ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

# Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms

S.Ramya, Dr. N.Radha

Research Scholar, Department of Computer Science, PSGR Krishnammal College for Women, Coimbatore, India

Assistant Professor, Department of Computer Science, PSGR Krishnammal College for Women, Coimbatore, India

**ABSTRACT:** Chronic Kidney Disease (CKD) is a gradual decrease in renal function over a period of several months or years. Diabetes and high blood pressure are the most common causes of chronic kidney disease. The main objective of this work is to determine the kidney function failure by applying the classification algorithm on the test result obtained from the patient medical report. The aim of this work is to reduce the diagnosis time and to improve the diagnosis accuracy using classification algorithms. The proposed work deals with classification of different stages in chronic kidney disease according to its severity. The experiment is performed on different algorithms like Back-Propagation Neural Network, Radial Basis Function and Random Forest. The experimental results show that the Radial basis function algorithm gives better result than the other classification algorithms and produces 85.3% accuracy.

**KEYWORDS**: Chronic Kidney Disease (CKD), Data mining, Machine Learning (ML), Back-Propagation Neural Network, Radial Basis Function and Random Forest**.**

## I. INTRODUCTION

Data mining plays a vital role in health care domain, nowadays. There is an increased need for an efficient analytical methodology to detect unknown and valuable information in health data. It produces huge amount of data about patients, diseases, diagnosis and medicines so on. By applying data mining techniques in health care domain, the administrators can improve the quality of service by discovering latent, potentially useful trends required by medical diagnosis. In the health care industry, the data mining is mainly used for predicting the diseases from the datasets [1].

Classification is one of the most widely used methods of data mining in healthcare organization. The classification technique predicts the target class for each data points. The classification methods such as Decision Tree, Support Vector Machine, K-Nearest Neighbour, Naïve Bayes and Neural Network. Decision Tree is widely used by many researchers in healthcare field like skin diseases and chronic kidney disease etc. The K-Nearest Neighbour is used to analyse the relationship between cardiovascular disease, hypertension and the risk factors of various chronic diseases in order to construct an early warning system. Multilayer Neural Network is used for diagnosis of various chest related diseases such as Lung Cancer, Asthma, and Pneumonia etc.

Chronic kidney diseases have become a major public health problem. Chronic diseases are a leading cause of morbidity and mortality in India. Chronic kidney diseases account for 60% of all deaths worldwide. Eighty percentage of chronic disease deaths worldwide occur in low- and middle-income countries [2]. The National Kidney foundation determines the different stages of chronic kidney disease based on the presence of kidney damage and glomerular filtration rate (GFR), which is measure a level of kidney function. There are five stages of chronic kidney disease. They are shown in the Table.1.

The remaining paper is organized as follows: Section II discusses literature survey of the research. In section III describes the methodologies used for classifying chronic kidney disease are discussed. Section IV deals with the experiments and its results for parameter measures used in classification, and the result obtained in each classification algorithms. Section V describes the conclusion of the proposed work along with the feature enhancement.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

**Table. 1 DIFFERENT STAGES OF CHRONIC KIDNEY DISEASE**

| Stages of Chronic Kidney Disease | | |
|---|---|---|
| **Stage** | **Description** | **Glomerular Filtration Rate (GFR)*** |
| **1** | Kidney damage (e.g., protein in the urine) with normal GFR | 90 or above |
| **2** | Kidney damage with mild decrease in GFR | 60 to 89 |
| **3** | Moderate decrease in GFR | 30 to 59 |
| **4** | Severe reduction in GFR | 15 to 29 |
| **5** | Kidney failure | Lesser than 15 |

## II. LITERATURE SURVEY

Lambodar Jena et. al [3] presented prediction of chronic kidney disease using Naive Bayes, Multilayer Perceptron, Support Vector Machine, J48, Conjuctive Rule and Decision Table. The performance of these algorithms is measured by classification accuracy, time taken to build model, time taken to test the model, mean absolute error, Kappa statistics and ROC Area. From the experimental result, the Multilayer Perceptron algorithm gives better than the other five algorithms with the classification accuracy of 99.7%.

Abeer et. al [4] presented a new clinical decision support system for diagnosing patients with chronic renal failure. The aim of this work is to improving performance of a previously reported Chronic Renal Failure diagnosis system which was based on Artificial Neural Network, Decision Tree and Naïve Bayes. Data Mining Classifier like Support Vector Machine and Logistic Regression is to compare the performance of these algorithms on the basis of its diagnostic accuracy, sensitivity and specificity. From the experimental result it is observed that the performance of the Support Vector Machine is better than the other algorithm with the accuracy of 93.1%.

Jerlin Rubini et. al [5] proposed a new chronic kidney disease dataset with three classifiers such as radial basis function network, multilayer Perceptron, and logistic regression. The obtained result of this experiment shows in terms of prediction accuracy, error rate, sensitivity, specificity and F-score. Accuracy of these three classifiers is evaluated. The performance of multilayer Perceptron is better than the other two algorithms with the accuracy of 99.7%.

Ruey Kei Chiu et. al [6] developed a best-fitting neural network model to detect various severity levels of chronic kidney disease. Back-Propagation Network (BPN), Generalized Feed Forward Neural Networks (GRNN), and Modular Neural Network (MNN) are used. The performance of these algorithms is measured by accuracy, sensitivity, and specificity. From the experimental result, the Back-Propagation Network provides better than the other two algorithms with the accuracy of 94.7%.

Jayalakshmi et. al [7] discussed a research to increase the performance of the network in terms of accuracy. The accuracy was increased by using three key concepts: missing data replacement, data pre-processing and introducing the Performance Vector (PV) in the search direction. The results of the network have been tested using Pima Indian Diabetes Dataset. The experimental system improves the performance more than 7% than the standard Gradient Descent method. The accuracy rate of improved Gradient Descent method is above 99%.

Koushal Kumar et. al [8] compared the performance of all three neural networks on the basis of its accuracy, time taken to build model, and training data set size. Learning vector quantization (LVQ), two layers feed forward Perceptron trained with back propagation training algorithm and Radial basis function (RBF) networks for diagnosis of kidney stone disease. The multilayer Perceptron with two hidden layers and back propagation algorithm is the better model for diagnosis of kidney stone disease. Its accuracy is 92% to diagnosis the kidney stone disease.

Rajalakshmi et. al [9] presented an efficient and effective model of forecast and classification of functional abnormalities of kidney using Associative Neural Network (ASNN) and Polynomial Neural Network (PNN). From the experimental result, the Associative Neural Network is better than the other classifier algorithm with the squared correlation coefficient of 0.98%.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

Vijayarani et. al [10] presented to predict kidney diseases by using Support Vector Machine (SVM) and Artificial Neural Network (ANN). The aim of this work is to compare the performance of these two algorithms on the basis of its accuracy and execution time. From the experimental results it is observed that the performance of the ANN is better than the other algorithm with the accuracy of 87%.

Shweta Kharya et. al [11] described various review and technical articles on breast cancer diagnosis and prognosis. The current research is being carried out using the data mining techniques to enhance the breast cancer diagnosis and prognosis. Among the various data mining classifiers and soft computing approaches, Decision tree is found to be better predictor with 93.62% accuracy on benchmark dataset (UCI machine learning dataset) and also on SEER dataset.

Abhinandan Dubey et. al [12] discussed the automated detection of diseases using Machine Learning Techniques. The K-Means Clustering Algorithm with a single mean vector of centroids, to classify and make clusters of varying probability of likeliness of suspect being prone to CKD. The results are obtained from a Real Case Data-Set from UCI Machine Learning Repository.

## III. METHODOLOGY

1. *Dataset:*

The dataset for diagnosis of chronic kidney disease is obtained from medical reports of the patients collected from different laboratories in Coimbatore. There are 1000 instances with 15 different attributes related to kidney disease like PID (patients ID), Age, Gender, Weight, Serium-albumin, Serium- sodium, Blood urea nitrogen, Serium creatinine, Serium uric acid, Sodium urine, Urine urea nitrogen, Urine creatinine, Urine uric acid,egfr and Kidney failure. The main contributing attribute to identify the chronic kidney disease is EGFR. Based on the value of the EGFR, the instances are classified as Low, Mild, Moderate, Normal and Severe.

2. *Classification Using R:*

R tool is developed by academicians who persistently provide libraries for new and developing statistical techniques. It is also the name of a popular programming language used by a growing number of data analysts inside corporations and academia. The machine learning algorithms such as Back Propagation Neural Network, Radial Basis Function and Random Forest are implemented as classification task for accurate diagnosis of CKD based on different performance evaluation measures.

The machine learning algorithms are listed below:

1. Artificial Neural Network
2. Back Propagation Neural Network
3. Radial Basis Function
4. Random Forest

*Artificial Neural Network:*

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase [13].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

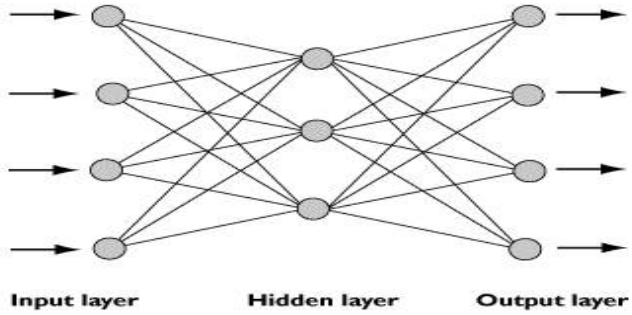


Fig .1 Artificial Neural Networks

Artificial neural network architecture has three layers in its structure. The three layers are input, hidden and output layer. The input layer corresponds to a single attribute. The output layers of a network contain the solution to a problem. The hidden layers are between the input and output neurons and correspondingly a large number of weights**.** The weights are the key elements in an artificial neural network. They express the relative strength of the input data. The summation function computes the weighted sums of all the input elements entering each processing element. A summation function multiplies each input value by its weight and totals the values for a weighted sum *Y*. The formula for n inputs is

$$Y = \sum_{i=1}^{n} X_i \quad w_i \quad \text{---------- eq. (1)}$$

For the j$^{th}$ neuron of several processing neurons in a layer the formula is

$$Y_j = \sum_{i=1}^{n} X_i \quad w_{ij} \text{ ---------- eq. (2)}$$

The summation function computes the internal stimulation. Based on this level, the neuron may or may not produce an output. The relationship between the internal activation level and the output can be linear or nonlinear. The relationship is expressed by one of several types of transformation (transfer) functions. The transformation (transfer) function combines (i.e., adds up) the inputs coming into a neuron from other neurons/ sources and then produces an output based on the choice of the transfer function. The sigmoid (logical activation) function (or sigmoid *transfer function*) is an S-shaped transfer function in the range of 0 to 1, and it is a popular as well as useful nonlinear transfer function:

$Y_T = 1/ (1 + e^{-Y})$ ---------- eq. (3)

Where $Y_T$ is the transformed (i.e., normalized) value of Y. A threshold value is a hurdle value for the output of a neuron to trigger the next level of neurons. If an output value is smaller than the threshold value, it will not be passed to the next level of neuron.

1. *Back Propagation Neural Network:*

Back propagation algorithm is a learning rule for multi-layered Neural Networks. Back Propagation networks are fully connected, layered, feed forward networks, in which activations flow from the input layer through the hidden layer(s) and then to the output layer. Back propagation uses supervised learning in which the network is trained using data for which inputs as well as desired outputs are known. In order to train a neural network to perform some task, the weight of each unit must be adjusted, in such a way that the error between the desired output and the actual output is reduced.

2. *Radial Basis Function Neural Network:*

Radial basis functions (RBF) networks are also feed forward, but have only one hidden layer. A RBF network has any number of inputs and has only one hidden layer with any number of units .It uses radial combination functions in the hidden layer, based on the squared Euclidean distance between the input vector and the weight vector and uses exponential activation functions in the hidden layer. It has any number of outputs with any activation function and has

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

connections between the input layer and the hidden layer, and between the hidden layer and the output layer. Gaussian RBF networks are said to be local-processing networks because the effect of a hidden unit is usually concentrated in a local area centered at the weight vector [5].

3. *Random Forest:*

Random Forest classifies all trees in the forest in the classification process by combination of the prediction of the tree structure each sampled according to the same distribution and the random vector values. It is an ensemble of decision trees based classifiers. Each tree is constructed by a bootstrap sample from the data. It uses both bagging and random variable selection for tree building. Once the forest is formed, test instances are percolated down each tree and trees make their respective class prediction. The error rate of a random forest depends on the strength of each tree and correlation between any two trees [14].

## IV. PSEUDO CODE

1. *Pseudo code for Back Propagation Neural Network:*

Step 1: Initialize the weights in the network (often randomly)
Step 2: Repeat
* For each example e in the training set do
1. O = neural-net-output (network, e); forward pass
2. T = output for e
3. Calculate error (T - O) at the output unit
4. Compute delta_wi for all weights from hidden layer to output layer; backward pass
5. Compute delta_wi for all weights from input layer to hidden layer; backward pass continued
6. Update the weights in the network

* End
Step 3: until all examples classified correctly
Step 4: return (network).

2. *Pseudo code for Radial Basis Function Neural Network:*

Step 1: Choose the number (m) and initial coordinates of the centers (R) of the RBF function.
Step 2: Choose the initial value of the spread parameter (σ) for each centre (R).
Step 3: Initialize the weights/coefficients (w) to small random values [-1, 1].
Step 4: For each epoch (e).
Step 5: For each input vector/pattern (x (p)).
Step 6: Calculate the output (y) of each output node (o).
Step 7: Update the network parameters (w, R, σ)
Step 8: end for (p = n)
Step 9: end for (e = total epochs)

3. *Pseudo code for Random Forest:*

Step 1: To generate c classifiers: for i = 1 to c do
Step 2: Randomly sample the training data D with replacement to produce i D
Step 3: Create a root node, i N containing i D Call Build Tree (i N) end for
Step 4: Build Tree (N): if N contains instances of only one class then return else
Step 5: Randomly select x% of the possible splitting features in N
Step 6: Select the feature F with the highest information gain to split on
Step 7: Create f child nodes of N , 1 N ,..., f N , where *F* has f possible values ( 1 F , … , f ) For *i* = 1 to *f* do
Set the contents of i N to *i* D, where *i* D is all instances in N that match *i F* Call Build Tree (i N*)*
End for, end if.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

## V. EXPERIMENTAL RESULT

### A. *Performance Evaluation:*

Diagnosis tests include different types of information, such as symptoms and medical tests. Doctor's conclusion of medical treatment rest on diagnosis tests which makes the accuracy of diagnosis is important in medical care. Providentially, the attributes of the diagnosis tests can be measured for a given disease condition. The best probable test can be chosen based on the attributes sensitivity, specificity, kappa and accuracy are widely used statistics to describe a diagnostic test.

*Kappa*: Kappa statistic is used to assess the accuracy of any particular measuring cases, it is usual to distinguish between the reliability of the data collected and their validity. Cohen's kappa measures the pact between two different measures which classify N items into C mutually exclusive categories.

$$K = \frac{Pr(a) - Pra(e)}{1 - Pr(e)} \quad \text{---------- eq. (4)}$$

Where; Pr (a) is observed agreement, and Pr (e) is the hypothetical probability of chance agreement.

*Sensitivity*: It also called the true positive rate, measures the proportion of positives which are correctly identified. The probability of positive test give that the patient is ill.

$$Sensitivity = \frac{TruePositives}{TruePositives + FalseNegatives} \quad \text{---------- eq. (5)}$$

*Specificity*: It also called the true negative, measures the proportion of negatives which are correctly identified. The probability of negative test given that the patient is well.

$$Specificity = \frac{TrueNegatives}{TrueNegatives + FalsePositives} \quad \text{---------- eq. (6)}$$

*Accuracy:* It is the amount of true results, one or the other true positive or true negative, in a population. It measures the degree of precision of a diagnostic test on a condition.

$$Accuracy = \frac{No.ofCorrectassesment}{No.ofallassesments} \quad \text{---------- eq. (7)}$$

### B. *Result:*

1. *Diagnosis Using Back Propagation Neural Network:*

Fig. 2 gives the result of back propagation neural network classifier. First, install the package back propagation neural network for 'nnet'. The dt is original dataset and cl is the class labels for selecting the databases. The class label name is kidney failure. It consists of a network training module (nnet ()) and a prediction module (predict). The dataset is trained back propagation and training labels with 15 classes. The performance of the classifier is evaluated in terms of kappa, sensitivity, specificity and accuracy. The predictive accuracy reported is 80.4%.

2. *Diagnosis Using Radial Basis Function Neural Network:*

Fig. 3 gives the result of radial basis function neural network classifier. First, install the package radial basis function neural network for 'Rsnss'. The dt is original dataset and cl is the class labels for selecting the databases. The class label name is kidney failure. It consists of a network training module (rbf ()) and a prediction module (predict). The dataset is trained rbf function and training labels with 15 classes. The performance of the classifier is evaluated in terms of kappa, sensitivity, specificity and accuracy. The predictive accuracy reported is 85.3%.

3. *Diagnosis Using Random Forest:*

Fig. 4 gives the result of random forest classifier. First, install the package random forest for 'randomForest'. The dt is original dataset and cl is the class labels for selecting the databases. The class label name is kidney failure. It consists of a network training module (rbf ()) and a prediction module (predict). The dataset is trained rf function and training labels with 15 classes. The performance of the classifier is evaluated in terms of kappa, sensitivity, specificity and accuracy. The predictive accuracy reported is 78.6%.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

**International Journal of Innovative Research in Computer and Communication Engineering**

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

5. *Comparative Analysis*

Table. 2 shows the performance evaluation measures like kappa, specificity, sensitivity and accuracy are given for each classifier in detail. Kappa Statistic is used to evaluate the accuracy of any particular measuring cases which is used to distinguish between the reliability of the data collected and their validity. The Kappa score for the Radial Basis Function is 0.8170. The Radial Basis Function provides better result. The Sensitivity and Specificity measures are used to calculate the true positive rate and true negative rate. The value of sensitivity for Back Propagation Neural Network it is 0.8353, Random Forest attained 0.8067 and Radial Basis Function obtained 0.8706. And the value for specificity for Back Propagation Neural Network is 0.8910, Random Forest obtained 0.8482 and Radial Basis Function gained 0.9233. Fig. 5 illustrates the chart comparison of different classifiers like back propagation, radial basis function and random forest with performance analysis of kappa, sensitivity specificity. Fig. 6 demonstrates the accuracy comparison of algorithms.

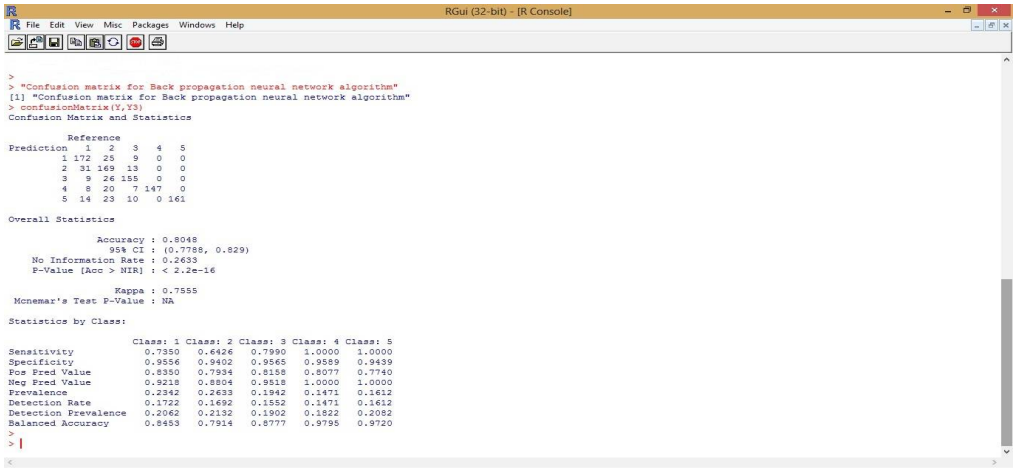


Fig. 2 Back Propagation Neural Network

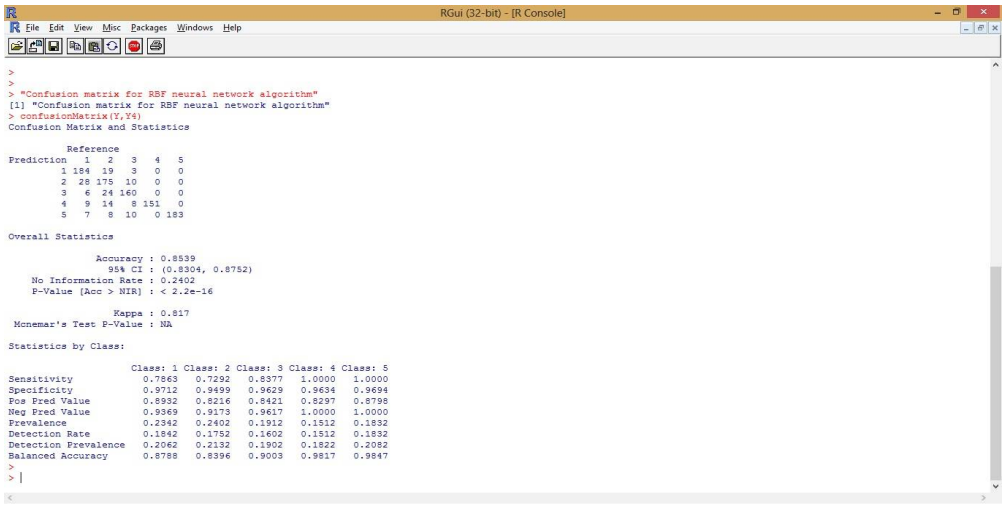


Fig. 3 Radial Basis Function

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

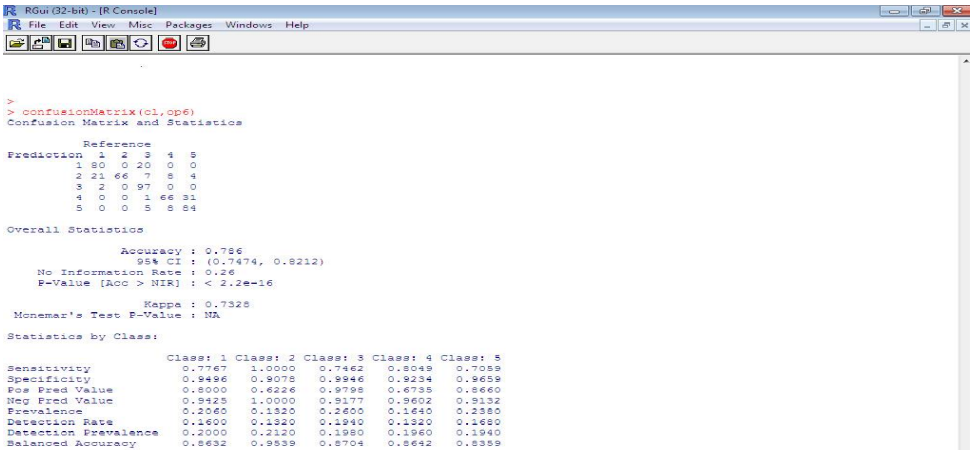


Fig. 4 Random Forest

TABLE. 2 PERFORMANCE MEASURES FOR VARIOUS CLASSIFIERS

| Classification / Performance Measure | BP | RBF | RF |
|---|---|---|---|
| **Kappa** | 0.75 | 0.81 | 0.73 |
| **Specificity** | 0.89 | 0.92 | 0.73 |
| **Sensitivity** | 0.83 | 0.87 | 0.80 |
| **Accuracy** | 80.4% | 85.3% | 78.6% |

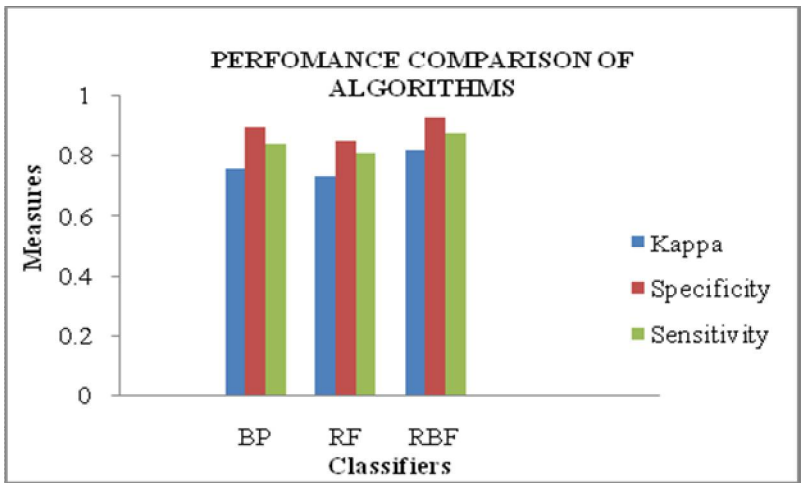


Fig. 5 Performance Comparison Chart

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 1, January 2016**

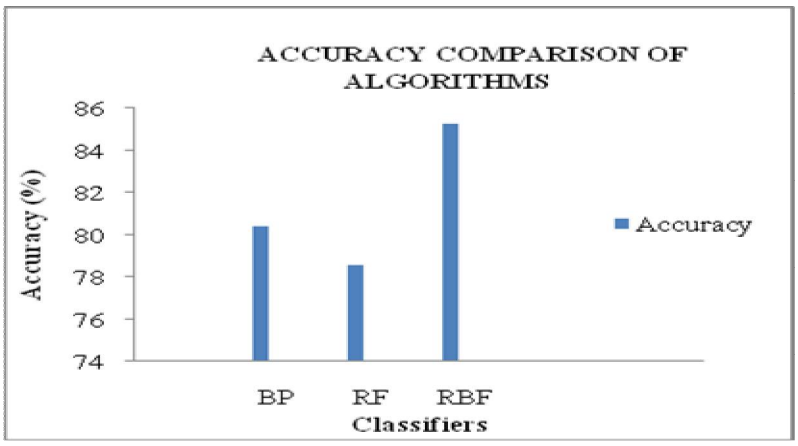


Fig. 6 Accuracy Comparison Chart

## VI. CONCLUSION

Prediction of chronic kidney disease is one of the essential topics in medical diagnosis. The proposed work is to classify the different stages of chronic kidney disease according to its severity. The classification algorithms that have been considered for predicting chronic kidney disease are Back propagation Neural Network, Radial Basis Function and Random Forest. The models are evaluated with four different measures like Kappa, Accuracy, Sensitivity and Specificity. From the experimental result, the Radial Basis Function is the better accuracy for predicting chronic kidney disease and it attains the accuracy of 85.3%.

## REFERENCES

1. Mohammed Abdul Khaleel and Sateesh Kumar Pradham, "A Survey of Data Mining Techniques on Medical Data for Finding Locally Frequent Diseases", International Journal of Advanced Research in Computer Science and Software Engineering, 2013, Vol.3, No. 8, pp. 149-153.
2. Veerappan, Ilangovan and Abraham Georgi," Chronic Kidney Disease: Current Status, Challenges and Management in India", Indian Journal of Public Health Research and Development, 2014, Vol.6, No.1, pp. 1694-1702.
3. Lambodar Jena and Narendra Ku. Kamila, "Distributed Data Mining Classification Algorithms for Prediction of Chronic Kidney Disease", International Journal of Emerging Research in Management and Technology, 2015, Vol. 4, No. 11, pp. 110-118.
4. Abeer, Ahmad and Majid," Diagnosis and Classification of Chronic Renal Failure Utilising Intelligent Data Mining Classifiers", International Journal of Information Technology and Web Engineering, 2014, Vol.9, No.4, pp. 1-12.
5. Jerlin Rubini and Eswaran," Generating Comparative Analysis of Early Stage Prediction of Chronic Kidney Disease", International Journal of Modern Engineering Research, 2015, Vol. 5, No. 7, pp. 49-55.
6. Ruey Kei Chiu and Renee Yu-Jing," Constructing Models for Chronic Kidney Disease Detection and Risk Estimation", Proceedings of 22nd IEEE International Symposium on Intelligent Control, Singapore, 2011, pp. 166-171.
7. Jayalakshmi and Santhakumaran, "Improved Gradient Descent Back Propagation Neural Networks for Diagnoses of Type II Diabetes Mellitus", Global Journal of Computer Science and Technology, 2010, Vol. 9, No. 5, pp. 94-97.
8. Koushal Kumar and Abhishek, "Artificial Neural Networks for Diagnosis of Kidney Stones Disease", International Journal Information Technology and Computer Science, 2012, Vol. 7, No. 3, pp. 20-25.
9. Rajalakshmi, Neelamegam and Bharathi, "Diagnosis and Classification of Level of Kidney Function Using Associative Neural Network and Polynomial Neural Network", Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2013, Vol. 4, No. 4, pp. 724-738.
10. Vijayarani and Dhayanand, "Kidney disease prediction using Support Vector Machine and Artificial Neural Network algorithms", International Journal of Computing and Business Research (IJCBR), 2015, Vol. 1, No. 3, pp.1765-1771.
11. Shweta Kharya, "Using data mining techniques for diagnosis and prognosis of cancer disease", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), 2012, Vol. 2, No. 2, pp. 55-66.
12. Abhinandan Dubey, "A Classification of CKD Cases Using Multi Variate K-Means Clustering", International Journal of Scientific and Research Publications, 2015, Vol. 5, No. 8, pp. 1-5.
13. Yashpal Singh and Alok Singh Chauhan," Neural Networks in Data Mining", Journal of Theoretical and Applied Information Technology, 2005, Vol.5, No.6, pp. 37-42.
14. Baylor, Konukseven and Koku," Control of a Differentially Driven Mobile Robot Using Radial Basis Function Based Neural Networks", World Scientific and Engineering Academy and Society Transcations on Systems and Control, 2008, Vol. 3, No. 12, pp. 1002-1013.
15. Shubham Bind, Arvind Kumar Tiwari and Anil Kumar Sahani, " A Survey of Machine Learning Based Approaches for Parkinson Disease Prediction", International Journal of Computer Science and Information Technologies, 2015,Vol. 6, No. 2, pp.1648-1655.