CLASSIFICATION OF CHRONIC KIDNEY DISEASE STAGES IN DIABETIC PATIENTS

S. Mahalakshmi[1], P. Menaka[2], R.S. Rajkumar[3]

[1,2]Student, [3]Assistant Professor

Department of Computer Science and Engineering,

Sri Ramakrishna Engineering College, Tamil Nadu, India

ABSTRACT

Chronic Kidney Disease is a type of kidney disease in which there is gradual loss of kidney function over a period of months or years. The two main causes of chronic kidney disease are diabetes and high blood pressure, which are responsible for two-thirds of the cases. Diabetic nephropathy (diabetic kidney disease) is kidney damage that results from having diabetes. Having high blood glucose levels due to diabetes can damage the part of the kidneys that filters our blood. The main objective of this work is to predict chronic kidney disease in diabetic patients using classification algorithms such as Radial Basis Function Neural Network (RBFN), Multilayer Perceptron (MLP) and Convolution Neural Network (CNN). This research work also focused on finding the best classification algorithm based on the classification accuracy.

Keywords:

Chronic Kidney Disease, Glomerular Filtration Rate, Radial Basis Function Neural Network, Multilayer Perceptron, Convolutional Neural Network.

1. INTRODUCTION

Chronic Kidney Disease [5] (CKD) is one of the worldwide public health problems due to the costly treatment at its end stage and high possibility of death. The two main causes of chronic kidney disease are diabetes and high blood pressure, which are responsible for up to two-thirds of the cases. According to recent report of Kidney Disease Improving Global Outcome (KDIGO), this disease is divided into five stages based on the defined range of Glomerular Filtration Rate (GFR): CKD1, CKD2, CKD3, CKD4 and CKD5. To avoid the death patients in CKD5 stage have to perform either kidney transplantation or dialysis. The most common cause of kidney failure in population-based studies is Diabetic nephropathy.

Diabetic nephropathy [11] (diabetic kidney disease) is a kind of kidney disease that damage the kidney function that results from having diabetes. One in 4 women and one in 5 men with type 2 diabetes develops diabetic nephropathy. It is even more common in type1 diabetes. Here, Artificial Neural Network is used for classifying kidney disease into its stages.

Artificial Neural Network [12] is a popular strategies for supervised machine learning and classification. An artificial neural network is an attempt to simulate the network of neurons that make up a human brain so that the computer will be able to learn things and make decisions in a humanlike manner. It has three layers such as input layer, hidden layer and output layer. Input layer receives various forms of information from the outside world. From the input layer, the data goes through one or more hidden layers. The hidden layer's job is to transform the input into something the output unit can use. The output layer's job is to process the input which is obtained by its hidden layer.

2. LITERATURE SURVEY

Veenita Kunwar et.al [1] proposed, Chronic Kidney Disease has been predicted and diagnosed using data mining classifiers: Artificial Neural Network and Naive Bayes. Performances of these algorithms are compared using Rapid miner tool. The obtained results showed that Naive Bayes is the most accurate classifier with 100% accuracy when compared to ANN having 72.73% accuracy.

Ammarah Farooq et.al [3] designed, Convolution network has been used for classifying structural MRI images to diagnose Alzheimer's disease. MRI images are preprocessed to get GM images which are then passed to CNN network. Networks are trained and tested using Google Net and Res Net models. About 4% increase in classification accuracy is achieved by this proposed approach.

Pinar Yildirim et.al [4] proposed, a comparative experiment is carried out on sampling algorithms to predict chronic kidney disease. The effect of class imbalance in training data on performance is evaluated for multilayer perceptron. Three sampling algorithms namely, Resample, SMOTE and Spread sub sample algorithms are used to analyze the data set and their performance is evaluated by multilayer perceptron by varying learning rate.

Thamaraiselvi. V. G et.al [6] developed, a novel prototype architecture has been designed for the disease prediction of infants, based on the combination of parent DNA using Artificial Neural Networks. The proposed system integrates Microarray Technology, Digital Image Processing and Artificial Neural Networks. By applying microarray technology to the parental DNA, a gene expression image can be obtained. The extracted gene image is further characterized with the help of Digital Image processing. A neural network is trained with the mutated value which lists the probability of diseases to the infants.

Raghu Raj. P et.al [10] developed few novel ideas and algorithms to systematically process the retinal images and infer on the nature and extent of disease. The image is processed and obtained in the required format. The Cotton Wool Spots are detected and quantized (if present) using the color splicing technique. The Hemorrhages are detected and quantized (if present) using color segmentation. Veins are extracted using the bottom hat transformation.

Patcharaporn et.al [5] developed a classification model for predicting transitional interval of kidney disease stages 3 to 5. The existing medical records of Hemodialysis patient from Pan Hospital, Chiang Ray, Thailand, have been exploited as the case study. Decision Tree, K-nearest neighbor, Naïve Bayes and Artificial neural networks were used for eliciting the knowledge and creating classification model with the selected set of attributes.

Abdulrahman Shamsan et.al [8] designed, predictive analytics techniques such as Decision Tress, Logistic Regression, Naive Bayes, and Artificial Neural Networks are used for predicting CKD. These predictive techniques are assessed and compared based on accuracy of prediction. The study provides a decision support tool that can help in the diagnosis of CKD.

Sahana B J et.al [7] proposed a medical sector application, which helps the medical practitioners in predicting the disease types based on the symptoms. Patients can also predict disease by entering symptoms in the form of sentences. It is automation for disease prediction and it identifies the disease, its types and complications from the clinical database in an efficient and an economically faster manner. Naïve Bayes algorithm is used for classification and C4.5 algorithm is used for stages prediction. This algorithm takes symptoms as input and predicts the disease based on old patients data.

S.Ancy et.al [2] designed RBFNN (Radial Basis Function Neural Network) to determine the condition of a patient as a normal or a kidney failure patient. This method reveals the stages of the kidney failure patient and treatment and clinical decision. This work, contributes the Big Data Reduction using Distributed Radial Basis Function Neural Network (DRBFNN)

3. ARCHITECTURE

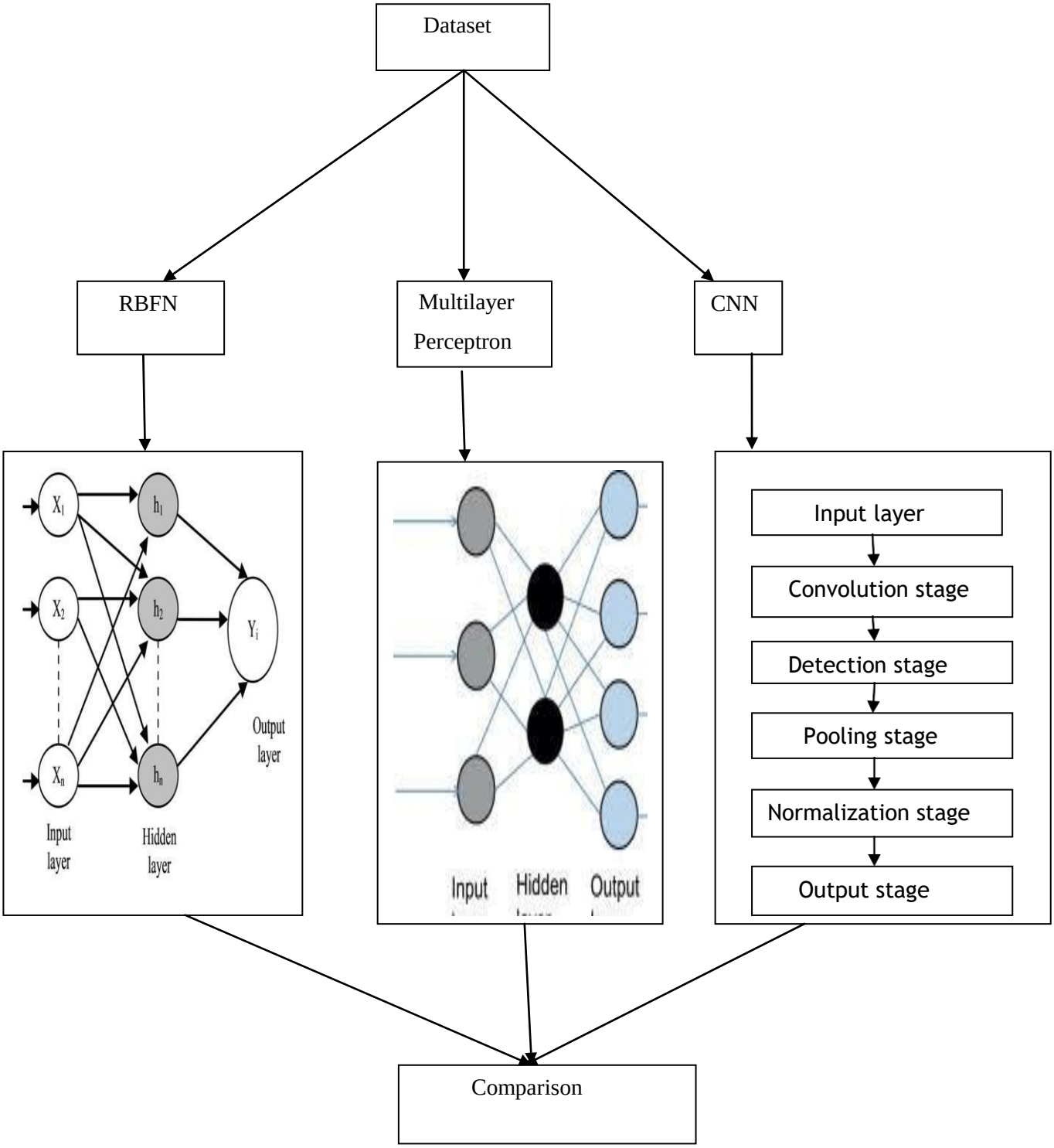

Fig. 1 Architecture of proposed system

4. GFR CALCULATION

Glomerular Filtration Rate is used for finding the kidney function level. It can be calculated using attributes such as age, gender, serum creatinine. GFR calculated using the below formula [5]:

$$GFR_{Female\ (Creatinine<0.7)} = 144*(Creatinine/0.7)^{(-0.329)} *(0.993)^{Age} \quad (4.1)$$

$$GFR_{Female\ (Creatinine>0.7)} = 144*(Creatinine/0.7)^{(-1.209)} *(0.993)^{Age} \quad (4.2)$$

$$GFR_{male\ (Creatinine<0.9)} = 141*(Creatinine/0.9)^{(-0.411)} *(0.993)^{Age} \quad (4.3)$$

$$GFR_{male\ (Creatinine>0.9)} = 141*(Creatinine/0.9)^{(-1.209)} * (0.993)^{Age} \quad (4.4)$$

Chronic Kidney Disease Stages are split using the value of GFR. Stages of chronic kidney disease are classified in below form.

- stage1: GFR>90
- stage2: GFR=60-89
- stage3: GFR=30-59
- stage4: GFR=15-29
- stage5: GFR<15

5. RADIAL BASIS NEURAL NETWORK

RBFN is supervised and three-layered feed forward neural network. The hidden layer of RBFN consist of a number of nodes and a parameter vector called a center which is considered the weight vector. The Euclidean distance is used to measure the distance of an input vector from the center [2]. The transfer function of a RBFN is mostly built up of Gaussian rather than sigmoid center. The Gaussian functions decrease with distance from the center.

The Euclidean length is represented by r_n that measures the radial distance between the datum vector p (p1, p2, p3…pi) and the radial center $p^{(j)}$ = (w1n, w2n…..w i n)

$$r_n = \|p-p^{(n)}\| = \quad _m\text{-}w_{in})^2)^{1/2} \quad (5.1)$$

Here $r_n = \| \|$ is the Euclidean norm, is the activation function and w_{mn} is the connecting weight between the m^{th} hidden unit and n^{th} output unit. A suitable transfer function is then applied to r_n to give $(r_n) = \|p-p^{(n)}\|$. Finally, the output layer (i-1) receives a weighted linear combination of (r_j),

$$P^{(i)} = w_o + \quad _n{}^{(i)} \ (r_n) = = w_o + \quad {}^{(i)}_n \ (\|p-p^{(n)}\| \ ) \quad (5.2)$$

Here c_n is the center of the neuron in the hidden layer and (r_n) is the response of the n^{th} hidden unit and w_o is the bias term.

6.MULTILAYER PERCEPTRON

Feed forward neural networks [4] are widely used models for classification, forecasting and problem solving. A typical feed-forward back propagation neural network model is proposed to classify the CKD. Our Multilayer Perceptron architecture is a feed forward neural network model that consist of three layers:

- Input Layer,
- Three Hidden Layers and
- Output Layer

For classification of CKD, two hidden layers with 20 neurons is created and trained. The two hidden layer helps in improving the network accuracy. The hidden layer neurons are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs. The output of the hidden layer can be represented as

$V_i = {}_{in}X_n$ (6.1)

Error calculation:

$E_o(i) = Y_{d.o}(i) - Y_o(i)$ (6.2)

Final hidden layer output:

$Y_i = (V_i)$ (6.3)

Here, W_{i1} …….. W_{ij} are the value of synaptic weight

W_{io} is the value of bias

$E_0(i)$ is the value of error

$Y_{d.o}(i)$ is the desired value of output

V_i is the output value before the process of activation function

Ω(.) is the activation function

Y_i is the final output for hidden layer.

7. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks have been utilized to great success in numerous vision classification problem and to extract information in medical imaging tasks such as mitosis detection in breast cancer and knee cartilage segmentation. In this proposed work, CNN is used to predict kidney function in patients with chronic kidney disease from high- resolution digital pathology scans of their kidney biopsies. Human visual cortex inspires convolutional neural networks. Each neuron responds to stimuli occurring in its receptive field [3].This operation is similar to convolution and input image patch works as its small receptive field. When input pattern matches with the convolving filter, response is produced in the form of feature maps.

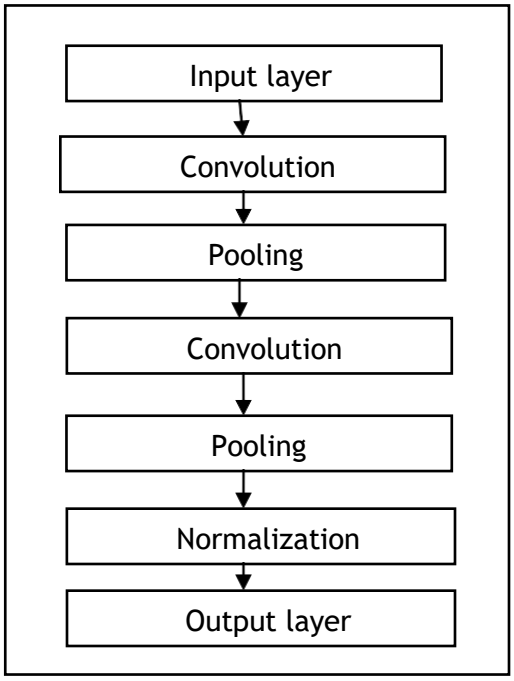

Fig. 2 Architecture of CNN

8. IMPLEMENTATION & RESULTS

The clinical data consist of 600 patient's records which are has been considered for the analysis and that dataset is taken from UCI Machine Learning Repository [11]. Totally 8 attributes are there in the dataset. Numerical and nominal values of attributes are considered. The attributes of the dataset are class, age, gender, blood sugar, serum creatinine, diabetes, Glomerular Filtration Rate, Stages of kidney disease. This dataset is used for Radial Basis Function Neural Network and Multilayer Perceptron Neural Network algorithms. MRI Kidney Disease images are used in Convolution Neural Network classification algorithm. Here, R tool is used for classification.

The accuracy & loss of the three algorithms are shown in table1:

Algorithms	Training set Accuracy	Training set Loss	Testing set Accuracy	Test set Loss
RBFN	88%	12%	85	15%
MLP	91%	9%	90	10%
CNN	94%	6%	93	7%

Table 1: Comparison of RBFN, MLP, CNN

When compared, Convolution Neural Network has higher accuracy and minimum loss.

9. CONCLUSION

In this study, a comparative experiment is carried out on learning algorithms to predict chronic kidney disease. RBFN, MLP and CNN performance are evaluated. Among the algorithms, CNN has better accuracy results on the data set than the others. This study reveals that ANN algorithms can improve the performance of classification rate.

10. REFERENCES

1. S.Ancy, K.Comelius, "A Predictive Model for Kidney Failure E-health", IEEE 3rd International Conference on Sensing, Signal Processing and Security (ICSSS), 2017.
2. Ammarah Farooq, Syed Muhammad Anwar, Muhammad Awais, Saad Rehman, "A Deep CNN based Multi-class Classification of Alzheimer's Disease using MRI", *IEEE Instrumentation and Measurement Society prior to the acceptance and publication.,* 2017.
3. Pinar Yildirim, "Chronic Kidney Disease Prediction on Imbalanced Data by Multilayer Perceptron", *IEEE 41st Annual Computer Software and Applications Conference,* 2017.
4. Veenita Kunwar, Khushboo Chandel, "Chronic Kidney Disease Analysis Using Data Mining Classification Techniques", *6th International Conference – Cloud System and Big Data Engineering (Confluence,)* 2017.
5. Patcharaporn, panwong, Natthakan, "Predicting Transitional Interval of Kidney Disease Stages 3 to 5 Using Data Mining Method", *Second Asian Conference on Defense Technology (ACDT),* 2017.
6. Sahana B J, Dr. Minavathi, "Kidney Disease Prediction Using Data Mining Classification Techniques and ANN", *International Journal of Innovative Research in Computer and Communication Engineering* Vol. 5, Issue 4, April 2017.
7. Abdulrahman Shamsan, Faisal Aqlan, "Data Mining for Chronic Kidney Disease Prediction", *Proceedings of the Industrial and systems Engineering Conference,* 2017.Guneetkaur, Ajay Sharma, "Predict Chronic Kidney Disease Using Data Mining Algorithms in Hadoop", *International Journal of Advanced Computational Engineering and Networking,* ISSN: 2320-2106, Volume- 5, Issue 6, Jun-2017.
8. Thamaraiselvi. V. G, Vishnu Kumar Kaliappan, "Infants Disease Prediction Architecture Using Artificial Neural Networks and Digital Image Processing", *World Congress on Computing Technologies,* 2017.
9. Vijayalakshmi. J, "Kidney Failure Due to Diabetics – Detection using Classification Algorithm in Data Mining", *International Journal of Data Mining Techniques and* Applications Volume: 06, Issue: 02, Page No.62-64 ISSN: 2278-2419, December 2017.
10. Tabassum. S, Jharna Majumdar, "Analysis and Prediction of Chronic Kidney Disease using Data Mining Techniques", *International Journal of Engineering Research in Computer Science and Engineering* Vol 4,Issue 9, September 2017.
11. Kamat. R. K, Kavita. S, "Big Data Analysis: Using R Analytical Tool", *IEEE International conference on I-SMAC,* 2017.
12. Raghu Raj. P, Gurudatha Bhai. K, "Algorithmic Approach for Prediction and Early Detection of Disease Using Retinal Images", *IEEE Computer Graphics, Imaging and Visualization,* 2017.
13. Md.Osman Goni Nayeem, Maung Ning Wan, Md. Kamurul Hasan, "Prediction of Disease Level Using Multilayer Perceptron of Artificial Neural Network for Patient Monitoring" , *International Journal of Soft Computing and Engineering,* ISSN: 2231-2307, Volume-5 Issue-4, September 2015
14. http://archieve.ics.uci.edu/ml/datasets/chronic_kidney_disease#.