International Journal of Engineering & Technology, 7 (1.3) (2018) 191-194

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Classification of non-chronic and chronic kidney disease using SVM neural networks

Ravindra B.V[1]*, N.Sriraam [2], M Geetha[3]

[1]*School of Information science, Manipal Academy of Higher Education, Manipal, India*
[2]*Centre for Medical Electronics and Computing, M.S. Ramaiah Institute of Technology, Bangalore, India*
[3]*Department of CSE MIT, Manipal Academy of Higher Education, Manipal, India*
**E-mail: ravindra.bv@manipal.edu*

Abstract

Chronic kidney disease (CKD) refers to the failure of the renal functionalities that leads to the deposition of wastes, electrolytes and other fluids in the body. It is very important to recognize the symptoms that cause the CKD and pathological blood and urine test indicates the key attributes. It is well fact that one has to undergo dialysis due to renal failure. The severity level of disease can be predicted as well as classified using appropriate computer aided quantitative tools. This specific study discusses the classification of chronic and non-chronic kidney disease NCKD using support vector machine (SVM) neural networks. The simulation study makes use of UCI repository CKD datasets with n=400. In order to train to train the attributes of kidney dialysis four cases were considered by including the nominal and numerical values. A radical basis kernel function was employed to train SVM. The performance of the proposed scheme is evaluated in terms of the sensitivity, specificity and classification accuracy. Results reveal an overall classification accuracy of 94.44% was obtained by combining 6 attributes. It can be concluded that the SVM based approach found to be a potential candidate for classification of CKD and NCKD.

Keywords: *Hemodialysis; CKD; NCKD; SVM;.*

1. Introduction

Chronic kidney disease (CKD) refers to the renal failure where typical symptoms will be observed such as leg swelling, vomiting, tiredness [1]-[5]. General diagnosis is performed through blood, urine, test to know the functionality of the kidney. Hence it is essential to introduce an earlier screening mechanism for identifying symptoms of CKD, so that precautionary measures can be taken to avoid any complications. Several attempts have been made in the past to predict and classify various kidney diseases [6]-[8]. Further artificial intelligence based computational tools have been applied to develop automated classification of CKD from non-chronic kidney diseases (NCKD) [6]-[10]. This study suggests the application of support vector machine (SVM) neural network for the classification of CKD and NCKD.

Dataset derived from open source DB database repository CKD data was considered for this study. Four classes were considered to train the SVM. The raw attribute value was initially normalized t0 ensure that, The SVM neural network perform better training. The performance of the SVM classifier is evaluated in terms of sensitivity, specificity and classification accuracy. Experimental simulate shows an overall classification accuracy of 94.4% was obtained which was found to be better than the result reported earlier in the literature. Fig.1 shows the proposed schematic flow diagram.

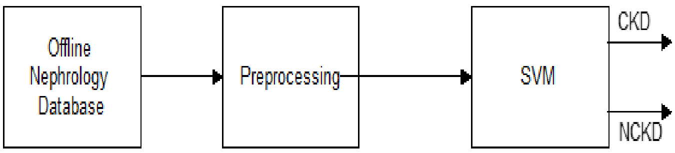

Fig.1: Proposed Scheme

2. Related work

A predictive approach in predicting Survival of kidney dialysis patients was reported [11]. Two data mining techniques are employed and a classification accuracy of 97.78% was reported. An enhanced version of data mining approach was proposed by the same research group [12]. Database of 188 patients based on 707 visits was created for the purposed study, two category, ‘alive’ and ‘diseased’ to make decision for survival period of the patient. Dialysis period of three was considered on the basis two data mining techniques roughest theory and decision tree was employed. The entire data base was divided into 8 trails and the classification was performed the rules were framed based on lower and upper approximation. Sixteen different classifiers using the data mining rules helps in predicting the survival rate. An overall classifier accuracy range 75% to 85% was achieved. It was concluded that the selection of kidney dialysis parameters plays a significant role in predicting the survival prediction.

Sriraam et al., investigated the urological related disorder using the data mining techniques. The two techniques namely, association mining and decision tree were proposed for predicting the survival period of kidney failure patients. Database from Malaysian hospital was used. It was shown that the decision tree mining yielded promising results suitable for clinical decision making [13],[14]. Pothraju and Sreedevi have proposed a framework called SMOTHE for prediction of kidney disease. The procedure minimizes the variations between the balanced and imbalanced dataset. An overall classification accuracy of 98.73% was reported [15]

Koklu &Tutuniu made an attempt to clearly to classify CKD and NCKD using four data mining methods[16]. Though results were shown, no specific details of the algorithm was reported. Sinha and Sinha have applied KNN and SVM techniques for the same classification problems [17]. The result were not as accurate as compared to the earlier reported iteration in [18]. A random forest machine learning algorithm was applied for the CKD classification problem. The application of appropriate feature selection technique was reported in the details for chronic kidney disease diagnosis [19]. Ravindra et al., have reported the importance of attribute selection for kidney dialysis survival prediction by making use of k-means clustering techniques [20],[21].

3. Study Data

This specific study makes use of open source data located in the University of California Irvine Machine learning repository [22],[23]. A total of 400 subjects data were stored where 250 were labeled as CKD and remaining 150 as NCKD. A total of 25 attributes were available. Table 1 shows the complete details [22].

4. Support Vector machine Classification

Support vector machine neural network was found to be an appropriate pattern classifier model for engineering and clinical application in the recent decade. Through a non-linear phenomenon, SVM project the given input pattern to higher –dimensional feature space and later linear decision surface is constructed [24]-[26]. The construction of efficient hyper plane leads to better pattern classification. Fig.2 illustrates the hyper plane separating binary classes [24].

Table 1: Attributes and their properties in patient dataset

No	Attribute	Type	Units or Values
1	Age	Numerical	Years
2	Blood Pressure	Numerical	mm/Hg
3	Specific Gravity	Nominal	(1.005,1.010,1.015,1.020,1.025)
4	Albumin	Nominal	(0,1,2,3,4,5)
5	Sugar	Nominal	(0,1,2,3,4,5)
6	Red Blood Cells	Nominal	(normal, abnormal)
7	Pus Cell	Nominal	(normal, abnormal)
8	Pus Cell clumps	Nominal	(present, not present)
9	Bacteria	Nominal	(present, not present)
10	Blood Glucose Random	Numerical	mgs/dl
11	Blood Urea	Numerical	mgs/dl
12	Serum Creatinine	Numerical	mgs/dl
13	Sodium	Numerical	mEq/L
14	Potassium	Numerical	mEq/L
15	Hemoglobin	Numerical	gms
16	Packed Cell Volume	Numerical	
17	White Blood Cell Count	Numerical	cells/cumm
18	Red Blood Cell Count	Numerical	millions/cmm
19	Hypertension	Nominal	(yes, no)
20	Diabetes Mellitus	Nominal	(yes, no)
21	Coronary Artery Disease	Nominal	(yes, no)
22	Appetite	Nominal	(good, poor)
23	Pedal Edema	Nominal	(yes, no)
24	Anemia	Nominal	(yes, no)
25	Class	Nominal	(CKD, NCKD)

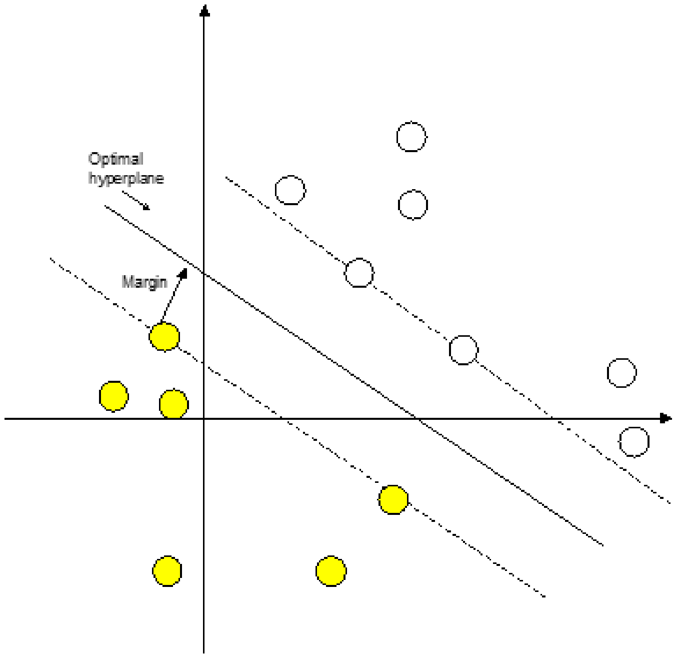

Fig. 2: Hyper plane separating binary class

The selection of kernel function is very crucial for pattern classification. In order to train datasets with SVM, four cases considered to the study. Table 2 shows the conditions considered for this study.

Table 2: Conditions considered for this study

Cases	Attributes
Case 1	Blood Pressure Specific gravity Serum creatinine
Case 2	Albumin Sugar Blood glucose random Hemoglobin
Case 3	Package cell Volume WBCC RBCC
Case 4	Albumin Sugar Blood glucose random Serum Creatinine Sodium Potassium Hemoglobin

In order to train the network with dialysis parameter, a simple normalization procedure is considered for this proposed study. The attribute normalization, Attribute (N) is defined as:

$$\text{Attribute (N)} = \frac{\{(\text{Original attribute value}) - (\text{upper limit of attribute})\}}{\{(\text{Upper limit of the attribute}) - (\text{Lower limit of the attribute}) + 1)\}} \quad (1)$$

The effect of normalization was found to be significant during the SVM training procedure.

In order to train the SVM network, radial basis kernel function was employed. The performance of the SVM classifier was evaluated in terms of the following parameters.

Specificity = $T_P / (T_P+T_N)$ (2)

Sensitivity= $T_N / (F_P+T_N)$ (3)

Classification accuracy =
$(T_P+T_N) / (T_P+F_P+T_N+F_N)$ (4)

Where

T_P = correctly recognized CKD attributes.
F_N = Incorrectly recognized as NCKD
T_N = Correctly recognized NCKD attributes
F_P = Incorrectly recognized as CKD

The results are translated in Table 3.

Table 3: Classification Results

Cases	SE (%)	SP (%)	CA (%)
Case 1	**100**	**8.6**	**65.75**
Case 2	**98**	**91.2**	**93.75**
Case 3	**80**	**88.8**	**85.8**
Case 4	**84.6**	**98**	**93**

Most of the work reported in the literature that makes use of UCIML datasets have employed the machine learning algorithms for the entire dataset [8], [12], [16]-[18]. The proposed work emphasize the importance towards the close correlation that

exists among the kidney dialysis parameter. The influence of albumin, sugar and blood glucose have contributed significantly towards the classification of CKD and NONCKD. The results were comparable to that of the work reported earlier [8], [12],[16]-[18].

It can be observed that the effect of albumin, sugar, blood glucose random and hemoglobin attributes found to be significant in performance, better binary classification of CKD (VS) NCKD. A maximum classification accuracy of 93.75% was obtained.

Fig.3 shows the effect of performance of SVM classifier with normalization and without normalization of the attribute. Only two cases were shown

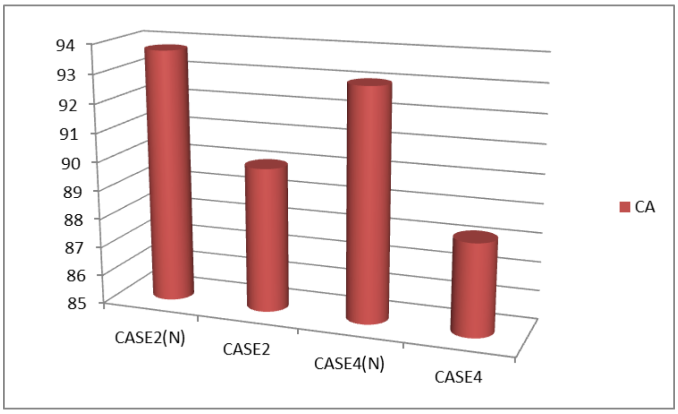

Fig. 3: Performance of SVM Classifier

Fig.4 shows the computational requirements for the SVM algorithm.

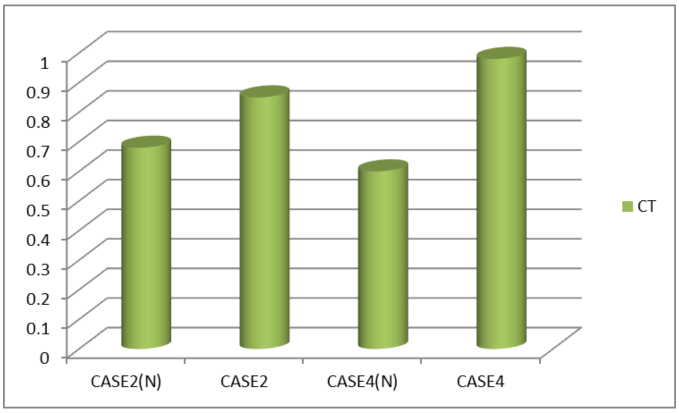

Fig. 4: Computational Complexity of the proposed SVM Classifier

It can be inferred from Figs 3 and 4 that the SVM classifier with attribute normalization outperforms than without normalization.

5. Conclusions

This specific study makes use of support vector machine (SVM) model for classification of CKD and NCKD. The total 24 attributes recorded from UCI machine repository, CKD database were categorized into four cases depending on its close association between the attributes. Stimulation shows an overall classification accuracy of 93.75% with SVM classifier that makes use of radial basis kernel function.

References

[1] Liao, Min-Tser; Sung, Chih-Chien; Hung, Kuo-Chin; Wu, Chia-Chao; Lo, Lan; Lu, Kuo-Cheng (2012). "Insulin Resistance in Patients with Chronic Kidney Disease". Journal of Biomedicine and Biotechnology. 2012

[2] What Is Chronic Kidney Disease?". National Institute of Diabetes and Digestive and Kidney Diseases. June 2017. Accessed 11 November 2017.

[3] V. A. Moyer, “Screening for chronic kidney disease: Us preventive services task force recommendation statement,” Annals of internal medicine, vol. 157, no. 8, pp. 567–570, 2012.

[4] L. C. Plantinga, D. S. Tuot, and N. R. Powe, “Awareness of chronic kidney disease among patients and providers,” Advances in chronic kidney disease, vol. 17, no. 3, pp. 225–236, 2010.

[5] Redmon JH, Elledge MF, Womack DS, Wickremashinghe R, Wanigasuriya KP, Peiris-John RJ, Lunyera J, Smith K, Raymer JH, Levine KE (2014). "Additional perspectives on chronic kidney disease of unknown aetiology (CKDu) in Sri Lanka – lessons learned from the WHO CKDu population prevalence study". BMC Nephrology. 15 (1): 125. doi:10.1186/1471-2369-15-12

[6] Abeer Y. Al-Hyari,Ahmad M. Al-Taee ,Majid A. Al-Taee, Diagnosis and Classification of Chronic Renal Failure Utilising Intelligent Data Mining Classifiers, International Journal of Information Technology and Web Engineering, Volume 9 Issue 4, October 2014,Pages 1-12

[7] José Neves ., M. Rosário Martins ., João Vilhena .,João Neves ., Sabino Gome.,António Abelha., José Machado.,Henrique Vicente, A Soft Computing Approach to Kidney Diseases Evaluation, J Med Syst (2015) 39: 131,DOI 10.1007/s10916-015-0313-4

[8] Naganna Chetty, et.al, Role of Attributes Selection in Classification of Chronic Kidney Disease Patients, IEEE conference, pp. 82-87, 2015

[9] M. Baumgarten and T. Gehr, “Chronic kidney disease: detection and evaluation,” American family physician, vol. 84, no. 10, p. 1138, 2011.

[10] Levey AS, de Jong PE, Coresh J, El Nahas M, Astor BC, Matsushita K, et al. The definition, classification,and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. Kidney Int,2011;80:17–28.

[11] Balasaravanan, K. "Detection of dengue disease using artificial neural network based classification technique." *International Journal of Engineering & Technology* 7, no. 1.3 (2017): 13-15.
[12] Kusiak, A., Dixon, B., & Shah, S., Predicting survival time for kidney dialysis patients. *Computersin Biology and Medicine*, *35*(4), 311–327. doi:10.1016/j.compbiomed.2004.02.004 PMID:15749092,2005.
[13] N. Sriraam, V. Natasha and H. Kaur,” data mining approaches forkidney dialysis treatment”, Journal of Mechanics in Medicine and Biology, Volume 06, Issue 02, June 2006.
[14] N Sriraam, V Natasha, H Kaur,Data Mining Techniques and MedicalDecision Making for Urological Dysfunction,Handbook of Research on Informatics in Healthcare and Biomedicine,2006
[15] Sai Prasad Potharaju &M. Sreedevi, An Improved Prediction of Kidney Disease using SMOTE, Indian Journal of Science and Technology, Vol 9(31), DOI: 10.17485/ijst/2016/v9i31/95634, August 2016
[16] Murat Kolu, Kemal Tutuncu, Classification of chronic disease with most known data mining methods
[17] Sinha, P, & Siha, P. (2015). Comparative Study of Chronic Kidney Disease Prediction using KNN and SVM. In International Journal of Engineering Research and Technology, Vol. 4, No. 12, ESRSA Publications.
[18] Kumar M. (2016). Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm, Journal of Computer Science and Mobile Computing, Vol.5 Issue.2, p. 24-33.
[19] Made Satria Wibawa, Made Dendi Maysanjaya, I Made Agus Wirahadi Putra, "Boosted classifier and features selection for enhancing chronic kidney disease diagnose", Cyber and IT Service Management(CITSM)20175th International Conference on, pp. 1-6, 2017.
[20] Ravindra B V , Sriraam. N ,and Ms. Geetha Mayya.Discovery of significant parameters in Kidney dialysis data sets by K-means algorithms International conference Proc. of International Conference on Circuits, Communication, Control and Computing (I4C),ISBN :978-1-4799 pp.452-454,2014
[21] B.V. Ravindra, N. Sriraam, Geetha Maiya,Finding Impact of Precedence based Critical Attributes in Kidney Dialysis Data Set using Clustering Technique,,International Journal of Biomedical and Clinical Engineering (IJBCE),Vol.4,issue1, pp.44-50,Jan-June 2015.
[22] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
[23] L.Jerlin Rubini, et.al “Generating comparative analysis of early stage prediction of Chronic Kidney Disease”, International Journal Of Modern Engineering Research (IJMER) Vol. 5, pp 50-55 2015
[24] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, NY, USA, 1995.
[25] Vapnik VN. An overview of statistical learning theory. IEEETrans Neural Network[
[26] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge University Press, 2000