ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 09, Issue, 11, pp. 31431-31434, November, 2019

RESEARCH ARTICLE OPEN ACCESS

ARTIFICIAL INTELLIGENCE IN PREDICTING CHRONIC KIDNEY DISEASE

***Vanessa D. Martins, Viviane S.Ferreira, Marta O. Barreiros, Ilka P. Belfort and Allan K. Barros**

Department of Electrical Engineering, Laboratory for Biological Information Processing (PIB), Federal University of Maranhao (UFMA) Sao Luis-MA, Brazil

ARTICLE INFO

Article History:
Received 17th August, 2019
Received in revised form
26th September, 2019
Accepted 11th October, 2019
Published online 30th November, 2019

Key Words:
Machine learning, Classifiers, Prediction, Chronic Kidney Disease.

Corresponding author: Vanessa D. Martins

ABSTRACT

The prediction of the future is becoming an increasingly easy and discussed task in the literature, especially in healthcare, with predictive analyzes of medical data using the machine learning, which evolved after the development of new informed technologies that originated multiple search fields. Much dedication is fulfilled periodically to deal with an explosion of medical data, to gain knowledge of it, to predict disease, and to anticipate healing. In order to extract useful knowledge and aid decision-making, researchers are increasingly applying technical innovations, including database analysis, predictive analysis, machine learning and learning algorithms. Thus, aimed to conduct a review of literature on the use of Artificial Intelligence in the prediction of Chronic Kidney Disease.

Citation: Vanessa D. Martins, Viviane S. Ferreira, Marta O. Barreiros, Ilka P. Belfort and Allan K. Barros, 2019. "Artificial intelligence in predicting chronic kidney disease", *International Journal of Development Research*, 09, (11), 31431-31434.

INTRODUCTION

Artificial intelligence (IA) is the area of computer science that aims to simulate the processes of human thinking, having the ability to learn, store knowledge and solve problems (Krittanawong, 2017). The use of IA techniques has become widely accepted in medical applications, showing a growing number of medical devices available on the market, along with a fast pace of medical journal publishing, with more than 500 academic publications each year (Gant, 2001). The medical field makes an extreme contribution to the magnitude of medical data because of some innovations in the field, such as cloud computing, laparoscopic surgery, and robotic surgery, which replaced classical surgery (Gabriel, 2010).There are also intelligent applications or software that can analyze body signals using integrated sensors for monitoring purposes, as well as technologies that support new biological, behavioral and environmental data collection methods. These include sensors that monitor phenomena with high precision (Steve, 2014). All of these innovations come from the grandiosity of medical data by multiplying electronic medical data sources and records containing diagnostic images, laboratory results, and biometric data (Steve, 2014; Weil, 2014 and Groves, 2013).Researchers have deduced that this explosion of medical data has the potential to improve point-of-care decisions. The physician will be able to extract relevant knowledge for each patient, which provides better decisions and outcomes (Huang, 2015). There are many classification and prediction algorithms that can be applied to predict various diseases such as breast cancer, heart disease, motor neuron, diabetes, chronic kidney disease, among others. There is ongoing research work using Artificial Intelligence techniques in the field of medical diagnosis for these diseases (Boukenze, 2017). Kidney Disease is currently considered a global health problem as it affects millions of people worldwide. This disease is considered dangerous if not treated immediately in time, and can be fatal. If doctors have a good tool that can identify patients who are likely to have kidney disease in advance, they can start treatment faster, thus avoiding complications of the disease (Levey, 2012 and Wang, 2016). Thus, the objective was to perform a literature review of research conducted with Artificial Intelligence in the prediction of Chronic Kidney Disease (CKD).

Research using Artificial Intelligence: Ho *et al.* (2012) presented a computer-assisted diagnostic implement based on ultrasound image analysis. The system was used to detect and classify different stages of CKD. They used the K-means machine learning algorithm to detect after the image preprocessing step.

Table 1. Classification Algorithms for CKD Prediction

Authors (Year)	Location	Database	Instance	Attributes	Methods	Accuracy (%)
Kusiak et al (2005)	USA	UIHC	188	50	DT	75
					RS	67
Abhishek et al (2012)	India	Hospitals	1199	7	BPA	81
					RBF	62
					SVM	60
Chiu et al (2012)	Taiwan	HEC	430	6	BPN MNN	94,75
					GFNN	93,23
						86,63
Vijayarani and Dhayanand (2015)	India	KFT	583	6	SVM	76,32
					NB	70,96
Charleonnan et al (2016)	Thailand	UCI	400	24	SVM KNN	98,3
					LR	98,1
					DT	96,55
						94,8
Boukenze et al (2016)	Marroco	UCI	400	24	C4.5	63
					SVM	60,25
					NB	57,5
Anantha and Parthiban (2016)	India	CDC	600	12	DT	91
					NB	86
Tazin et al (2016)	Bangladesh	UCI	400	15	DT	99
					SVM	98
					KNN	97
					NB	96
Manish (2016)	India	UCI	400	25	RF	100
					SMO	97
					NB	95
					RBF	98
					MLPC	98
					SLG	98
Chimwayiet al (2017)	India	UCI	400	24	Neuro-Fuzzy	97

The study collected multiple ultrasound images of patients with kidney disease, and selected representative CKD images were applied for pre-analysis and comparison training. They concluded that transition sites calculated as reference indicators could provide physicians with an objective and auxiliary computational aid diagnostic tool for CKD identification and classification. Valderrama*et al.* (2014) suggested the feasibility study of using a distributed approach for alarm management of patients with kidney disease. They handled alarms related to monitoring CKD patients within the eNefro project.

The results section shows, through the proof of concept studied, the feasibility of Data Distribution Service (DDS) for enabling emergency protocols in terms of prioritization and personalization alarm, as well as some observations on security, privacy and performance real-time communication.Rosmani*et al.* (2015) developed self-care guidelines for CKD patients using Adobe Flash CS5.5. This CKD patient self-care site was developed using Adobe Dreamweaver, and has been helping to manage CKD patient self-care daily by creating a more effective channel of information designed for them.Hsieh *et al.* (2014) suggested that a real-time system for analyzing chronic kidney disease could be developed using ultrasound images only. The learning set was also used to classify chronic kidney disease by constructing a classifier using Support Vector Machine (SVM) to predict and classify the stage of CKD with ultrasound images. Singh *et al.* (2014), showed different methods to leverage the hierarchical structure in ICD-9 codes for CKD and heart failure assessment through high dimensionality data. This research proposed and evaluated a new feature of the engineering approach to leverage this hierarchy while improving the performance of predictive methods.

Classification Techniques

Classification and prediction is a data mining technique that first uses training data to develop a model and then the resulting model is applied to test data to obtain predictionresults (Mandli, 2014). Several classification algorithms were applied to data sets for the diagnosis of chronic kidney disease and the results were considered very promising (Table 1).

Kusiak*et al.* (2005) used preprocessing, transformations and data mining to gain insight into the interaction between many of the measured parameters and patient survival. Two different data mining algorithms were employed to extract knowledge in the form of decision rules. These rules were used by a decision-making algorithm, which predicts the survival of new hidden patients.They used Reed-Solomon (RS) and Decision Tree (DT) algorithms in 188 patients at the University of Iowa Hospitals and Clinics (UIHC). They totaled 50 important parameters identified by data mining which were interpreted by specific physicians. The decision tree algorithm (DT) produced 75% and RS with 67% correct predictions for the test data set.They introduced a new concept in their research, which was applied and tested using data collected at four sites with dialysis patients. The approach presented in his paper reduced the cost and effort of selecting patients for clinical studies. Patients can be chosen based on the prediction results and the most significant parameters discovered.

Abhishek et al. (2012) used three neural network techniques: the Back Propagation Algorithm (BPA), Radial Basis Function (RBF) and a nonlinear Support Vector Machine (SVM) classifier and compared them according to their efficiency and accuracy. They used the WEKA 3.6.5 implementation tool to find the best technique among the three algorithms for kidney stone diagnosis. The main objective of his thesis work was to propose the best diagnostic tool, such as kidney stones identification, to reduce diagnostic time, efficiency and accuracy. The data set for kidney disease was obtained from medical reports of patients from different hospitals and 1199 patients with 7 attributes each were used: age, sex, lymphoctins, monocytes, eosinophils, neutrophils, creatinine. From the experimental results they concluded that the Back Propagation Algorithm (BPA) significantly improved the conventional classification technique for use in the medical field with 81% accuracy over RBF and SVM with 62% and 60%, respectively. Chiu et al. (2012) presented an intelligent model for detecting kidney disease and assessing the severity of a patient. This intelligent model utilizes three types of artificial neural networks including back-propagation networks (BPN), generalized feeding neural networks (GRNN) and modular neural networks (MNN). The input data set for the development of neural networks was collected from the health examination cases provided by this study's collaborative hospital (HCE), which used 430 patients with 6 instances each: creatinine, glucose, systolic pressure, proteinuria, hematuria and urea. The best performing model was chosen for system development.The BPN algorithm obtained the highest accuracy of 94.75% in relation to MNN (93.23%) and GRNN (86.63%). The system developed in line with the best model was deployed on Google's cloud platform, leveraging the Google Application Engine.

Vijayarani and Dhayanand (2015) aimed to predict CKD using Support Vector Machine (SVM) and Naive Bayes (NB). The objective was to compare the performance of these two algorithms based on their accuracy and execution time. A synthetic kidney function test (KFT) dataset was created for the analysis of kidney disease. The data set with 584 instances and 6 attributes used in the comparative analysis were: Age, Sex, Urea, Creatinine and Glomerular Filtration Rate (GFR). This data set consists of affected kidney disease in formation. From the experimental results they observed that the SVM performance was better (76.32% accuracy) compared to the other algorithm (70.96% accuracy). Charleonnan et al. (2016) used four machine learning methods including the K-nearest neighbors (KNN), support vector machine (SVM), logistic regression (LR) and decision tree classifiers (DT) to predict kidney disease with the aid of the WEKA tool, with a database collected from the UCI Machine Learning Repository (University of California Irvine), consisting of 400 attributes and 24 instances (age, blood pressure, severity specific, albumin, sugar level, red blood cells, pus cells, agglomeratesof pus cells, bacteria, blood glucose, blood urea, serum creatinine, sodium,potassium, hemoglobin, cell volume, white blood cell count,red blood cells, hypertension, diabetes, coronary artery disease, appetite, edema andanemia).From the experimental results, they concluded that the SVM classifier provided higher accuracy and, moreover, the SVM has higher sensitivity after training and testing by the proposed method. The SVM classifier showed the highest accuracy than others with 98.3%, while the KNN, Logistic Regression (RL) and Decision Tree (DT) can produce the average accuracy of 98.1%, 96.55% and 94.8%, respectively. Boukenze *et al.* (2016) used machine learning algorithms such as Support Vector Machine (SVM), Decision Tree (C4.5), and Naïve Bayes (NB). These predictive models are constructed from the chronic kidney disease (UCI) dataset using Weka. Simulation results showed that the C4.5 classifier proved its predictive performance with better results in terms of accuracy and execution time obtained accuracy of 63% followed by SVM (60.25%) and NB (57.5%). Anantha and Parthiban (2016), obtained high accuracy using Decision Tree for early detection of CKD.

In their work, they aimed to predict early detection of chronic kidney disease for diabetic patients with the help of machine learning methods and finally suggested a decision tree to arrive at concrete results with desirable accuracy, measuring their performance to their specification and sensitivity. The Clinical Foundation HeartDisease's available data set of 600 clinical records was collected from a major Chennai-based diabetes research center with 12 instances each: gender, age, heredity, weight, smoking, blood pressure, fasting glucose, postprandial glucose, glycolyzed hemoglobin test, LDL, HDL, VLDL. They tested the data set for classification using Naïve Bayes (NB) and the Decision Tree (DT) method.By comparing the classification algorithms, they concluded that the accuracy is up to 91% for the Decision Tree classification compared to 86% for Naïve Bayes. In order to increase the accuracy of the prediction result, they also used neural network algorithms and clustering data that helped a lot in the mission and also provided room for future research. Tazin*et al.* (2016) used classification algorithms Supporting Machine Vector (SVM), Decision Tree (DT), Naïve Bayes (NB) and K-Nearest Neighbor (KNN), in the analysis of Chronic Kidney Disease Data collected. In the UCI repository to predict the presence of kidney disease. In the study, the decision tree (DT) shows promising results (99% accuracy) when implemented using the WEKA data mining tool, followed by SVM, KNN and NB with 98, 97 and 96% accuracy values, respectively.

The classification algorithm provides vital improvements in classifications with appropriate numeric attributes. Manish (2016) in his work predicted the risk of chronic kidney disease by comparing numerous algorithms that were implemented using the WEKA tool. The researcher focused on the application of several classifier algorithms including Random Forest (RF), Minimal Sequential Optimization (SMO), Naive Bayes (NB), Radial Basis Function (RBF) and Multilayer Perceptron Classifier (MLPC) and Simple Logistic (SLG), and obtained high accuracy values of 100, 97, 95, 98, 98 and 98, respectively, comparing them with the numerous methodologies applied. The researcher also used validation to classify each classifier. Chimwayi*et al.* (2017) applied the neuro-fuzzy algorithm to determine the risk of CKD in patients. Predictions made using neuro-fuzzy gave 97% accuracy from the chronic kidney disease (UCI) dataset. Using selected resources, prediction for chronic kidney disease is designed to identify the risk. Prediction results are grouped to identify the percentage of patients at high risk for kidney disease who are most likely to be diabetic. Using hierarchical grouping, three groups formed show that there is a strong relationship between chronic kidney and diabetes. Therefore, classification methods are a good solution because they provide a more accurate prediction about an individual health because it is a process that separates data into groups whose members have one or more characteristics in common. In addition, artificial intelligence using machine learning is an excellent working tool for health professionals in decision-making (Lenart, 2016).

Final Considerations

There is an extreme need to develop a new classification technique that can accelerate and simplify the process of diagnosing chronic diseases. According to research, it has been observed that CKD can be predicted using various classifiers in data mining as well as predicting disease stage using Artificial Intelligence. The different experiences observed have shown that most classifiers provide high accuracy values, above 90%, which can be implemented in an easy-to-run interface to assist physicians and healthcare professionals in decision making and the accuracy of patient outcomes of patients.Many technology companies, such as IBM, Apple, and Google, are investing heavily in healthcare analytics to make disease management easier. It is important to note that IAwill not replace physicians, but it is important for physicians to know how to use IA sufficiently to generate their hypotheses, perform "big data" analysis, optimize IA applications and software in clinical practice to bring the era of precision medicine.

REFERENCES

Abhishek, G.S.M.T and Gupta, D. 2012.Proposing Efficient Neural Network Training Model for Kidney Stone Diagnosis.*International Journal of Computer Science and Information Technologies,* vol. 3, 3900-3904.

AnanthaPadmanaban,K.R. and Parthiban, G. 2016.Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease.*Indian Journal of Science and Technology,* 2016, vol. 9(29).

Boukenze, B., Haqiq, A. and Mousannif, H. 2017.Predicting Chronic Kidney Failure Disease using Data Mining Techniques.*Advances in Ubiquitous Networking, Springer*, pp 701-712.

Boukenze, B., Mousannif, H. andHaqiq, A. 2016. Performance of Data Mining Techniques to Predict in Healthcare Case Study: Chronic Kidney Failure Disease. *International Journal of Database Management Systems* (IJDMS), vol.8, n.3.

Charleonnan, A., Fufaung, T., Niyomwong, F., Chokchueypattanakit, W., Suwannawach, S. and Ninchawee, N. 2016.Predictive analytics for chronic kidney disease using machine learning techniques.Management and Innovation Technology International Conference (MITicon), IEEE.

Chimwayi, K.B., Haris, N., Caytiles, R.D. and Iyengar, N.C.S. 2017.Risk Level Prediction of Chronic Kidney Disease Using Neuro- Fuzzy and Hierarchical Clustering Algorithm (s). *International Journal of Multimedia and Ubiquitous Engineering,* vol.12, n.8.

Chiu.R.K, Chen, R.Y, Wang, S, Jian, S.2012.Intelligent systems on the cloud for the early detection of chronic kidney disease.*Machine Learning and Cybernetics,* IEEE; p. 1737 – 1742.

Gabriel I. Barbas, Sherry A. Glied,2010.New Technology and Healthcare Costs - The Robot Assisted Surgery Case "; *The new England Journal of Medicine,* n. 363, p. 707-704.

Gant, V., Rodway, S. and Wyatt, J. 2001. Artificial neural networks: Practical considerations for clinical application. In R. Dybowski& V. Gant, Clinical applications of artificial neural networks (pp. 329-356). Cambridge, MA: Cambridge University Press.

Groves, P and Kayyali, B. 2013.The bigdate revolution in health.McKinsy and Company Health System Reform Center EUA. Available in: http://digitalstrategy.nl/wp-content/uploads/E2-2013.04-The-big-data-revolution-in-US-health-care- Accelerating-value-and-innovation.pdf.

Ho, C., Pai, T., Peng, Y., Lee, C., Chen, Y. 2012. Ultrasonography Image Analysis for Detection and Classification of Chronic Kidney Disease," IEEE Complex, Intelligent and Software Intensive Systems, p. 624 – 629.

Hsieh, J.W., Hung, C., Lee, Y., Chih, C., Shan Lee, W., Fen Chiang, H. 2014.Stage Classification in Chronic Kidney Disease by Ultrasound Image. International Conference on Image and Vision Computing New Zealand, ACM, p. 271-276.

Huang, T. and Lan, L. 2015. Promises and Challenges of Big Data Computing in Health Sciences. Big Data Research, vol. 2, pp 2-11.

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., Kitai, T. 2017 Artificial intelligence in precision cardiovascular medicine.*J. Am. Coll. Cardiol.*Vol. 69, n. 2657–2664.

Kusiaka, A., Dixonb, B. and Shah, S. 2005.Predicting survival time for kidney dialysis patients: a data mining approach. Computers in Biology and Medicine, p. 311 – 327.

Lenart, M., Mascarenhas, N., Xiong, R., Flower, A. 2016 Identifying Risk of Progression for Patients with Chronic Kidney Disease Using Clustering Models. *IEEE Systems and Information Engineering Design Conference* (SIEDS '16) p. 221-226.

Levey, A.S. and Coresh, J. 2012.Chronic kidney disease. Lancet, 379, pp 165–180.

Mandli, I. and Panchal M. 2014 Selection of Most Relevant Features from High Dimensional Data using IG-GA Hybrid Approach. *International Journal of Computer Science and Mobile Computing,* vol.3 Issue. 2, p. 827-830.

Manish, K., 2016.Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm.*International Journal of Compute Science and Mobile Computing,*vol 5, Issue 2, p. 24-33.

R. Weil, 2014.Big Data in Health: A New Era for Research and Patient Care Alan R. Weil. Health Affair, vol. 33, n. 7, pp 1110.

Rosmani, A., Mazlan, U., Ibrahim, A., Zakaria, D. 2015. I-KS: Composition of Chronic Kidney Disease (CKD) Online Informational Self-Care Tool. Computer, Communication, and Control Technology, IEEE, p. 379 – 383.

Singh, A., Nadkarni, G., Guttag, J. and Bottinger, E. 2014.Leveraging hierarchy in medical codes for predictive modeling.*Bioinformatics, Computational Biology and Health Informatics,* ACM, p. 96-103.

Steve G. Peters, James D. Buntrock, 2014.Big Data and the Electronic Health Registry.*Ambulatory Care Manage,* vol. 37, n. 3, pp. 206–210.

Tazin. N., Sabab, S.A, Chowdhury, M.T. 2016.Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique.*International Conference on Medical Engineering, Health Informatics and Technology* (MediTec).

Valderrama, M.A.E, Barroso, T.A., Roa, L.M, Hernández, D.N., Tosina, J.R., Fosalba N.A., Martín, J.A.M. 2014. A Distributed Approach to Alarm Management in Chronic Kidney Disease," IEEE Transl. Biomedical and Health Informatics, vol. 18, p. 1796 – 1803.

Vijayarani, S., Dhayanand, S. 2015. Data mining classification algorithm for kidney disease prediction.*International journal on cybernetic and information,* Volume 4, Issue 4, p.14-24.

Wang, H., Naghavi, M., Allen, C. 2016.GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and causespecific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet, 388, pp 1459–544.

****