International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

CHRONIC KIDNEY DISEASE PREDICTION USING NEURAL NETWORKS

S.Priya[1], S.Nirmal Kumar[2], G.Sibi Saravanan[3], E.Pandiyan[4], Ashuthosh Kumar Pandey[5]

[1]Assistant Professor, Department of Computer Science Engineering, Coimbatore Institute of Technology, Tamil Nadu, India
[2,3,4,5]UG Student, Dept. of Computer Science Engineering, Coimbatore Institute of Technology, Tamil Nadu, India

---

Abstract - *Chronic kidney disease (CKD) is a leading cause of mortality around the world. Providing diagnostic aid for CKD disease by using a set of data that contains only medical information obtained without advanced medical equipment can help people who want to discover the disease or the risk of disease at an early stage. The aim of our project is to classify chronic kidney disease (CKD) by developing a system using machine learning. Our method is implemented by classification approach using artificial neural network (ANN), Keras python Library for sequential model creation. The model used is feed forward network with back propagation algorithm. The system can assist medical practitioners in the already existing diagnosis systems. It can also help the patients to know earlier if they are having CKD or likely to have by using certain attributes.*

Key Words: *Chronic Kidney Disease (CKD), Artificial Neural Network (ANN), etc.*

1. INTRODUCTION

Chronic Kidney Disease, also known as chronic kidney failure, is a gradual loss of kidney function. The kidneys filter waste and fluids from our blood which are excreted through urine. The kidneys balance the salts and minerals such as calcium, phosphorus, sodium, and potassium that circulate in the blood. They also make hormones that help control blood pressure, make red blood cells and keeps the bones strong. When chronic kidney disease reaches an advanced stage, dangerous levels of fluid, electrolytes and wastes can build up in our body. The early stages of chronic kidney disease is characterised by few signs or symptoms, which makes it difficult to identify the disease. The symptoms may become apparent only when the kidney function is significantly impaired [1]. If left untreated, Chronic kidney disease can progress to end-stage kidney failure, which is fatal without artificial filtering (dialysis) or a kidney transplant.

The most challenging task in medical field is to develop a diagnostic method using fast and accurate algorithm which runs faster and gives correct results. So it has become a demanding issue to create a reliable and powerful medical diagnosis system to support the existing diagnostic processes. The conventional diagnosis is complex hence computing methods like artificial neural networks can be utilized in development of diagnosis. In diagnosing a disease due to time and data constraints, even partial detection of disease can be much helpful. Thus artificial neural networks are a promising means for diagnosis.

2. ARTIFICIAL NEURAL NETWORKS INTRODUCTION

Artificial Neural Networks which is also known as ANN is basically a computational model. These computing models are a complex network of basic components or a set of nodes called neurons. The nodes are interconnected with each other. Each connection between the nodes has a specific weight associated with it. The basic structure of any neural network consists of three layers. The first layer is called the input layer and the middle layer called the hidden layer and the final layer called the output layer. Every neural network will definitely have all these three layers with one or multiple nodes. A simple and basic neural network will have one hidden layer. But a complex neural network can have several hidden layers. The inputs to the network are fed through the initial input layers. The computation is done in the hidden layers and the respective output is popped out through the output layers. They are many different kinds of neural networks available. Nowadays neural networks with multiple hidden layers are popular and play a vital role in several huge computational purposes. One great example of neural networks with multiple hidden layers is the Multi-Layer Perceptron (MLP). An MLP can even solve a complex problem which a simple neural network cannot.

Multi Layer Perceptron (MLP)

The Multi-Layer Perceptron consists of many perceptrons. MLPs have the capability to compute non-linear functions. Hence they are primarily used in supervised learning for regression and classification. In the above network structure, inputs to the network are fed through three input nodes, computing is done throughout two successive hidden layers.

Finally the output is achieved at the output layer as a binary value of yes or no based on if the patient is infected with the disease or not. Figure1 shows multi layer perceptron structure.

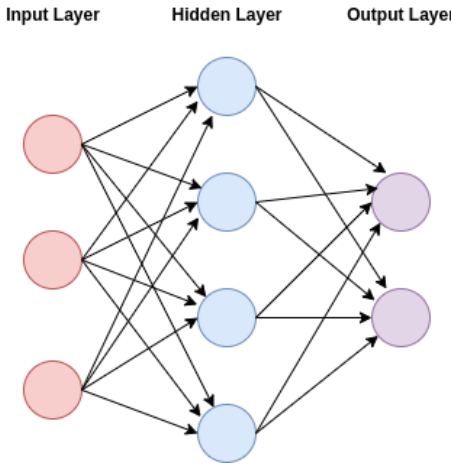


Figure 1: Multilayer Perceptron Architecture

Working of ANN

The working concept of the artificial neural network is inspired by the working principle of the human brain. The human brain consists of several billion neurons and passes information between them. Likewise is the concept of ANN with several nodes in it.

The interconnections between the nodes are called synapses. Each connection has a weight associated with it. The weights of each connection get updated accordingly after every iteration of the computational process in the hidden layers.

These days the ANNs are widely used for the purpose of diagnosis of diseases. Due to is wide learning capabilities and fault tolerance, it is most popular in medical diagnosis. One of the most popular structures of networks used is the feed-forward network(FFN). In FFN the passing of data or information is allowed only in the forward direction from one node in the current layer to one or more nodes in the next layer. A back propagation neural network is a type that is used in the classification process to classify between a person who is infected and the one who is not.

Feed Forward Network (FNN)

The connections in the Feed Forward Network (FNN) do not form loops or cycles in the network of nodes and their connections. These Feedforward networks are basically used for supervised learning in case of sequential data. One form of the feed-forward network used is the Multi-layer perceptron.

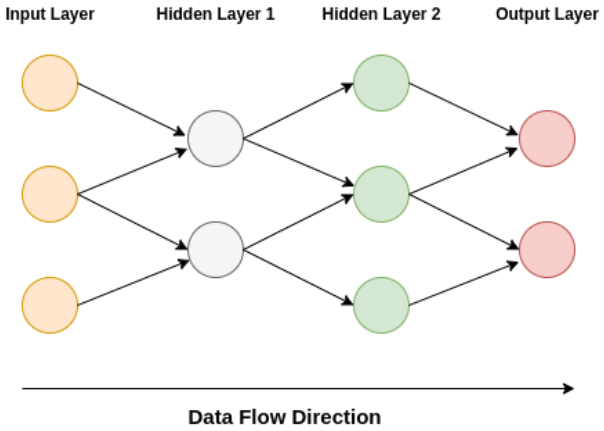


Figure 2: Feed Forward Architecture

3. RELATED WORK

There is a continuous study and research going on in the field of medical diagnosis. A lot of work has been done on diseases like Cancer, Diabetes, Heart attack etc. using neural networks. Berina Alic et al. (2015) presents "Machine Learning Techniques For Classification Of Diabetes And Cardiovascular Diseases" this paper presents the overview of machine learning techniques in classification of diabetes and cardiovascular diseases (CVD) using Artificial Neural Networks (ANNs) and Bayesian Networks (BNs). It was concluded that the higher accuracy was achieved with ANN. Suresh Kumar , Sathiya Priya et al.(2018) presents" Chronic Kidney Disease Prediction Using Machine Learning " which analysed two machine learning algorithms Decision Tree(DT) algorithm, Naive Bayesian (NB) algorithm .The performance of Decision tree method was found to be 99.25% accurate compared to naive Bayes method. El-housaany Rady and Ayman Anwar et al. (2019) presents "Prediction of Kidney Disease Stages Using Data Mining Algorithms" compared the results of applying Probabilistic Neural Networks(PNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM) and Radial Basis Function (RBF). The findings showed that PNN gives an overall classification accuracy of 96.7% and took longer execution time whereas MLP took just 3s. Ravindra ,Sriram et al. (2018) presents "Chronic Kidney Disease Prediction Using Back Propagation Neural Network Classifier" used four kidney dialysis attributes for classifying using an ANN. The study revealed that application of normalization procedure increased training efficiency. Asif Selekin, John Stankovic et al. (2016) presents "Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes" studied on selecting important predictive attributes using wrapper and LASSO regularization and identified a cost effective highly accurate detection classifier using only five attributes. Pinar Yildrim et al.(2017) presents "Chronic Kidney Disease Prediction On Imbalanced Data By Multilayer Perceptron" .The study revealed that sampling algorithms can improve the performance of classification algorithms and learning rate is a a crucial parameter in MLP.

4. KIDNEY DISEASE DATASET

The dataset used in this work for the diagnosis and prediction of chronic kidney disease is one hundred percent legitimate data collected from several different real patients over a period of time. The dataset is obtained from the UCI repository of datasets. The data set consists of records of 400 different patients under 25 different attributes. But in this project, we use only 14 attributes after eliminating irrelevant features.

Out of those 14 features used, 13 of them are given as inputs to the model and the 14th attribute is the classification in which the output is arrived at.

Table I: Attributes used for classification

Attribute	Data Type
Red blood cells	Normal, Abnormal
Pus cells	Normal, Abnormal
Pus cell clumps	Present, Not Present
Bacteria	Present, Not Present
Packed cell volume	Numeric values
White blood cell count	Numeric values
Red blood cell count	Numeric values
Hypertension	Yes, No
Diabetes mellitus	Yes, No
Coronary artery disease	Yes, No
Appetite	Good, Poor
Pedal edema	Yes, No
Anemia	Yes, No
Classification	Yes, No

5. ARTIFICIAL NEURAL NETWORKS CLASSIFIER

Multilayer Perceptron with BPA

A multilayer perceptron is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate output. It is a modification of the standard linear perceptron in that it uses three or more layers of neurons (nodes) with nonlinear activation functions, and is more powerful than the perceptron in that it can distinguish data that is not linearly separable, or separable by a hyper-plane.

Backpropogation is the most widely applied learning algorithm for multilayer perceptron in neural networks. Backpropogation employs gradient descent to minimize the squared error between the network output values and desired values for those outputs. These error signals are used to calculate the weight updates which represent knowledge learnt in the networks.

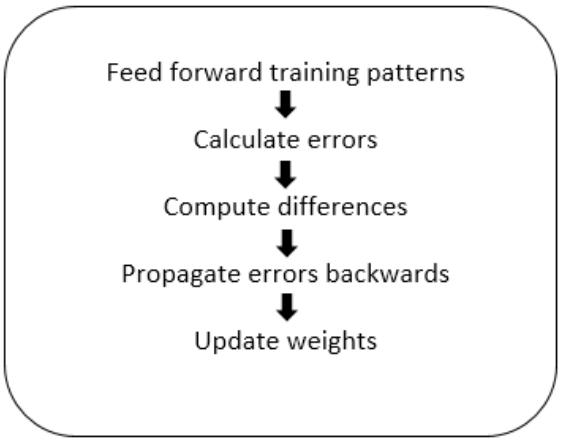


Figure 3: Backpropagation epoch cycle

6. EXPERIMENT RESULTS AND DISCUSSION

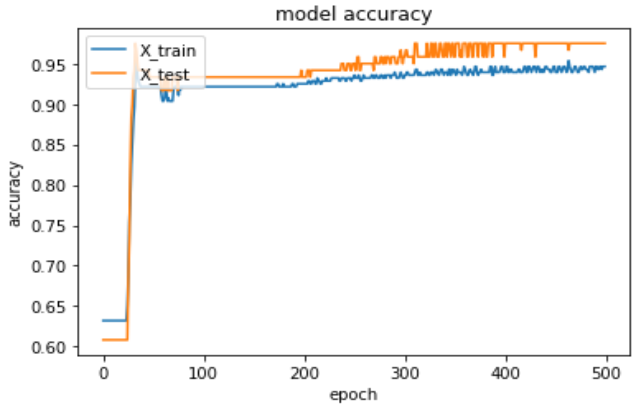
A neural network model is chosen and trained with examples of all classes. After successful training, the system is able to diagnose the unknown cases and to make predictions. The optimizer used for neural network model is Adam optimizer. The loss function binary crossentropy is used since the prediction is yes/no decisions i.e, binary classification. The neural network model is trained with 280 instances or data items and testing of neural networks is done with 120 data instances. The number of hidden layers used is 1 which is sufficient. The number of epochs is set at 500 .

Parameters	Values
Number of Training data	280
Number of Testing data	120
Number of Hidden Layers	1
Accuracy	95.36

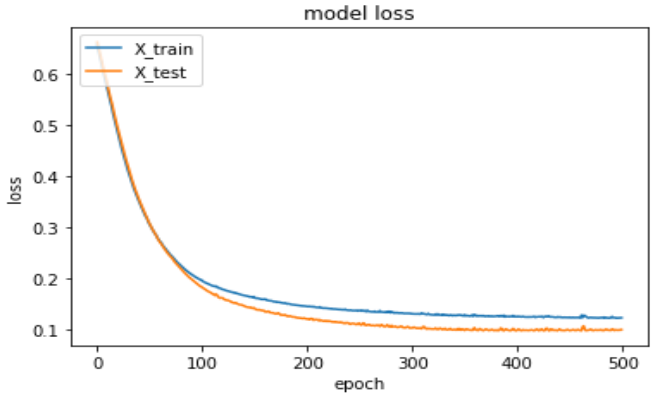
Table II: MLP Parameters

Table III: Result

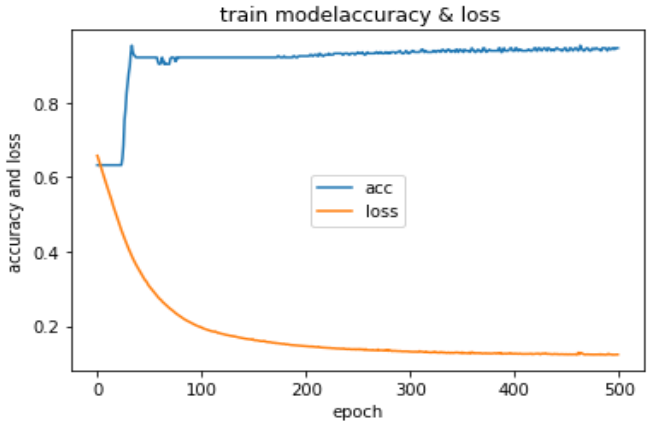
Correctly classified instances	116
Incorrectly classified instances	4
Accuracy	95.36%



The accuracy of model with the training data lies around the neighbourhood of 90-95%. But as it gets trained we can see the accuracy while testing it well beyond the 95% mark.We settle at an average of 95.36%.



As we can see the loss is way to high at the beginning but as the training and testing goes, it goes way down to less then a percent.



Initially when the model starts execution the accuracy is pretty much lower than loss, but as the model iterates through several hundred epochs, we can see a drastic jump in the level of accuracy upto 99% and the loss going down to as much as negligible.

Siddheshwar Tekale, Pranjal Shingavi, Sukanya Wandhekar, Ankit Chatorikar et.al (2018) presented the paper "Prediction of Chronic Kidney Disease Using Machine Learning Algorithm". From the results analysis, it is observed that the decision tree algorithms gives the accuracy of 91.75% and support vector machine (SVM) gives accuracy of 96.75%.

Thus we can come to a conclusion that neural network model used is very efficient and accurate in classifying the disease.

CONCLUSION

In the work presented, neural networks is used for diagnosis of chronic kidney disease . With the above results we have achieved our objective to find the best model for CKD diagnosis. The multilayer perceptron with back propagation algorithm is a good model for diagnosis of CKD, Its accuracy is 95.36%. The error rate is also considerably low. Thus we have come to conclusion that multilayer perceptron trained with back propagation is one of the most suitable and efficient algorithm for kidney disease diagnosis. Limitation of the application is we do not have data of strength because of size of data set and missing attribute values.

REFERENCES

[1] www.mayoclinic.org/diseases-conditions/chronic-kidney-disease/symptoms-causes

[2] Koushal Kumar, Abhishek et al."Artificial Neural Networks for Diagnosis of Kidney Stones Disease "2012

[3] Berina Alic et al. "Machine Learning Techniques for Classification of Diabetes and Cardiovascular Diseases" 2015

[4] Suresh Kumar, Sathiya Priya et al. "Chronic Kidney Disease Prediction Using Machine Learning "2018

[5] El-housaany Rady and Ayman Anwar et al. "Prediction of Kidney Disease Stages Using Data Mining Algorithms" 2019

[6]Ravindra, Sriram et al. "Chronic Kidney Disease Prediction Using Back Propagation Neural Network Classifier" 2018