International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 23 (2017) pp. 13498-13505

Review of Chronic Kidney Disease based on Data Mining Techniques

S.Dilli Arasu
Research Scholar, Department of Computer Applications (Ph.D.),
Bharath Institute of Higher Education and Research (BIHER),
(Declared as Deemed-to-be-University under Section 3 of UGC Act 1956)
 # 173 Agharam Road, Selaiyur, Chennai - 600 073, Tamil Nadu, India.
Orcid Id: 0000-0002-4961-8130

Dr. R.Thirumalaiselvi
Research Supervisor / Assistant Professor, Department of Computer Science,
Government Arts College for Men (Autonomous)- Nandanam, Chennai -600 035, Tamil Nadu, India.

ABSTRACT

The Chronic Kidney disease is the most important health issues concerning the people as a whole. Chronic diseases lead to morbidity and increase of death rates in India and other low and middle income countries. The chronic diseases account to about 60% of all deaths worldwide. 80% of chronic disease deaths worldwide also occur in low and middle income countries. In India, probably the number of deaths due to chronic disease found to be 5.21 million in 2008 and seems to be raised to 7.63 million in 2020 approximately 66.7%. Data mining is the process of extraction of hidden information from the large dataset. Major data mining techniques such as clustering, classification, association analysis, regression, time series and sequence analyses were used to predict kidney diseases. The techniques that were introduced so far had minor drawbacks in the quality of preprocessing or at any other stages. In this paper, the various data mining techniques are surveyed to predict kidney diseases and major problems are briefly explained.

Keywords: Chronic Kidney Disease (CKD), Risk factors of CKD, Challenging issues of CKD, Data mining and Machine Learning (ML) algorithms.

INTRODUCTION

The massive advancement due to the amount of biological data available has raised a gloomy question of being classified, managed effectively and to be transfer the raw data to meaningful information. The emergence of this colossal amount of calls into question the pattern of modern computation. It can be addressed using data mining algorithms. Machine Learning surely stands to capture a major fraction of the problem and thus accounts for the latest progress in the field of bioinformatics, computational biology and application of machine learning methods on prominent problems in human biology and behavior. The algorithms and mathematical techniques allow us to go beyond a mere depiction of the data and makes logical results in the form of mathematically testable models. The notions of supervised and unsupervised learning make this process easy and comprehensible. By simplifying abstraction the statistical predictions of a system is obtained which constitutes a model.

As people's day to day life become more and more modernized and extended life span in the society, Chronic Kidney Disease (CKD) also found common and result in degradation in the functionalities of kidney function. Once any person gets CKD, they may suffer from the disease which may decrease their working capability as well as living quality. It is also rapidly results in other chronic diseases like high blood pressure, anemia, weak bones due to poor nutritional health and nerve damage. In the meantime, kidney disease maximises the patient risk of contracting heart and blood oriented diseases. Chronic kidney disease even causes other chronic disease such as diabetes, high blood pressure and other disorders. High risk groups are classified as person with diabetes, hypertension, and hereditary. It is possible to get rid of chronic kidney disease through early detection and proper treatment once the progress of the disease is observed it may greatly leads to kidney failure.

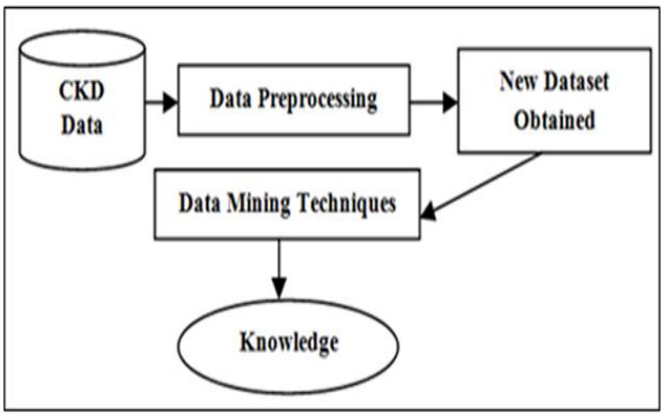

Figure 1: Overview of Research Work

Figure 1 above specifies the overview of the research work. CKD dataset is taken from the UCI Repository. Data preprocessing is performed to fill the missing values, noise

removal, data cleaning, etc. At last, a new dataset will be obtained. Data Mining techniques are applied on it to evaluate the performance in case of predicting whether the person has kidney diseases or not. Kidney disease is a growing problem. For most people, kidney damage occurs slowly over many years, often due to diabetes or high blood pressure. This is called chronic kidney disease. When someone has a sudden change in kidney function because of illness, or injury, or has taken certain medications this is called acute kidney injury. This can occur in a person with normal kidneys or in someone who already has kidney problems. The main risk factors for developing kidney disease are: Diabetes, High blood pressure, cardiovascular (heart and blood vessel) disease, and family history of kidney failure.

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 (National Kidney Foundation, 2002). 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. The health care dataset contains missing values. To address this problem, pre processing techniques will be used in healthcare datasets. These missing values can degrade the performance of abnormality detection. Several methods were proposed to fill up these missing values. An existing classification framework used a data preprocessing method where data cleaning method was used to fill in the missing values and to correct the erroneous ones.

A recalculation process is performed on the chronic Kidney disease (CKD) stages and the values were recalculated and filled in for unknown values. Though this method is efficient, the influence of expert in the field of healthcare dataset values for CKD is needed. So to avoid this need and improve the preprocessing as a layman, an ensemble-learning based sparse-data modeling framework is proposed. The dataset transformation is performed to convert the nominal features to numerical features of the data set to make it support the generation of subsets. That is positive responses, such as yes, good, and present, were defined as 1, whereas those negative responses such as no, poor and not present, were defined as 0.

The proposed work generates subsets of the data, and then it trains models using subsets and then combines the candidate models into an ensemble learner for making predictions and also to generate a score for the importance of missing values in each feature in the dataset. Figure 2 illustrates the classification of kidney disease. It is of two type namely acute kidney disease and chronic kidney disease. Here in our research work we concentrate on factors supporting chronic kidney diseases respectively.

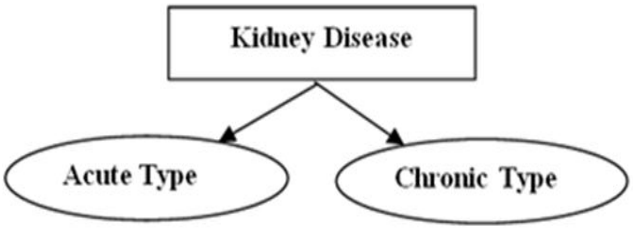

Figure 2: Classification of Kidney Disease

LITERATURE REVIEW

To measure the different methods of CKD the influence factors are used as the inputs for network modeling. Next, the detection of CKD classification and risk evaluation are evaluated. The three different neural network models are employed for detection. Kidney disease influences the kidneys damage and fails to filter blood. This damage can cause wastes to remain in the body.

The majority people, kidney damage occurs slowly over years, due to diabetes or high blood pressure called as chronic kidney disease. When anyone feels a sudden change in kidney function due to illness, injury, has taken certain medications is called acute kidney injury. This can occur in a person with normal kidneys or in anyone who already suffers from kidney problems. It seems to be a growing problem. 10% of the population in worldwide is affected by this disease.

Due to CKD millions of people die each year because they do not undergo proper treatment. The Global Burden of Disease study held in 2010 determines that chronic kidney disease ranked 27th position in list of causing total number of deaths worldwide around 1990 and can be rises to 18th position in 2010. Risk factors of kidney disease are smoking, obesity, high cholesterol, diabetes (types I and II), autoimmune disease, obstructive kidney disease, bladder obstruction caused by Benign Prostatic Hyperplasia, atherosclerosis, cirrhosis and liver failure, narrowing of the artery, kidney cancer, cancer occurring in bladder, kidney stone and related infections (Kirubha & Manju Priya, 2016). Kidney diseases are predicted and compared using SVM and Artificial Neural Network algorithm based on the accuracy and execution time.

Result shows that ANN outperforms with reduced execution time (Vijayarani & Dhayanand, 2015). CKD dataset is taken from UCI repository and performance are evaluated using the algorithms such as Naive Bayes, Multilayer Perceptron, SVM, J48, Conjunctive rule and Decision tree. It shows that multilayer perceptron algorithm gives better classification accuracy rate in prediction of chronic kidney diseases (Lambodar Jena & Narendra Kamila, 2015). K-Means Clustering Algorithm along with a single mean vector of centroids have been formulated to classify the clusters of varying probability of likeliness suffers from CKD. The results are obtained from a real case dataset (UCI Repository) to show the probability of disease causing factors

(Abhinandan Dubey, 2015). Machine learning algorithms like AD Trees, J48, KStar, Naive Bayes, Random Forest algorithms are used to predict the kidney disease. The performance result of the Naive Bayes shows better accuracy rate compared to other algorithms (Swathi Baby & Panduranga, 2015).

Evolution of big data in healthcare field is evaluated by Support Vector Machine, Decision Tree and Bayesian Network machine learning algorithms (Basma Boukenze, et al., 2016). Chronic Kidney Disease dataset is used to predict patients with chronic kidney failure and normal person. C4.5 algorithm provides better results with less execution time and accuracy rate. Performances are judged by Basic concepts of Decision Tree, Bayes classification, Rule based classification, Back Propagation, Support Vector Machine and K-Nearest Neighbour algorithms. To address the same problem, classifiers namely Multilayer Perceptron, Random forest, Naive Bayes, SVM, K-Nearest Neighbour and Radial Basis Function are also involved (Pushpa M. Patil, 2016).

Three different neural network models have been implemented for chronic kidney disease prediction which includes back propagation neural network, generalized feed forward neural network and modular neural network. This research shows that all these models influences genetic algorithm in to their respective neural factor. All three models give better accuracy more than 85%. Compared to other models, back propagation neural network has the highest accuracy (Ruey Kei Chiu & Renee Yu-Jing, 2011). Chronic kidney disease are predicted using six different data mining algorithms like Random Forest Classifiers, Sequential Minimal Optimization, Naive Bayes, Radial Basis Function, Multilayer perceptron classifier and Simple Logistic.

Totally, 400 records are used for training set to perform prediction. Among these, Random Forest outperforms well. Real time collected from Apollo Hospital dataset involves many imbalanced data. To address this issue a rebalancing algorithm called SMOTE has been presented to predict kidney disease of patients (Sai Prasad Potharaju & Sreedevi, 2016). To improve the accuracy of prediction result the classification algorithms like Naive Bayes and Decision tree algorithms are compared. But the accuracy of Decision tree classification is seems to be 91%. (Anantha Padmanaban & Parthiban, 2016).

RISK FACTORS OF CKD

Susceptibility Factors

It results in increase susceptibility to kidney damage. These symptoms are seen in aged people, hereditary facts; reduce kidney mass, less weight baby, low income and educational levels.

Initiation Factors

It specifies the factors that directly involves in kidney damage. Its symptoms are diabetes mellitus, high blood pressure, autoimmune diseases, systemic (affecting the whole body) infections, urinary tract infections, urinary stone, obstacle of lower urinary tract, toxic drugs.

Progression Factors

It leads to worse fact of kidney damage and rapid decline functionalities once the damage gets started. The symptoms are found to be high level proteinuria, high BP, less glycemic control due to diabetes and smoking.

End- Stage Factors

Increase in morbidity and mortality due to kidney failure results in lower dialysis dose, temporary vascular access, anemia, less serum albumin level and late referral dialysis

The following table 1 shows the stages of CKD with GFR rate and action plan that must be taken eventually at each stage.

Table 1: Classification of Chronic Kidney Disease and Evaluation Plans

S.No	Stages of CKD	Glomerular Filtration Rate (GFR)	Action Plans
1	Kidney damage with normal GFR	90 or above	Diagnosis and treatment of comorbid conditions, disease progression, reduction of risk factors for cardiovascular disease
2	Kidney damage with mild decrease	60 to 89	Estimation of disease progression
3	Moderate decrease	30 to 59	Evaluation and treatment of disease complications
4	Severe reduction	15 to 29	Preparation of kidney replacement treatment (dialysis, transplantation)
5	Kidney failure	Less than 15	Kidney replacement therapy

Challenges In Chronic Kidney Disease Management In India

The average global occurrence values for treating End Stage Renal Disease (ESRD), dialysis and transplant patients were found to be 280, 215 and 65 patients per million (ppm) respectively.

In India, the average occurrence values for treated ESRD (not diagnosed), dialysis and transplant patients were 70, 60 and 10 ppm respectively. This global increasing in number is 7% every year. It is estimated that only 10 to20% of ESRD patients in India undergoes long-term renal replacement therapy.

It is estimated that in India within one year, there are 3,500 renal transplant, 3,000 Continuous Ambulatory Peritoneal Dialysis initiation and 15,000 Maintenance Hemo Dialysis patients (Agarwal & Srivastava, 2009).

Data Mining Task Primitives

A mining task can be specified as query used to input the data mining system. A data mining query implies the data mining task primitives. To allow the user, interactively communicate with the data mining system to generate the mining process, or examine the findings from different angles or depths. The required data can be mined using the database or the set of data in which the user is interested.

It includes the database attributes or data warehouse dimensions of interest. Knowledge can be mined using characterization, discrimination, association or correlation analysis, classification, prediction, clustering, outlier analysis, or evolution analysis. The background knowledge is used in the discovery process for evaluating the patterns found. Concept hierarchies are a popular form of background knowledge, which allow mined data at multiple levels of abstraction. This interestingness measures and thresholds for pattern evaluation guide the mining process to evaluate the discovered patterns. Different kinds of knowledge may have different interestingness measures like support and confidence for association rules (Charu).

Advantages And Disadvantages Of Data Mining In Medical Field

Advantages

- Data mining in healthcare detects fraud and abuse.
- Help physicians to identify effective treatments and best practices.
- Patients exploit better and greater affordable healthcare services.
- Increases in the speed of working with large datasets and rapid report generation, faster analysis, improved operational efficiency and reduced operating cost.
- Data Mining can extract predictive knowledge from large databases.

Disadvantages

- Heterogeneous medical complex physician's interpretation with poor mathematical classification.
- Ethical, Legal and Social Issues.
- Data Ownership issues.
- Privacy and Security related to Human Data Administration.
- It Involves privacy issues and security issues and
- Misuse or incorrect information.

Data Mining Algorithms And Techniques

A data mining algorithm is a well define procedure that takes data as input and generates as output. It involves patterns in the form of models.

It encompasses several algorithms and techniques like Classification, Clustering, Prediction, Association Rules, Neural Networks etc., to perform knowledge discovery from databases.

Classification

Classification is the simplest and one of the popular data mining techniques. Where objects are divided and assigned to the other groups called classes. Each object has to be dispersed exactly to one class and not more than one and never to no classes at all.

The classification algorithms are,

Naive Bayes Algorithm

The Naive Bayes algorithm combines the prior probability and conditional probabilities in a single formula, so that it can be used to calculate the probability of each of the possible classifications in turn.

To make this, choose the classification with the largest value from a given a set of k mutually exclusive and exhaustive classifications with prior probabilities and n attributes followed by values of the instance.

The posterior probability of class that occurs for the specified instance can be shown to be proportional to prior probabilities along with respective values. Based on the assumption that if the attributes are independent, the value of expression can be calculated using the product by calculating this product for each value of from 1 to k the classification with largest value can be chosen.

Nearest Neighbor

It is the kind of distance based algorithm used mainly when all attribute values are continuous, it may be modified according to categorical attributes.

To estimate the classification of an unnoticed instance using the classification of the instance or instances that are closest to it. Even more instances in the training set applies the same principle to classify the k nearest neighbour or most nearest one known as k- Nearest Neighbour.

Decision Tree Based Algorithm

This approach is most useful to solve classification problems where a tree is constructed to model the classification process.

As soon as the tree is built, it is applied to each tuple in the database to generate classification.

Clustering

It examines data to find groups of items that are similar to each other as partition-based, hierarchical and agglomerative methods.

Partition-Based Method

This method examines the overall grouping of the input database objects into k partitions as cluster. Clusters are created by optimizing a parameter to calculate cluster distance based on similarity measure.

It enables similar objects placed within a cluster and unrelated objects are grouped under different clusters. It involves two types of clustering namely,

K-Means Clustering

This is a centroid-based approach that takes the number of partitions k as input and creates k clusters of the input database consist of n objects or records by optimizing the rule of clustering.

The resulting clusters with high similarity kept as intra-cluster and low similarity as inter-cluster. In this algorithm primarily a random set of k objects are chosen as the cluster centers to compute the mean value representing the cluster mean or centre.

Each object can be involved in the cluster to exhibits high similarity. Updating of new cluster is made till there is no change found in the cluster structure.

K-Medoids Clustering

It eliminates the sensitivity using medoid as a measure for similarity by choosing the mean value of objects in a cluster, the most centrally located object within a cluster is chosen.

This position may be changed while maintaining the distance-based similarity measure. Randomly k objects are chosen as medoid and remaining objects are assigned to the cluster based on the medoid distance and higher iterations reveal the cluster mediods.

Hierarchical Clustering Algorithm

Hierarchical clustering technique focuses on the principle of decomposing databases either in a top-down or bottom-up fashion.

It is divided into divisive and agglomerative techniques.

Agglomerative Clustering

Each individual object is treated as a cluster. Then, clusters are merged according to the similarity of objects until a single cluster is formed.

Finest case of clustering is made when all objects in the database influences same cluster type or when the terminate condition is reached.

Divisive

It operates in the reverse direction or bottom-up and considers the whole database as a single cluster.

Clusters splits based on similarity and dissimilarity measures, until each object formed as individual cluster or a terminating condition is reached (Margaret H. Dunham & Sridhar, 2006).

Prediction

This technique predicts or guesses the data values for attributes as missing or dislocate.

Linear Regression-Based Prediction

Linear regression is the simplest among the different types of regression where data are modeled by straight lines.

Data being modeled by straight lines could be interpreted as the growth of the data being in a linear form or following a straight line. Regression coefficients are estimated using the least squares method.

Association Rules

An association rule is basically an expression of the form X → Y where X and Y are item-sets.

Association rule mining is controlled by Support and Confidence.

Support gains the statistical signification of a rule and confidence forms the degree of certainty for detecting the associations.

Apriori Algorithm

Apriori is a type of candidate generation algorithm proceeds in a level-wise order. The apriori algorithm follows the join and prune steps. Join step constructs new candidate sets and prune step helps in filtering out candidate item-sets based on the anti-monotonic property.

Neural Networks

A neural network forms an interconnected group of artificial neurons to processes information and computation is done using associated weights. These weights are updated or adjusted during prediction of input records.

An ANN is a kind of adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks posse's non-linear statistical data modeling tools used to model complex relationships between inputs and outputs. It is the most prestigious model used in any application models to facilitate artificial intelligence.

Propagation

Propagation is a kind of processing approach where tuple of values are given as input to the neural network, at each node in the input layer.

Then the summation and activation functions are applied simultaneously at each node, an output value created from each output arc of the node in turn sent to the subsequent nodes until a tuple of output values are reached from the nodes of the output layer.

Back Propagation

Back Propagation is a learning technique that adjusts weights through propagating weight that changes backward from the sink to the source nodes in neural network structure.

It is purely a generalized delta rule approach known as a feed-forward back propagation network. It constitutes a supervising learning Multi Layer Perceptron model.

A training set of input patterns is applied to the network to computes the output pattern, and if any error occurs, the difference between actual and desired output patterns is taken into account by adjusting its weight. Since the real uniqueness or effectiveness of the network exists in the values of the weights between neurons, it is necessary to adjust the weights among them.

Radial Basis Function Network

A radial basis function is a type of functions whose value decreases or increases from a central point at a proper the distance.

The Gaussian activation function in an RBF network is typically a neural network with three layers. The input layer is used to input the data, Gaussian activation function is specified at the hidden layer and linear activation function is involved in the output layer.

Perceptrons

The simplest neural network is called a perceptron. It is a single neuron with multiple inputs and one output developed to promote activation.

Perceptron can be classified into two classes using a unipolar activation function, where output 1 can be used to classify into one class, while an output 0 can be used to pass in the other class. The simplest feed forward can neural network is called a multilayer perceptron that forms the network of perceptrons

Discussions

Data Mining gains greater importance in extracting hidden information from large set of database. Though we need a technique to intelligently extract hidden information data mining techniques is not seems to be straight forward. To apply data mining techniques it is significant to understand the nature of data. The most important step involves preprocessing of data, if the pre-processing is not carried out correctly then the entire decision making process may go worse (Jiawei Han & Micheline Kamber, 2003).
It involves cleaning, integration and transformation of data into understandable format and reducing the size to improve the quality of pattern found. In the above section, we have discussed about the different data mining techniques, various tasks that are used for the diagnosis of diseases and among these, classification is the most widely used algorithm as it has many advantages while comparing to other algorithms and yields better classification results.

The advantages and disadvantages of data mining techniques have been discussed above. Future research paper addresses the issue evolved by keeping preprocessing as a layman, an ensemble-learning based sparse-data modeling framework is proposed. The dataset transformation is performed to convert the nominal features to numerical features of the data set to make it support the generation of subsets.

That is positive responses, such as yes, good, and present, were defined as 1, whereas those negative responses such as no, poor and not present, were defined as 0. The proposed work generates subsets of the data, and then it trains models using subsets and then combines the candidate models into an ensemble learner for making predictions and also to generate a score for the importance of missing values in each feature in the dataset.

Regarding to the measurement of serious levels of CKD, presently glomerular filtration rate is the most commonly measuring factor used in healthcare field to estimate kidney function. The calculation of GFR is made by physician in the healthcare field through patient's blood creatinine, age, race, gender and other factors depending upon the type of formal-recognized computation formulas is employed. The GFR indicate how health of a patient's kidney regarding the disease issue. The physician checks these measures to determine the different stage of kidney disease of a patient (Firman, 2009; Manish Kumar, 2016).

CONCLUSIONS

This paper aims to analyze the various data mining techniques in medical domain and some of the algorithms used to predict kidney diseases eventually. From the above survey, it is proven that results may vary for different stages of kidney disease diagnosis based on the tools and techniques used. Data mining provides better results in disease diagnosis when appropriate techniques used. Thus, data mining is the significant field for healthcare predictions.

REFERENCES

[1] Abhinandan Dubey. (2015). A Classification of CKD Cases Using Multivariate K-Means Clustering. International Journal of Scientific and Research Publications, 5(8), 1-5.

[2] Agarwal, S.K., & Srivastava, R.K. (2009). Chronic kidney disease in India: challenges and solutions. Nephron Clin Practice,; 111(3):197-203.

[3] Anantha Padmanaban K.R., & Parthiban, G. (2016). Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease. Indian Journal of Science and Technology, 9(29), 1-5.

[4] Basma Boukenze, Hajar Mousannif, Abdelkrim Haqiq. Performance of Data Mining Techniques to Predict in Healthcare Case Study: Chronic Kidney Failure Disease. International Journal of Database Management Systems, 2016; 8(3), 1-9.

[5] Charu C. Aggarwal, (2015). Data Mining, Springer Cham Heidelberg New York: Springer International publishing Switzerland.

[6] Firman, G. (2009). Definition and Stages of Chronic Kidney Disease, Released on Feb. 28, http://www.medicalcriteria.com/site/index.php?option=com_content&view=article&id=142%3Anefckd&catid=63%3Anephrology&Itemid=80&lang=en.

[7] Jiawei Han & Micheline Kamber, (2003). Data Mining, concepts and Techniques, Second (Eds.), Elsevier Publication.

[8] Kirubha, V. & Manju Priya, S. (2016). Survey on Data Mining Algorithms in Disease Prediction. International Journal of Computer Trends and Technology, 38(3), 24-128.

[9] Lambodar Jena & Narendra Ku.Kamila. (2015). Distributed Data Mining Classification Algorithms for Prediction of Chronic- Kidney-Disease. International Journal of Emerging Research in Management &Technology, 4(11), 110-118.

[10] Manish Kumar. (2016). Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm. International Journal of Computer Science and Mobile Computing, 5(2), 24-33.

[11] Margaret H. Dunham & Sridhar, S. (2006). Data Mining: Introductory and Advanced Topics, Pearson Education, Dorling Kindersley (India) Pvt Ltd.

[12] National Kidney Foundation (NKF), (2002). Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification and Stratification. American Journal of Kidney Disease, 39: S1-S266.

[13] Pavithra, N. & Shanmugavadivu, R. (2016). Survey on Data mining Techniques used in Kidney related Diseases. International Journal of Modern Computer Science, 4(4), 178-182.

[14] Pushpa M. Patil. (2016). Review on Prediction of Chronic Kidney Disease Using Data Mining Techniques. International Journal of Computer Science and Mobile Computing, 5(5), 135-141.

[15] Ruey Kei Chiu., Renee Yu-Jing, (2011). Constructing Models for Chronic Kidney Disease Detection and Risk Estimation. Proceedings of 22nd IEEE International Symposium on Intelligent Control, Singapore, 166-171.

[16] Sai Prasad Potharaju & Sreedevi M. (2016). An Improved Prediction of Kidney Disease using SMOTE. Indian Journal of Science and Technology, 9(31), 1-7.

[17] Swathi Baby, P. & Panduranga, T. (2015). Vital, Statistical Analysis and Predicting Kidney Disease Using Machine Learning Algorithms. International Journal of Engineering Research and Technology, 4(07), 206-210.

[18] Vijayarani, S. & Dhayanand, S. (2015). Kidney disease Prediction Using SVM and ANN Algorithms. International Journal of Computing and Business Research, 6(2), 1-12.