A Neuro-Fuzzy Based Intelligent System For Diagnosis Of Renal Cancer

Nikita, Balwinder Kaur, Dr. Harsh Sadawarti, Dr. Jimmy Singla

Abstract: The renal cancer is very serious, and rapidly growing disease. It is very necessary that the abnormalities in renal or kidney of a patient will be identified so that cancer will not reach to its advanced stage or phase. Therefore, the early detection of renal cancer is very important to save the life and to increase the life span of particular patient. In this study, an intelligent system using adaptive neuro-fuzzy inference system (ANFIS) is proposed to diagnose renal cancer. This proposed system helps to classify the various stages of renal cancer such as no cancer, stage 1, stage 2, stage 3 or stage 4 cancer. Hence, the prime objective of this study is to diagnose renal cancer by using adaptive neuro-fuzzy inference system (ANFIS). There are seven input variables that are given to the system. These are haematuria (blood in urine), red blood cell count, flank pain, tumour size, Von Hippel-Lindau gene, high blood pressure and trichloroethylene exposure. Similarly, the output corresponding to the given inputs is degree of renal cancer. The output can be no cancer, stage 1, stage 2, stage 3 or stage 4 of renal cancer. The MATLAB software is used to implement this adaptive neuro-fuzzy inference system (ANFIS). The developed medical intelligent system for the detection of renal cancer shows the correct and accurate results with accuracy 96%.

1. INTRODUCTION

The renal cancer is also called renal cell carcinoma, renal adenocarcinoma, hypernephroma or kidney cancer. This renal cancer often found in the adults. As the kidney is a vital organ of our body that helps to excrete waste products during the filtration of blood. The renal cancer occurs in the tubules of kidney. These tubules are small tubes in the kidney, which helps to filter the body blood, get rid of waste and helps to make urine. The renal cell cancer will start growing rapidly and uncontrollably and can spread to other organs of the body like lungs. The renal cancer hardly shows the sign of it in initial phase. In addition, there is no any test till now that will detect the renal cancer without its symptoms. As the symptoms become more worse, the stage or phase of renal or kidney cancer will also increased. It is very difficult to diagnose renal or kidney cancer at its worse stage or stage 4. There are no any exact risk factors for the renal cancer by medical experts. In this study, the seven risk factors are considered for the diagnose of renal cancer. These are as follow:

- Haematuria
- Red blood cell count
- Flank pain
- Tumor size
- Von Hippel-Lindau syndrome
- High blood pressure
- Trichloroethylene exposure

Haematuria is the presence or appearance of red blood cells (RBC) in the urine. It can be visible or microscopic haematuria. This symptom is normal if the count of RBC per high-power field is less than 5. If it is 5 or more than 5 then the Haematuria is chronic. The red blood cell count is also called erythrocyte count. It is the blood test in which the number of red blood cell present in the blood is calculated.

The count should not too low or too high. It must be in normal range. The RBC count is normal if the range is in between 3.5 – 5.50 MIL/UL. Flank pain is the pain in the one side of the body between the abdomen and back. If the flank pain is normal then there is no any issue. However, if the pain is continues and severe then it will indicate the serious medical conditions. If the flank pain is persistent then it can be symptom of renal cancer. Tumor size is the abnormal growth in the kidney or renal. The size of tumor will increase rapidly. If the size of tumor is less than 7.0 cm then it is normal. If the size will increase than 7.0 cm then it will be a serious issue to be consider. Von Hippel-Lindau syndrome is a symptom for renal cancer in which the hereditary conditions of a patient will be considered. This means that the development of risk of a cancer is due to VHL features that passed from generation to generation in a particular family. The tumor in the renal or kidney in a particular patient can be due to VHL genes. High blood pressure is also a risk factor for kidney or renal cancer. It is of two types: Diastolic blood pressure, which is the bottom number or lower number of blood pressure and the normal range of diastolic blood pressure is lower than 80 and Systolic blood pressure, is the top reading of blood pressure and the reading of systolic blood pressure below than 120 is normal. The both blood pressure of a patient should be normal. It cannot be too high. Trichloroethylene is the harmful and toxic chemical. If a person inhales this deadly chemical in large or excess amount then it will cause many harmful effects in a human body. It also affect on the renal or kidney. Hence, the trichloroethylene exposure is also a main symptom for the kidney or renal cancer

Artificial Neural Network

An artificial neural network is a model, which is designed to perform various tasks, or operations similar the way as a neuron in human brain do. The numerous capabilities like reasoning, remember and apply previous experiences of brain is provided by the biological neurons present in it. There are 1000 billion neurons in a human brain, which are connected with each other. The power or ability of a human brain is grow from these neurons and association among them. It is an electronic model that resembles with human brain and tries to do task as it is done by brain. In ANN, the nodes are considered as artificial neurons which helps to process complex and nonlinear data. These artificial
neurons have the ability to do their job in parallel and distributed. It also mimics the brain for the construction of various architectures, methods for reasoning and number of techniques used for functioning. The very first neuron network was proposed by McCulloch and Pitts in 1943 in which there are only 2 states of neurons depending upon their threshold value. After this invention, the number of models of neural network have been developed but the proposed model by McCulloch and Pitts is really a new and good opportunity for intelligent machines. The artificial neural network is applied in various domains such as engineering, medical field and many more. However, the main cause for study of artificial neuron network is to figure out how a human brain work and to develop a machine that helps to solve those complex problems which cannot be solved a normal computer that operate sequentially. In the case of artificial neural network, the input is given to the artificial neurons. These given inputs are multiplied with synaptic weights corresponding to them. The external bias is also provided to the system. This bias and weighted input will then send to summing function and then this value is fed to activation function. At the end, the final output is given by activation function. The artificial neural network can be modelled as shown in figure 1.

Adaptive Neuro-Fuzzy Inference System
The adaptive neuro-fuzzy inference system (ANFIS) is the mixture of two methods of soft computing: artificial neural network and fuzzy inference system. The adaptive neuro-fuzzy inference system (ANFIS) works in Sugeno fuzzy inference system. The structure of ANFIS is also resemble with multilayer feed-forward neural network, the only difference is in adaptive neuro-fuzzy inference system the signal flow directions are represented by links and also in ANFIS there are no any synaptic weights. The adaptive neuro-fuzzy inference system (ANFIS) works in Sugeno fuzzy inference system. The structure of ANFIS is also very much alike with multilayer feed-forward network. The only difference is that in adaptive neuro-fuzzy inference system the signal flow directions are represented by links and also in ANFIS, there are no any synaptic weights connected to the links. The rules used in ANFIS are as follow:

Rule 1: if x is X, and y is Y, then f₁ = a₁x+b₁y+c₁
Rule 2: if x is X₂ and y is Y₂, then f₂ = a₂x+b₂y+c₂

Where X, Y, X₁, Y₁, X₂ and Y₂ are membership functions and a₁, b₁, c₁, a₂, b₂, c₂ are determined during the process of training and known as designing parameters.

There are also four main steps in the architecture of fuzzy inference system

1. Fuzzification: The inputs given to the system are crisp values. In the fuzzification process, these crisp values are converted into fuzzy values.
2. Knowledge base: All the required knowledge, collected by the knowledge engineer, is stored in the knowledge base. All the rules and facts are generated by using this knowledge stored in it.
3. Inference engine: The inference engine extracts the knowledge from the knowledge base and performs various manipulations on that acquired knowledge to reach at a particular solution. It uses those rules and facts generated in knowledge base.
4. Defuzzification: The output generated by inference engine is in fuzzy value. The defuzzification process transform these fuzzy values to the crisp values.

The architecture of fuzzy inference system is illustrated in figure 2.
the clinicians to cure the patients suffering from terrible disease renal cancer under their guidance.

- Linguraru et al. (2009) introduce a tool that was used to distinguish renal tumors from each other. The CT images were first acquired, then from that images the tumors were distinguish and then the classification of that tumors had been done accordingly. In each step, the dimensions and shapes of lesions was processed to refine them. The CT images were segmented and histogram was constructed according to those segmented lesions. The renal tumors are classified according to their features extracted from the segmented CT images.

- Kong (2010) proposed a method that helps to track the movement of dimensions from the gathered images by using the technique known as image morphing. The developed method was carried out the feature of requirement from the target image by edge detection. For the mapping purpose of features, the near-neighbour algorithm was used. Thus, these extracted features were used to track the measurements of dimensions.

- Abhilash and Chauhan (2012) detects the movement of urological organs by using the prediction techniques, it also predicts the movement of internal kidney. Adaptive neuro-fuzzy inference system is used for the detection of position and hence, the non linear mapping was earned by the system. This research was accepted that the used algorithm and technique gave more accurate results with 94% accuracy.

- Haas and Nathanson (2014) recognized those genes that are correspond with hereditary kidney cancer. The kidney cancer that occurs because of genes can also become the reason behind various subtypes of kidney cancers. This is because the hereditary kidney cancer occurs due to mutations in genes. Hence, these genes can also lead to increase in pathological subtypes of kidney cancer. In his work, the two types of kidney cancer are analyzed with their other syndromes.

- Schwartz et al. (2015) investigated the change in kidney cancer after predicting five year by uni-model and multi model analysis the two machine learning methods were used for this investigation. These two methods are support vector machine (SVM) and k-nearest neighbour method (KNN). The accuracy is calculated by comparing the results of these two methods. After comparison, the result of this study justified that the machine learning method k-nearest neighbour method (KNN) is more accurate than support vector machine.

- Tander, Ozmen, and Ozden (2015) introduced various estimation models for the re-occurrence of kidney cancer. In this study, the Multilayer Perceptron Neural Network are designed to simplify and integrate the various estimation techniques. These post-operative estimation techniques are used by physician for the regular check-up of those patients of kidney cancer where there is chances of re-occurrence of renal cancer. The two Multilayer Perceptron Neural Networks are trained as well as tested Kattan's and Sorbellini's nomogram data.

- Vasilescu et al. (2017) gave some facets about those organs, which are close and affected by kidney cancer if it is not detected at initial stages. The thermographic detection can be used to prevent the affect of kidney cancer on other paired organs of body. It also analyzed the symmetry of thermal field correlated with each and every organ to which it was applied for detection. The result for this method is not concluded accurately yet because there are various errors in measurement.

- Jones et al. (2017) provided various mixtures of priorities. The doctors, expertise or professionals to give the suggestion about the health care in an effective manner can use these priorities. This paper elaborates the top ten ambiguities and research priorities. These ambiguities and priorities are: the biomarkers used for the detection of the disease, implementation of a medical diagnostic system that can help the doctor in the detection of particular disease, risk factors of that disease and many more like these.

- Li et al. (2017) classified the correlation of metformin in case of kidney cancer. For this classification, the eight datasets are used in which 254,329 total numbers of patients are present. The classification was done by analysing the various dissimilarities between the patients of kidney cancer associated with the metformin. The metformin helps to improve the overall survival (OS) and cancer-specific survival (CSS) of patients of kidney cancer.

- Deng et al. (2017) created a fused network which helps to figure out the fruitful information about the various kinds of data. This data helps to diagnose of a disease more correct and accurate. The generated fused network was basically used to determine the stage of renal cancer. It also merges the gene expressions with DNA methylation, which led for better detection of renal cancer. It classified the renal cancer into three categories and also its accuracy is better than all the methods used for the detection of renal cancer by using fused network.

- Que et al. (2018) gave the precise and organized review of articles of kidney cancer. In this paper, 35 different articles are reviewed and 4532 patients are involved in this articles for the detection of renal cancer. A meta-analysis has been done and the result of this meta-analysis suggested that decrease in overall survival was associated with low phosphatase and tensin homolog expression.

- Tuncer and Alkan (2018) developed a system which helps to detect the renal cancer by using the pictures of tissues of renal cancer and abdominal. The output achieved by this system help to make the decision more accurately and correctly. In this proposed system, there are two main steps. The first step is the image segmentation in which an image is captured and divided it into various parts. After that, the required features of kidney have been extracted by using the image processing...
techniques. In the second step, the detection of stage of kidney cancer has been done by using the classification methods named support vector machine with 92% accuracy.

- Aljouie (2018) examined the number of cases of kidney cancer for their cross-study validation and cross-validation. In the case of Kidney Renal Papillary Cell Carcinoma (KIRP) and Kidney Chromophobe (KICH), the accuracy gained by using the support vector machine is 71% and 72% respectively. However, when a model is trained by using the training dataset for the detection of Kidney Renal Papillary Cell Carcinoma (KIRP), then the observed accuracy is 66%.

- Technopole (2019) gave the brief explanation about the diagnose and treatment of kidney cancer. In this paper, it is described that in the early 2000s, the time period for the treatment of patient of kidney cancer is only about 15 months approximately to save the life of that particular victim. The main field of research for the renal cancer is to find out the exact biomarkers of kidney cancer that will help the expertise or professionals for the detection of kidney cancer at the early stages. The need of providing better management and proper drugs to the patients of renal cancer or kidney cancer is also there.

3. EMPIRICAL STUDIES

Data Description
The dataset was built by authors with the help of expertise and doctors and then used in this work. The used dataset contains 200 data of different patients suffering from the renal cancer. There is no any missing value in used dataset. Each instance of dataset indicates the risk factor or symptoms of renal cancer and results corresponding to given instances. The risk factors of renal cancer that has been considered in this proposed work are Haematuria, Red blood cell count, flank pain, tumor size, Von Hippel-Lindau syndrome, high blood pressure and trichloroethylene exposure.

Experimental Setup
The software used for the implementation of this developed work is MATLAB with ANFIS classifier. Initially, the dataset was created by taking samples of patients of renal cancer. As the dataset contains 200 samples, this data was partitioned into two parts using direct partitioning method and the partitioning ratio is 75:25 respectively. First part is training data. For the training process, the 75% of data from entire dataset has been chosen. Hence, the training phase has been done with adaptive neuro-fuzzy inference system by using this data. Similarly, the second part is testing data. The remaining 25% of data is used for the testing process. The training data and testing data is different from each other. Figure 3 illustrate the flow diagram of developed intelligent system.

4. DEVELOPMENT OF NEURO-FUZZY SYSTEM

An intelligent system for the detection of renal cancer is developed by using adaptive neuro-fuzzy technique. The input variables considered in the development of this medical intelligent system are elaborated in the introduction section. The dataset that was used for training and testing phase is also described in the previous section. The input variables in the developed diagnostic system for renal cancer are assigned as input1, input2, input3, input4, input5, input6 and input7 and the output variables are assigned ad output. The figure 4 shows the structure of proposed medical intelligent system for renal cancer by using adaptive neuro-fuzzy inference system.

Membership functions
For every input variable, the triangular membership function has been used. Three triangular membership functions are used for input1, input2, input3 and input 7. Similarly, for input4, input5 and input6, two triangular membership functions are used for each. For example, for input variable red blood cell count (RBCC), the three groups are made. These three groups are low, medium and high. The adaptive neuro-fuzzy inference system will automatically generate these groups of input variables from the loaded training dataset. Similarly, the membership function is generated for the output variables corresponding to the given input variables. There are four membership functions.
that are generated from training data during the training phase as four stages of cancer of renal or kidney. Figure 5, figure 6, figure 7 and figure 8 shown the membership functions for input1, input2, input3 and input4 respectively. These membership functions are not defined by manually. These are generated from the loaded dataset of training during the training phase of adaptive neuro-fuzzy inference system. Also, the rules used to achieve the required output by the system is automatically setup by the system.

Figure 5: Membership function for input variable 1 i.e. Hematuria

Figure 6: Membership function for input variable 2 i.e. RBCC

Figure 7: Membership function for input variable 3 i.e. Flank Pain

Figure 8: Membership function for input variable 4 i.e. Tumor size.

Rules
All the probable mixture of used input variables for the detection of renal or kidney cancer in adaptive neuro-fuzzy inference system are used for the development or generation of rules. The total number of rules that are used in developed ANFIS for renal cancer are 648. The total number of rules can be calculated as follow: Total rules= membership function of input1 * membership function of input2 * membership function of input3 * membership function of input4 * membership function of input5 * membership function of input6 * membership function of input7.

Hence, total rules= 3*3*3*2*2*2*3 =648 rules

Training
The training phase of adaptive neuro-fuzzy inference system (ANFIS) by using loaded training dataset is described in this section. The figure 9 shows the training error at the epoch 3. Form the total dataset of renal cancer, the expertise or specialist taken 75% of it for the training phase of adaptive neuro-fuzzy inference system. As the adaptive neuro-fuzzy inference system is a hybrid system, therefore it is used to enhance the speed of diagnostic system.

Testing and Validation
After the training phase, now the time is to test the develop system by testing phase. The rest of 25% of total dataset of renal cancer has been used for the testing phase. The dataset used for testing data must be different from the training dataset to validate the accuracy of developed inference system for the detection of renal cancer by using neuro-fuzzy inference system (ANFIS). During the validation of this system, the main evaluation has been done on how much the developed system has been trained to classify the patients, between renal cancer patients and non renal cancer patients, correctly and accurately. In the validation, it is also figure out that the developed diagnostic system will also gave the correct and accurate output for unseen or new data or not. The diagnostic system must be trained in such a way so that it can give proper and appropriate output for any kind of data given to the diagnostic system as the input.
5. RESULT

After acquiring the output from the develop diagnostic system for renal or kidney cancer by using adaptive neuro-fuzzy inference system (ANFIS), now the specialist doctors or expert detects the obtained output by comparing it with the actual or target output. Here, the target output is that output which is evaluated by the expertise for a particular input and same inputs are given to diagnostic system. After the comparing between obtained and target results, it is noticed that the developed diagnostic system for renal cancer gives the similar output as the target outputs. The performance of the diagnostic system for the renal cancer is calculated by considering the various parameters. These parameters are sensitivity, specificity, precision and classification accuracy. The table 1 shows the performance of developed diagnostic system for renal cancer by using adaptive neuro-fuzzy inference system by considering the various parameters.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Model</th>
<th>Membership function</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Precision (%)</th>
<th>Classification Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renal Cancer</td>
<td>Adaptive neuro-fuzzy inference system</td>
<td>Triangular</td>
<td>95.8</td>
<td>96.1</td>
<td>95.8</td>
<td>96</td>
</tr>
</tbody>
</table>

These calculated parameters is shown in the form of graph in figure 10.

6. CONCLUSION AND FUTURE WORK

In this study, an intelligent medical diagnostic system for the diagnose of renal cancer by using adaptive neuro-fuzzy inference system has been proposed. This proposed method can help the nephrologists in the detection of stage of renal cancer. It provides the efficient way for the diagnose of renal cancer. This medical diagnostic system can be used as a supportive tool for the expertise or doctors to save the life of their patient from the life-threatening disease by detecting the renal cancer at initial stage. Professionals as well as non-professionals can use this developed system for renal cancer. The proposed system is trained by using the dataset during the training phase. Similarly, it is tested and validated in the testing phase. The performance is also evaluated by considering various parameters such as sensitivity, specificity, precision and classification accuracy. After these calculations, it is observed that the proposed intelligent medical diagnostic system for renal cancer by using adaptive neuro-fuzzy inference system (ANFIS) gives the better results with classification accuracy 96%. The MATLAB software is used for the implementation of this diagnostic system. In the future work, this research work can also be pursued by adding more risk factors or input variables for the detection of renal cancer. This will leads to enhancement in the performance of the system and can be used in those areas where there is a lack of resources so that the valuable life of patients can be saved.

7 REFERENCES


