Radial Basis Function Neural Network Based Predictive Model with Classifier for Kidney Failure in E-Health

S. Ancy[a] and K. Cornelius[b]

[a]Assistant Professor, Department of IT, Jeppiaar Institute of Technology, Chennai, India. Email: ancys2014@gmail.com
[b]Assistant Professor, Department of IT, St. Peters Engineering College, Chennai, India. Email: cornelius851@gmail.com

Abstract: Big data is the large set of dataset. It involves extraction, selection, analyzing and interpolation of data. Big data is used wide assortment in medical fields for analyzing the patient's medical history, prediction of future effects and clinical decision making. It can also be used as a tool to store large number of data. It helps us to understand the diseases and also paves way to predict the disease and its future effects caused by the disease. In this paper we use RBFNN (Radial Basis Function Neural Network) with classifier algorithm with the use of parameters to determine the condition of a patient as a normal or a kidney failure patient. The proposed method reveals the stages of the kidney failure patient and treatment and clinical decision.

Keywords: Big data, Radial basis function neural network, kidney failure, classifier algorithm, clinical decision.

1. INTRODUCTION

The upcoming Big Thing in the world of IT will be Big Data. The term Big data is used to describe both structured and unstructured that are of vast volume of data that deluges a business on a day-to-day basis. Like many technologies in IT, big data phases intense reductions in cost, extensive enhancements in the time complexity for a task to be computed, or for offering a good service and launching a new product. Big Data generates meaning from the stowage and dealing out of huge quantities of digital material that cannot be examined with old-style calculating technique. The big data faces are volume, velocity and variety. Big data is different from usual computing that is robotically generated by a device rather they are typically an entirely new data source. It is not designed as friendly, it may not have much values. Instead it is used to examine huge amount of data, appropriate information, identification of unknown correlations, hidden patterns, competitive advantage, effective marketing, customer satisfaction, better business decisions: strategic and operational, increased revenue. Data can be taken from any source and can be analyzed to enable with the features of cost reductions, reductions in time, development of new product and optimized offerings and smart decision making. Data set in digital form is the data that are used in the analysis in healthcare. These data are enormous and intricate, they are hard to succeed with usual traditional S/W and/or H/W; data management tools are also finding difficult in these data

to manage these and their methods[16], Big data analytics tenders, in healthcare yield gain of the burst in data to extract intuitions for the best informed conclusions [17-19] and as a category of research. Healthcare of Big data is devastating not only because of its size but also because of the variety of data and the along with the speed at which it must be treated and fared [16]. Analytics in this bigdata has the latent to improve attention, safe guard lives, minor budgets. This analytics is too proactively for one who would benefit from preventing the tricky and care for the advancement or changes in routine; broad scale diseases which are outlined as broad scale to identify foretelling events. It also support prevention initiatives; this is done by assembling, publishing the data that are based on medical procedures.

Kidney failure is the silent death and non communicable disease. Physicists, chemists and clinicians works to foresee the kidney's working that gives doctors a deep intervene for saving the life of the patient and helps to recover the patient. It's the measuring of GDP in the economy: to understand the dynamics of the system that is measuring a large quantity of compound activity is needed. Development and validation can be done using the prediction models, using demographic, medical, and data of the patient form the laboratory. Clinical and nonclinical data sources, to develop predictive models (PMs) identify patients with a high probability of not. This leads to the conclusion of a number that ranges in the sequential scale of 0 to 1. The result is thus used for finding the likelihood of the organ's success or failure of any patient at any given time.

Big data affects organizations across practically every industry.

- *Health Care:* Managing the records of patients like plans for treatment, information about prescription will surely bring out the secret insights which will be useful to improve the health of the patient. Traditional databases cannot handle these details as these are not as small as to be maintained by normal DB. On time processing, accuracy and transparency are the tough regulations of the industry in health care is concerned.

- *Education:* The data of any students that paves the way for the insight can make a noteworthy development in the system of education, students and also the curriculums. By analyzing these student's big data, they can identify unprotected students and making sure these students are making appropriate progress

- *Government:* Management of the utilities, running agencies, congestion problem in traffic or prevention of crime are able to be done by applying big data in the ground of government. This leads to significant growth. Thought there is very much advantage in using big data in government they the problem is to maintain the privacy and transparency of the data.

- *Manufacturing:* Big data analyzing in manufacturing sector paves way for the manufacturers to improve the quality of the product and also the waste produced during manufacturing can be reduced. Waste management is the biggest problem in manufacturing sector. Most of the manufacturers are indulged in analyzing and processing the huge data in industry so as to make fast and more responsive decisions in business.

- *Banking:* The sources of data in banking sector with large amounts of information is countless. Most of the data in banking sector are live or data that are streaming, banks has to find new and advanced ways to manage big data. The most important aspect is to understand the customers and increasing their satisfaction, minimizing the risk and detecting the fraud is it's equally important. Regulatory compliance maintenance is the base for the growth of the bank.

- *Retail:* Building the customer relationship is acute to the retail industry – and the best way to manage is to use big data technology. The Retailer has to understand the customers to enable best marketing.

The big data brings the most effective way to handle transactions. It also takes care of the most strategic way to bring back lapse business. Big data remains at the heart of all those things.

- *Fraud management:* The productivity is increased by detecting fraud earlier and improving prevention, to uncover the hidden trends and regulate and describe the patterns in large amounts of data. This is used to detect fraud in almost real time and also reduces false positives – and increase productivity. The most powerful simulation of features that is used to quickly adapt to the evolving patterns of fraud. This enhances the prevention with improved fraud strategy management. And also Reduces the risk of fraud with advanced analytic abilities and greater perceptibility

- *Demand signal management:* To become a truly claim-driven organization one have to manage the demand signal management software powered by our lightning-quick in-memory platform. External market and the data of the retailer as to be captured in real time. This is to be combined with the data of the internal business and state-of-the-art analytics. hence to nous, weigh, and rejoin to demand signals quicker than ever before

2. LITERATURE SURVEY

1. N. Pombo, P. Araujo and J. Viana "Knowledge discovery in clinical decision support system for pain management: A Systematic review "Artificial intelligence in Medicine, Vol. 60, No. 1, pp. 1-11, 2014.

   Clinical decision support systems (CDSSs) can be critical for improving the accuracy of pain assessment and offering better support for clinical decision-making. This review is focused on computer technologies for pain management that allow CDSSs to obtain knowledge from the clinical data produced by either patients or health care professionals. It is focused on computerized systems that constituted CDSSs and include data or results related to pain symptoms from patients with acute or chronic pain. The algorithm used is rule-based algorithms, ANN, nonstandard set theory, and statistical learning algorithms Computer technologies that are applied in CDSSs are important but not determinant in improving the systems' accuracy and the clinical practice, as evidenced by the moderate correlation among the studies. However, these systems are the base in the design of computerized systems oriented to a patient's symptoms as is required for pain management. Due to the lack of integration with mobile devices, the data are to be manually entered.

2. L. Kononenko, "Machine learning for medical diagnosis: history, state of art and perspective," Artifical Intelligence in medicine" Vol. 23, No. 1, pp. 89-109, 2001.

   Machine learning perspective and the development of intelligent data analysis is the basic reason for artificial Intelligence entering the medical field. The Algorithm used in this literature paper are the naïve Bayesian classifier, decision trees and neural networks. Naïve Bayesian Classifier is a recently developed method which deals with the reliability of decisions of classifiers. This seemed to be the most promising intelligent data analysis in medicine. It is one of the suitable approach to include machine learning for verifying the mysterious phenomena from the complementary medicine in diagnosing the future medical disease and its treatments.

3. A. Azemi, V.R. Sabzevari, M. Khademi, H. Gholizade, A. Kiani, Z.S. Dastgherib "Intelligent arrhythmia detection and classification using ICA" in engineering in medicine and biology society, 2006 EMBS 28th Annual International conference of IEEE, 2006, pp. 2163-2166.

   In this proposed literature, RR-interval signals are obtained from the ECG records. Detection and Classification of arrhythmic beat are done using RR-signals relying on the knowledge based

methodology. An algorithm is constructed for classification of arrhythmic beat which uses a three RR-interval sliding window method. There are four categories of beats which are categorized as normal, premature ventricular contractions, ventricular flutter/fibrillation and 2 degrees heart block. The input of a knowledge-based deterministic automaton is the beat classification. This is used to achieve arrhythmic episode detection and couplet, ventricular tachycardia, ventricular flutter/fibrillation and 2 degrees heart block. The MIT-BIH arrhythmia database is used for the evaluation of this method. It method produces a high show of 98% correctness for arrhythmic beat classification. For arrhythmic episode detection and classification the rate of accuracy is sustained as 94%. The ultimate way of classification for arrhythmia beat and episode classification using the RR-interval signal is the only constraint.

4.   H.L. Lu, KONG and P. CHIA "An automated ECG classification system based on neuro - fuzzy system", in computer in Cardiology 2000, 2000, pp. 387-390.

In this literature work, 12-lead Electrocardiogram (ECG) ----a hybrid neuro-fuzzy approach is used in the diagnosis of Acute Myocardial Infarction (AMI) at its early stage. This is a new approach presented to the diagnostic classification of 12-lead ECG. This design is a combination of fuzzy logic and neural network theory and it consists of two phases: (a) Fuzzy logical use to establish the diagnosis system, (b) neural network use for back propagation network which automatically adjusts the parameters of the system. A total of 124 ECGs from patients (with or without AMI) is used by this hybrid system. From the waves obtained from ECGs---Factors like Sensitivity, Specificity, partial and total accuracy is evaluated. As a result of which the neuro-fuzzy system accurately identifies patients with AMI to be 89.4% and without AMI to be 95.0%. In this system, the cons are the factors including time consumption, reasonable accuracy and variation occurring depending on clinical circumstances.

5.   T. Poggie and F. Girosi, "Networks for approximation and learning "proceeding of the IEEE, Vol. 78, pp. 1481-1497, 1990.

The authors of this paper explain about the problems in (a) approximation of nonlinear mapping, (b) i-regularization theory (c) theoretical framework for approximation which is based on regularization techniques. This hints to a class of three layer networks called as regularization networks which is related to a radial basis functions, particularly used for strict interpolation tasks. This extension leads to splines, clustering, and Bayes formulation. Also, the task-dependent clustering and dimensionality reduction are generalized and formulated via this network regularization process. Hence, the major disadvantage found in this process is that representation of the approximation of linear and nonlinear mappings are exactly appropriate in terms of simpler functions of fewer variables.

6.   Kayyan Najarian, Daniel A. Beard in "European Journal of scientific research" told about "the Prediction of Kidney Failure Using Artificial method Neural Network" in 2011.

Many persons are affected due to Kidney failure this End stage kidney disease and require some form of Renal Replacement Therapy to sustain life. More than 50 Parameters may be monitored while providing a kidney dialysis treatment. Artificial neural network which is used for predicting the failure of kidney explored and analyzed the massive data generated for several years from the kidney patients using artificial neural network to select the input variables, weight and connectivity structure to determine Optimal Network Topology and to train the system for learning. These features are used to determine and predict the life of patient. The expectancy is to detect the presence of renal failure in patient and also the survivals of the kidney dialysis can also ne predicted. The advantage is it can automatically assess the dialysis service. The Main disadvantage in this paper is there are

50 parameters take more time and multiple assumption makes the doctor in confusion to predict the kidney failure

7.   Gerardo in "Cardiac Sudden Death Risk Detection Using Hybrid Neuronal Fuzzy Networks"

The Cardiovascular Disease is the main cause of mortality. To save the life a person with accurate prediction is considered need of the hour. This literature paper designs a non-invasive technique for the analysis of HRV (Heart Rate Variability) regularly from ECG record of 1440minutes. HRV ----a measurement of the interval between R peaks of two consecutive QRS complex signals. Interrupting Noise signals are eliminated using an Adaptive filter. Also, power spectral densities are computed and measured signals filtered using HRV are characterized into High Frequencies, Low Frequencies and very Low Frequencies. The inputs are taken in time and frequency domain. By summing up these inputs into algorithm, quantification and characterization of HRV is obtained. Hence, high/low cardiac risk shall be identified. Here we find that by using this technique, the accurate and approximate values could not be found and immediate/sudden determination of predication of cardiac risk could not be found by doctors

8   A proposed theory on "A Survey on RBF Neural Network for intrusion Detection System" by Shruthi.

Network Security is one of the hottest issue faced by all the internet users across the world. Hackers have been adopting a new technology by advancement of existing technologies to create new attacks in the internet world. IDS—software used for inspecting all incoming and outgoing traffic which detects any unusual status in the network. It also uses various methods such as data mining, neural network and statistical approach. Amongst all, neural network is the best approach for detecting the malwares or any unusual statuses .RBF is also one of the feed forward and supervises technique of Neural Network. It is also used for anomaly detection and misuse detection. Besides all the benefits of IDS, this system carries certain disadvantages on optimization of RBF parameters and on response time. A proper Optimization is required with increase in time factor. Also, in IDS parallel layered approached shall be preferred than Sequential Layered approach.

9.   Messadag told about "An approach for ECG using Wavelets classification and Markov model"

Cardiac attacks can be easily monitored using ECG. It says about the relaxation and contraction of heart it gives a graphic and democratic view of a patient heart. This paper deals about the classification of ECG signal. This method is done using the Ergodic Markov Model. The MIT data are of 2 a) Discrete wavelet transform is calculated for each beat that is isolated 2) vectotial Quantization is applied. This is based on a Nonlinear Transformation. The parameters can be taken as a P,Q,R and s. and then the ECG can be calculated. The calculation of are done in different no of states in Markov model. This paper disadvantage is it cannot be easily predict and finalized with ECG alone and time gets consumed

10.   NicosMaglaveras Told About "ECG Pattern Recognition and classification using linear transformation and Non linear transformation"

Nicos Maglaveras (NMV) ----- proposed a theory that by using Linear and Non-Linear transformation, the Electrocardigram (ECG) pattern shall be easily recognized and classified. Electrocardiogram (ECG) ----- one of the extensively used signal waves in its related clinical practices which depicts data about the electrical functioning of the heart. This clear information is sketched out by modifying the shape of its integral waves (P, QRS and T). Thus the process of ECG is to reliably recognize the electrical waves and also accurately to measure the significant parameters from the temporal distribution of the

ECG Integral Waves. In this NMV theory--- while reviewing the results, the ECG is found to have non-linear transformed signals. Using the principal component analysis (linear & non-linear) and neural networks (NN) based techniques; the non-linear transformed data is mapped into n-dimensional spaces. And thus the ECG pattern is recognized. This ECG depicts accurate non-linear transformed signals and hence it isn't much easier to predict the cause of the failure of kidneys.

3. OUR APPROACH

Objective: This work, contributes the Big Data Reduction using Distributed Radial Basis Function Neural Network RBFNN[20]. This is used to predict the kidney failures in clinical decision support Systems based on the Distributed Radial Basis Function Neural Network (DRBFNN).

4. BACKGROUND

Distributed Radial Basis Function Neural Network (DRBFNN) is to predict the kidney failures [20]. There are 2 categories of algorithms that works (1) Descriptive or unsupervised learning. (2) Predictive or supervised learning (i.e., classification, regression). However, these algorithms lack in depth insight to sustain. This is mainly due to the problem to handle the data with special characteristics. These are data in health monitoring systems obtained from sensor. This new model exposed to be precise and apt on healthcare and wellbeing context [1]. The model presented by taking large number of patient records with the parameters of age, sex, hb, esr, bp, br, albumim, urea, sodium bicarbonate, rt kidney, lt of kidney

The gap between the given input vector and pre-defines center vector can be computed by using a radial basis function which is a multidimensional function. There are many ways of radial basis Function. A Gaussian called normalized Gaussian function is the one which is used as the radial basis function that is [1].

$$f(x) = \sum_{i=1}^{n} c_i G(\|x - t\|_w)$$
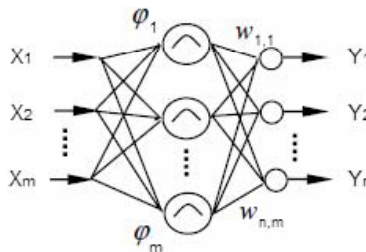
and $G(\|x - t\|_w)$ - Radial basis function

Figure 1: Depict the Distributed Radial Basis Function Neural Network

A. Architecture Diagram

The data from the source is uploaded in the dataset and then these data are then analyzed by the anomaly Detector and Batch Anamoly Score. This Anomaly Detection, which can help to automate a number of predictive tasks for fraud detection, security, quality control, diagnoses and more. These modules are checked and then those data will reach the Batch anomaly scores. The real time scores of Batch Anomaly score help increase performance of other modeling tasks. Any fraud detection that is said to be efficient should be able to sift through enormous amount of data that have various sources. Real time scores are the one which the company opts for both to perform antifraud guidelines. This is used to recognize suspicious links for realistically visualizing the connections. These

data form batch anomaly source. These are taken inside the data with scores. The clients expected probability of default is computed using the machine learning, Big Data Scoring's platform. The Decision using data scoring is decided and then the data which are not relevant are filtered and then the graph is shown with the collected data. The parameters are then compared and linear classifier is applied to these data. The classifier that is used to separate a set of objects into a line of groups is the linear classifier. The classification are mostly complex. This complex structure is needed to get the efficient solution. In hadoop using HDFS in which the name nodes and data nodes are combined together and then mapper and reducer is done. Mapper is the program that is to gather the various mappers running independently and consolidate the collected output to the reducer. Output collector is responsible for this task. The information that is collected from the map tasks are provided in the function called reporter function. All these process are carried out in many independent nodes.
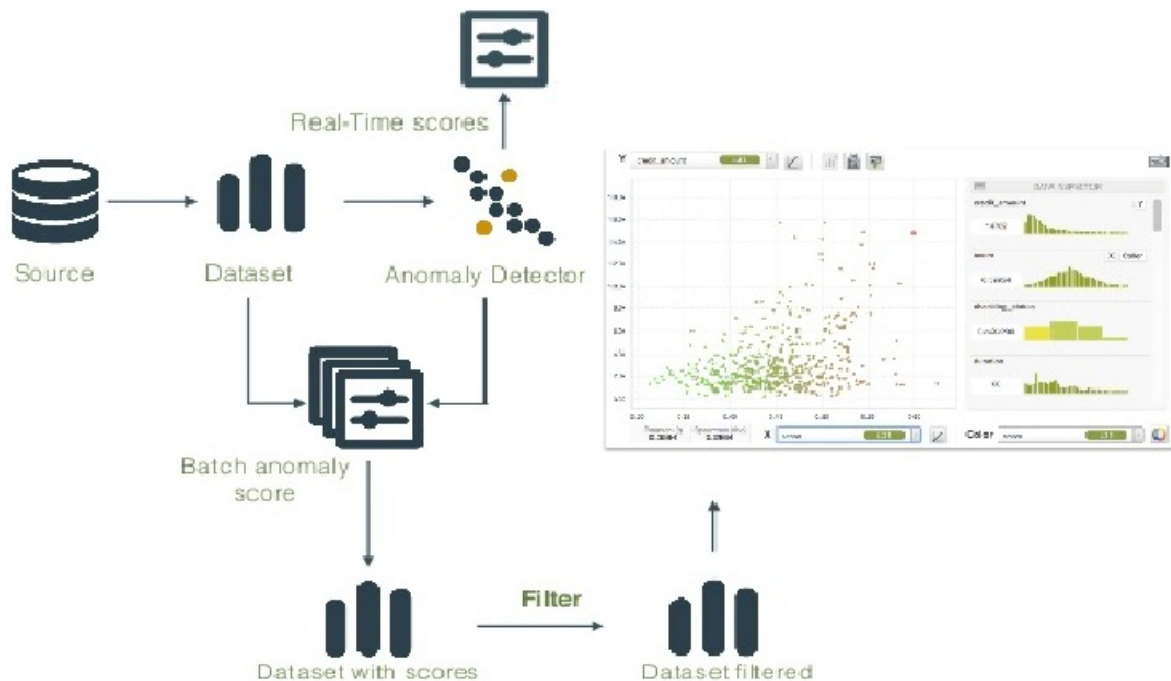

Figure 2: Working Methodology

B. Modules

1. **Anomaly Detection:** Anomaly is word coined to mean suspicious behavior. Anomaly detection is pattern finding in data. This anomaly is widely used in various applications as they relays to the great deal of information in many applications.

2. **Prediction:** Anomaly happen because of various reasons, such as (i) a processing component is given insufficient resources (ii) the input data rate exceeds the processing capacity of the component (iii) the component contains software bugs

3. **Diagnosis/decision making:** This is the very important phase. If this is not efficient then the performance of the entire system gets affected. The decision of this phase is further used for development.

4. **Other Big Data Mining Tasks**

    1. **Preprocessing**- Preprocessing is a procedure to clean and transform the data before it is passed to other modeling procedure

2. **Feature Extraction/Selection-** Feature extraction is for building various combination of different variables used to solve these problems provided the accuracy of the data is still maintained.
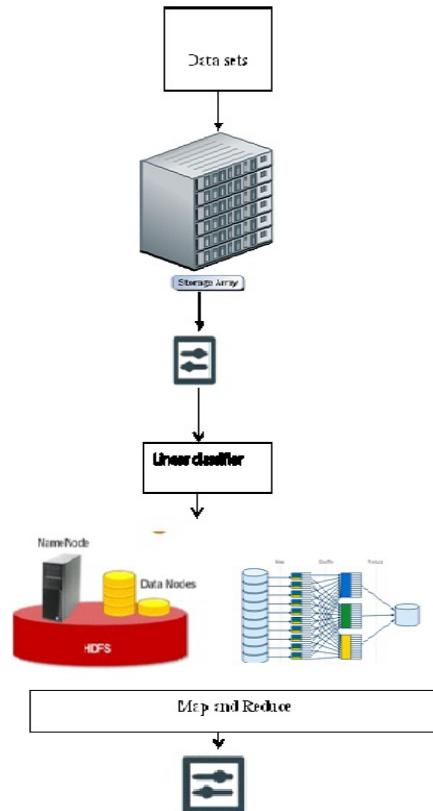

Figure 3: Architecture Diagram

5. RESULTS AND DISCUSSION

The patient prone to kidney failure is found in this method. The details of ePatient must be taken. The parameters are entered in text box i.e., Cretanine, Albumin, Sodium bicarbonate, Urea, Rt kidney, Lt kidney, age, sex, potassium and platelets are uploaded.
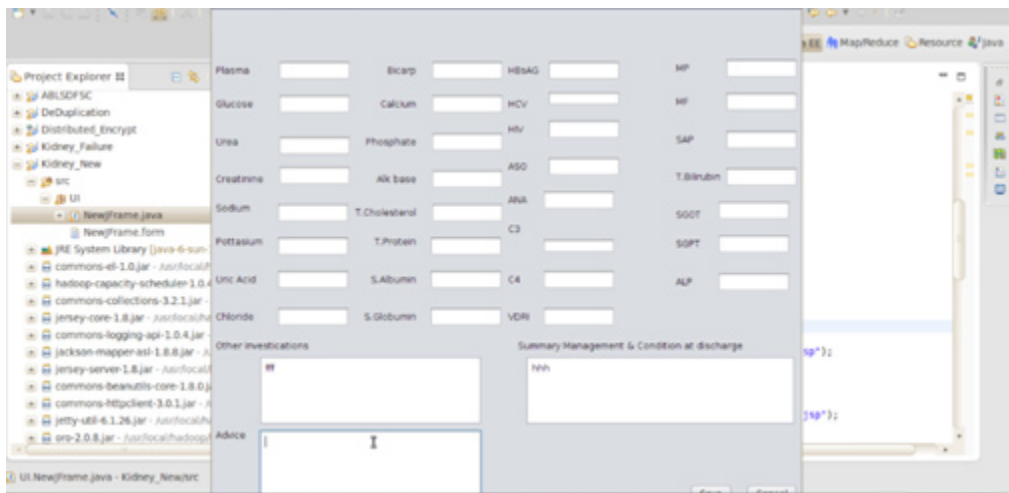

Figure 4: Sample page to update the patients details

Once the details are uploaded the system performs the prediction process and tells us whether the patient is prone to kidney failure or not. We have used a sample dataset for our use. Figure 5 depicts the sample dataset

Name	Age	Sex	Neph.No	Hb	Pcv	Tc	Dc	Esr	Platelets	urinalaysis	PR	Bp	Weight	Urea	
Mani Samy	35	male	55885	5.5 gm	17.5		8800	90/10	100/140	1.15 lakh	alb++	120/mfs	140/100 mm/Hg	73	197
Madhan	65	male	55896	5.9 gm	16.5		9100	90/6/4	50/112	1.18lakh	alb++	100/mfs	180/110 mm/Hg	65	134
joseph	34	male	57903	14 gm	43.6		7200	77/20/37	100/140	79000	alb nil	120/mfs	140/100 mm/Hg	45	226
karuppasamy	48	male	58467	7.5 gm	10000		10000	90/10	150/162	1.7lakh	alb++	100/mfs	140/100 mm/Hg	46	252
ramakrishan	51	male	58353	10.3 gm	16.5		9900	77/23	20/50	1.8lakh	alb+	120/mfs	180/110 mm/Hg	36	299
anaiappan	27	male	56511	10 gm	31		6800	60/28/12	22/70	1.46lakh	alb++	100/mfs	180/100 mm/Hg	76	143
velladurai	23	male	56523	5.9 gm	17.5		7100	80/10/4	100/140	1.14lakh	alb+	120/mfs	140/100 mm/Hg	56	152
jeyapratha	23	male	55364	5 gm	16		6000	62/32/6	50/100	1.98lakh	alb++	120/mfs	131/100 mm/Hg	45	158
manimuthu	43	male	55749	8.4 gm	15.5		10100	78/18/4	150/162	3.07lakh	alb nil	120/mfs	150/90 mm/Hg	76	218
santhosam	65	male	55536	11.3 gm	14.5		8200	90/10	110/135	1.03lakh	alb nil	120/mfs	190/100 mm/Hg	45	146
kumar	25	male	55436	82 gm	15.6		7800	80/14/6	60/110	2.9lakh	alb nil	100/mfs	140/100 mm/Hg	75	130
raj	50	male	55869	7.3 gm	16.9		8400	85/13/7	70/100	3.45lakh	alb++	120/mfs	121/90 mm/Hg	65	269
samy	40	male	56325	5.8 gm	15.8		8000	68/30/2	65/90	1.8lakh	alb+	100/mfs	120/30 mm/Hg	45	116
vivek	19	male	55693	5.9 gm	16.4		5700	63/33/4	25/52	1.4lakh	alb++	120/mfs	139/90 mm/Hg	56	278
munniyasamy	47	male	55539	6.3 gm	15.3		12000	82/12/6	20/70	1.96lakh	alb++	120/mfs	130/80 mm/Hg	46	288
chinnasamy	47	male	55469	2.58 gm	10.6		8200	72/26/2	20/48	1.8lakh	alb+	100/mfs	126/90 mm/Hg	76	163
durai	49	male	55639	6.7 gm	11.2		6100	90/9/1	100/142	5.5lakh	alb+	100/mfs	143/90 mm/Hg	55	188
deva	60	male	55123	5.9 gm	13.2		5100	74/22/4	55/120	2.98lakh	alb+	100/mfs	140/100 mm/Hg	45	259
chandran	36	male	56369	3.6 gm	10.6		4300	88/11/1	120/192	2.49lakh	alb nil	120/mfs	120/100 mm/Hg	76	211
rahul	27	male	54963	6.5 gm	14.6		5900	70/27/3	50/120	3.06lakh	alb nil	100/mfs	126/50 mm/Hg	56	92
saravanan	40	male	58963	6.2 gm	18.9		5700	83/16/11	120/140	1.75lakh	alb nil	120/mfs	129/58 mm/Hg	66	107
kanagarai	54	male	56398	6.5 gm	17.1		10100	80/14/6	60/115	1.79lakh	alb nil	100/mfs	130/110 mm/Hg	79	221

Figure 5: Sample dataset

Once the data is uploaded the system will tells us that the information is uploaded successfully. Figure 6 depicts the screen which tells us that the information is uploaded sucessfully.
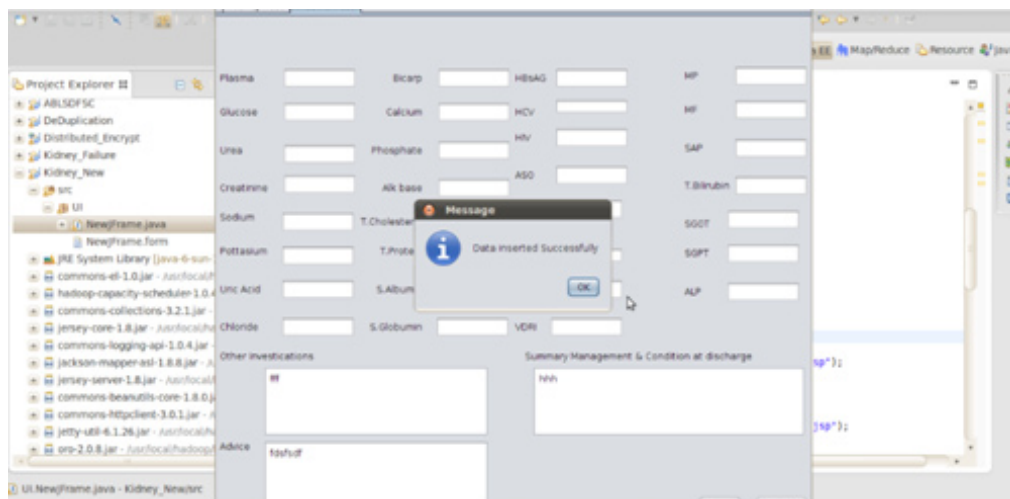

Figure 6: Screen after the data upload

The above figure displays a dialog box with the message dataset uploaded successfully. After this process in which the patient details are uploaded the system then does the prediction. The system further use CLASSIFIER and then HDFS to determine the stages of a patient affected by a kidney failure. The RBFNN with classifier algorithm with the given parameters will determine the condition of a patient as a normal or a kidney failure patient. Thus the proposed method reveals the stages of the kidney failure patient and also provides details about treatment and clinical decision.

6. CONCLUSION

We conclude that it is useful for handling large volumes of clinical dataset for storing and retrieve it. The comparison of patient dataset results may promise to yield knowledge about patient records and to find the root cause of the kidney failures. In future it will be used to predict blood clots in veins, brain tumour, weather forecasting, heart attack, diabetics patient dataset. In Phase II, the complexity evolves due to data reduction can

be overcome with the use of Map Reduce. So work of phase two is to implement DRBNN algorithm for big data reduction. Accurate results are produced by predictive modelling and it increases efficient handling of large volumes of patient dataset.

REFERENCES

[1]   N. Pombo, P. Araujo, and J. Viana, "Knowledge discovery in clinical decision support systems for pain management: A systematic review," Artificial Intelligence in Medicine, Vol. 60, No. 1, pp. 1-11, 2014.

[2]   I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," Artificial Intelligence in Medicine, Vol. 23, No. 1, pp. 89-109, 2001.

[3]   A. H. Gerardo and R. C. M. Antonio, "Cardiac Sudden Death Risk Detection Using Hybrid Neuronal-Fuzzy Networks," in Electrical and Electronics Engineering, 2006 3rd International Conference on, 2006, pp. 1-4.

[4]   H. L. Lu, K. Ong, and P. Chia, "An automated ECG classification system based on a neuro-fuzzy system," in Computers in Cardiology 2000, 2000, pp. 387-390.

[5]   D. J. Messadeg, C. Snani, and M. Bedda, "An approach for ECG classification using wavelets and Markov Model," in Information and Communication Technologies, 2006. ICTTA '06. 2nd, 2006, Vol. 1, pp. 1910-1913.

[6]   A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.," Anal. Chem., Vol. 36, No. 8, pp. 1627-1639, Jul. 1964. "Evida, website da platforma, https://evida.pt/(last access 5 May 2014)."

[7]   S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

[8]   T. Poggio and F. Girosi, "Networks for approximation and learning," Proceedings of the IEEE, Vol. 78, No. 9, pp. 1481-1497, 1990. "ALLab-Assisted Living Computing and Telecommunications Laboratory, http://allab.it.ubi.pt/mediawiki/index.php/Main-Page (last access 5 May 2014).", "Plux wireless biosignals, 'bioPlux', from www.plux.info/www.plux.info/http://www.plux.info/(last access 5 May 2014).".

[9]   A. Godfrey, R. Conway, D. Meagher, and G. OLaighin, "Direct measurement of human movement by accelerometry," Medical Engineering & Physics, Vol. 30, No. 10, pp. 1364-1386, Dec. 2008.

[10]   N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, and M. Strintzis, "{ECG} pattern recognition and classification using non-linear transformations and neural networks: A review," International Journal of Medical Informatics, Vol. 52, No. 3, pp. 191-208, 1998.

[11]   Oates, Tim, and David Jensen. "Large Datasets Lead to Overly Complex Models: An Explanation and a Solution." *KDD*. 1998.

[12]   Turinsky, A., and R. Grossman. "A framework for findingdistributed data mining strategies that are intermediate between centralized strategies and in-place strategies." Proceedings of Workshop on Distributed and Parallel Knowledge Discovery at KDD-2000. 2000.

[13]   Lazarevic, Aleksandar, and Zoran Obradovic. "Data reductionusing multiple models integration." European Conference on Principles of Data Mining and Knowledge Discovery. Springer Berlin Heidelberg, 2001.

[14]   Freund, Yoav, Robert Schapire, and N. Abe. "A short introduction to boosting." Journal-Japanese Society For Artificial Intelligence 14.771-780 (1999): 1612.

[15]   Quinlan, J. R. "Learning efficient classification proceduresand their application tu chess end games." Machine Learning,Micha1ski, RS, Carbone11, JC and Mitce11, TM (Eds.), Springer - Werlag, New York, r" p 463 (1983).

[16]   Yang, Ying-Kuei, et. al., "A novel self-constructing radial basis function neural-fuzzy system." Applied Soft Computing 13.5 (2013): 2390-2404.

[17] Sun, Jimeng, and Chandan K. Reddy. "Big data analytics for healthcare." Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013.

[18] Carlton, M. "How big data analytics reduced medical re-admissions." IBM Corporation (2012).

[19] Balan, Shilpa, and Joseph Otto. "Big Data Analysis of Home Healthcare Services." Information Technology and Management Science 19.1 (2016): 53-56.Knowledgent. Big Data and Healthcare Payers. 2013.

[20] Yang, Ying-Kuei, et. al., "A novel self-constructing radial basis function neural-fuzzy system." Applied Soft Computing 13.5 (2013): 2390-2404.