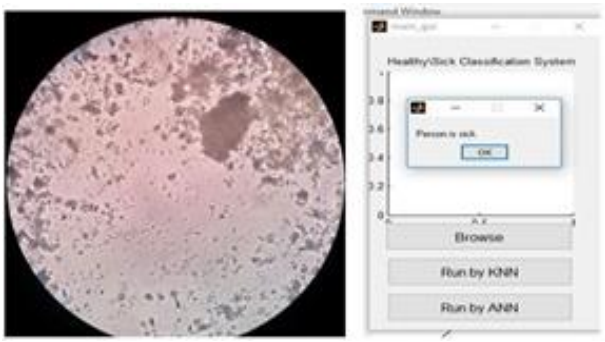

Fig. 9: Group (2), Image Number 5; Classify Sick.

Fig. 10: Group (2), Image Number 6; Classify Sick.

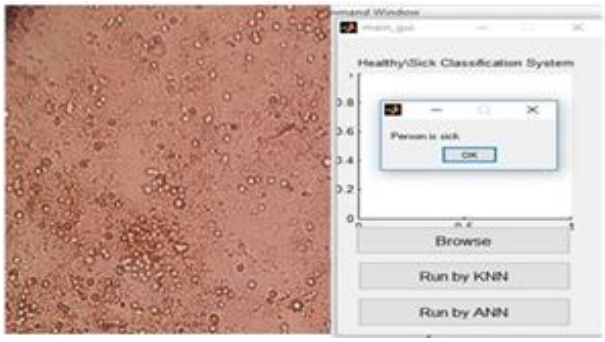

Fig. 11: Group (2), Image Number 7; Classify Sick.

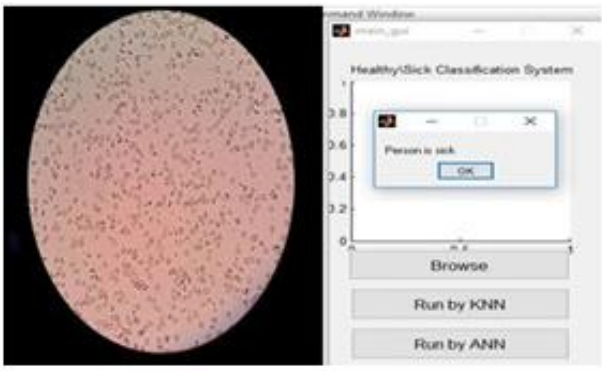

Fig. 12: Group (2), Image Number 8; Classify Sick.

In addition, performance comparison of ANN and KNN results with FFT & PSD techniques were obtained from the research, which are well illustrated, in table 2.

The dataset created has clearly stipulated the classification of healthy and unhealthy persons. The application of the artificial neural networks in the research study in distinguishing the healthy persons and unhealthy persons based on the selected symptoms clearly indicated the network ability to depict the patterns which correspond to the symptoms of the person. The network was integrated into the testing analysis. With this network, 99% of the cases were classified in the testing set. The images obtained during the research study clearly depict the distinction between the healthy and unhealthy persons.

Table 2: Performance Comparison Data for ANN and KNN.

K and N	KNN		ANN	
	Test success with FFT (%)	Test success with PSD (%)	Test success with FFT (%)	Test success with PSD (%)
1	76.2	85.7	69.2	47.6
2	76.2	85.7	42.9	57.1
3	57.14	57.14	71.4	71.4
4	61.9	61.9	71.4	47.6
5	52.38	57.14	61.9	90.5
6	42.86	61.9	76.5	71.4
7	28.57	57.14	82.1	81
8	42.86	57.14	57.1	81
9	47.68	52.38	38.1	85.7
10	38.1	52.38	76.3	85.7
11	38.1	52.38	61.9	47.6
12	42.86	47.62	49.2	81
13	52.38	47.62	42.9	76.2
14	42.86	42.86	42.9	66.7
15	42.86	38.1	85.7	76.2
16	42.86	33.33	71.4	71.4
17	42.86	38.1	71.4	66.7
18	47.62	33.33	57.1	47.6
19	52.38	33.33	71.4	61.9
20	52.38	33.33	42.9	66.7

In this research, the training data set was employed for generating the KNN as well as ANN. The variable K, as well as hidden neuron (N), were used for performance comparison between the KNN and ANN. While coming up with the performance comparison, the FFT, as well as PSD techniques, were utilized in both KNN and ANN. The performance comparison between KNN and ANN was evaluated in terms of the variable K value and Neuron count, respectively. This is best illustrated by the following figures;

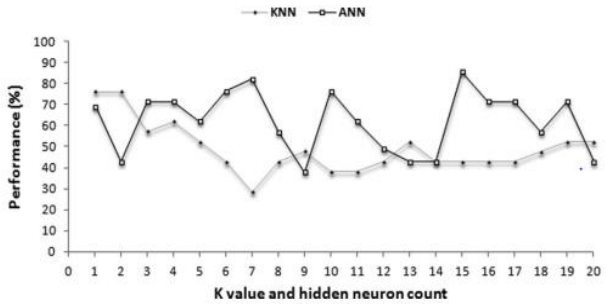

Fig. 13: Performance Comparison between KNN and ANN with FFT Technique.

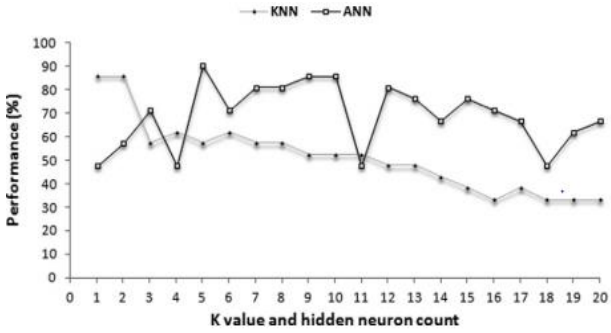

Fig. 14: Performance Comparison between KNN and ANN with PSD Technique.

6. Conclusion

In this research study, the diagnosis of the urinary disease based on KNN algorithm as well as ANN have been adequately verified. This paper clearly indicates the adoption of the new methods of carrying out a diagnosis of the urinary systems would contribute greatly towards improving the administration of medical treatments. Different classifications of images have been made to distinguish between the healthy and the sick person. The proposed diagnosis model has clearly specified the steps required for the conduction of the research. The KNN has clearly classified all the samples selected for the research. In addition, the performance comparison between the KNN and ANN has been adequately described. In this case, the Variable K values, as well as Neuron

Count N, have been utilized in evaluating the performance comparison between KNN and ANN.

References

[1] M.Sordo, "Introduction to Neural Networks in Healthcare," Open Clinical, [Online].

[2] S. a. A.Kusiak, (2002), "Cancer gene search with data mining and genetic algorithms," Computers in Biology and Medicine.

[3] T. L.Ozyilmaz, (2003), "Artificial neral networks for diagnosis of hepatitis disease." in Proceedings of the International Joint Conference on Neural Networks.

[4] C. a. G. Heckerling, (2007)"Prediction of Urinary tract inflection based on the artificial neural networks and genetic algorithms." International Journal of Informatis, pp. 289-296. https://doi.org/10.1016/j.ijmedinf.2006.01.005.

[5] Y. a. H.Chen, (2007)"Computer Aided-Diagnosis of uridynamic stress incontinence with vector -based perineal ultrasound using neural networks," Ultrasound in Obsterrics and Gynechology. pp. 1002-1006.

[6] Moallen, Monadjemi, (2008), "Automatic Diagnosis of Particular Disease using a Fuzzy-Neural Approach.," International Review on Computers & Software, pp. 406-411.

[7] P.J.G.Lisboa, (2002), "A review of evidence of health benefit from artificial neural networks in medical intervention", [Online].

[8] M. H. Kadhem and A. M. Zeki, (2014), "Prediction of Urinary System Disease Diagnosis: A Comparative Study of Three Decision Tree Algorithms," 2014 International Conference on Computer Assisted System in Health, Kuala Lumpur, pp. 58-61. https://doi.org/10.1109/CASH.2014.25.

[9] Jamie J. D'Costaa, Douglas G. Ward, Richard T. Bryan, (2016), "Urinary biomarkers for the diagnosis of urothelial bladder cancer", New Horizons in Translational Medicine, Vol. 3, pp. 221–223.

[10] R. a. J. J.S.Snchez, (2007), "An analysis of how training data-complexity affects the nearest neigbour classified." Patterns Analysis and Applications, p. 3.

[11] J. B. Siddharth Jonathan and K.N. Shruthi, A Two Tier neural inter -Network Based Approach to Medical Diagnosis using K- Nearest Neighbor Classification for Diagnosis Pruning.

[12] R. a. J. J.S.Snchez, (2007), "An analysis of how training data-complexity affects the nearest neigbour classified.," Patterns Analysis and Applications, p. 3.

[13] Saleh Alaliyat (2008) "Video - based Fall Detection in Elderly's Houses " Master of Science in Media Technology Department of Computer Science and Media Technology, Gjøvik University College.