CHRONIC KIDNEY DISEASE PREDICTION

[1]Dr.B.V.V.Siva Prasad, [2]Eslavath Latha, [3]Nithya Sairy, [4]Bhavana Tulasi

[1]*Associate Professor in School Of Engineering (CSE), Anurag University, Hyderabad.*
[2,3,4]*B.Tech Student in School Of Engineering (CSE), Anurag University, Hyderabad.*

Abstract:

One of the most serious illnesses nowadays is chronic kidney disease, for which a correct diagnosis must be made as soon as possible. The use of machine learning in healthcare has increased. The clinician can identify the ailment early with the use of machine learning classifier algorithms. This article has examined chronic kidney disease prediction from this angle. The dataset for chronic kidney disease was collected from the UCI repository. In this study, seven classifier algorithms were used, including the artificial neural network, C5.0, the logistic regression, the linear support vector machine with penalty L1 and L2 as well as the random tree. The dataset was also subjected to the significant feature selection approach. The results have been calculated for each classifier based on the following features: (i) full features; (ii) correlation-based feature selection; (iii) Wrapper method feature selection; (iv) least absolute shrinkage and selection operator regression; (v) synthetic minority over-sampling technique with least absolute shrinkage and selection operator regression selected features; and (vi) synthetic minority over-sampling technique with full features. From the findings, it can be seen that the synthetic minority over-sampling strategy with complete features uses LSVM with penalty L2 and achieves the maximum accuracy of 98.86%. The graph compares the outcomes of several algorithms and displays accuracy, precision, recall, F-measure, area under the curve, and GINI coefficient. The best results were obtained after using complete features and the synthetic minority over-sampling approach with least absolute shrinkage and selection operator regression. Again, the linear support vector machine provided the maximum accuracy of 98.46% in the synthetic minority over-sampling approach with the least absolute shrinkage and selection operator chosen features. On the same dataset, one deep neural network was used in addition to machine learning models, and it was found that this model had the greatest accuracy (99.6%).

I. Introduction

Your kidneys are damaged and not modifying your blood as they should if you have chronic kidney disease (CKD). If a person has CKD, it indicates that wastes are accumulated in the body since the fundamental function of the kidneys is to transform surplus water and waste from your blood to make urine. Because the harm was done gradually over a lengthy period of time, this illness is chronic. It is nice to say that it is a widespread ailment [1]. CKD may cause certain health issues. CKD has several causes, including diabetes, high blood pressure, and heart disease. In addition to these serious conditions, CKD is influenced by age and gender [2]. You may have one or more symptoms, such as back discomfort, nausea, diarrhoea, fever, nosebleeds, and vomiting, if your kidneys are not functioning properly. CKD is mostly caused by two illnesses: diabetes and high blood pressure [3]. Therefore, preventing CKD is achieved by the management of these two disorders. Typically, CKD does not show any symptoms until the kidney is severely damaged. According to research, the prevalence of CKD is fast rising, while the global death rate is staying the same [4]. Meanwhile, hospitalisation cases are rising by 6.23% annually. There are just a few diagnostic procedures to determine the stage of CKD:

Estimated glomerular filtration rate (eGFR), urine test, and blood pressure are the first three. Blood pressure is measured by a doctor because it reveals how well your heart is pumping blood. The patient has end-stage renal disease if the eGFR result is less than 15. Dialysis and kidney transplants are the only viable therapies at this time. Age, gender, the frequency and length of dialysis sessions, physical activity levels, and mental health are all factors that affect a patient's quality of life after dialysis [7]. The only option available to the doctor if dialysis is not feasible is kidney transplantation. But the price is astronomical [8]. As a result, it is crucial to recognise merit in early illness detection, monitoring, and management. Due to CKD's dynamic and mysterious nature in the early stages, it is crucial to accurately forecast its progression.

Depending on the stage, CKD requires medical therapy. If not, it is critical to identify the infection's organisational structure since it provides some clues. It supports the certainty of essential prayers and treatments. A major application area for intellectual intelligence systems is medical treatment [10]. In order to uncover hidden information from the extensive patient medical and treatment dataset that doctors regularly acquire from patients in order to get knowledge about the symptomatic data and to carry out accurate treatment plans, data mining may then play a significant role. Data mining is a technique for locating hidden information in a big dataset. Data mining techniques are interconnected and widely applied in many situations and fields. We can anticipate, categorise, filter, and cluster data using data mining techniques. The goal specifies how the algorithm will analyse a training set that includes a number of characteristics and goals. Data mining is appropriate for data mining in large datasets, however machine learning may also be used to perform it with a small dataset. Data analysis and pattern recognition are further capabilities of machine learning [9]. Machine learning methods are great at enhancing the accuracy of diagnostic prediction since there are several health datasets available [11]. Machine learning algorithms are becoming increasingly widespread in healthcare as the amount of electronic health records increases quickly. [12]. In order to diagnose CKD, Qin et al. In order to diagnose CKD, [13] suggested data assertion and a sample diagnosis are possible. Data assertion makes use of KNN. Logistic regression, random forest, support vector machine, K-nearest neighbour, naive Bayes classifier, and feed-forward neural network are six classifier algorithms that are used to determine the accuracy of a diagnosis. Random forest provides superior accuracy in these classifiers, at 99.75%.

II. Literature Review

On the basis of a dataset of 40000 cases, Vasquez-Morales et al. [14] created a neural network model with a 95% accuracy for risk prediction of the development of chronic kidney disease. Three models were used by Chen et al. [15] on the UCI dataset. They employed these classifiers to discover the patient's risk calculation using KNN, SVM, and soft independent modelling of class analogy (SIMCA). The SVM and KNN models both achieved the highest accuracy of 99.7%, while the SVM model is best able to withstand noise disruption. Due to the invasiveness and expense of CKD, many patients have reached their last stages without receiving therapy. Therefore, it is still crucial to find this condition early. Additionally, SVM machine learning classifier experiment results were provided by Amirgaliyev [16]. The use of machine learning classifier algorithms for the early diagnosis of CKD in diabetic patients was proposed by Padmanaban and Parthiban [17]. They used Naive Bayes and Decision trees to analyse the dataset after collecting the data from a diabetic research centre in Chennai. They utilised the Weka tool to measure accuracy and came to the conclusion that the Nave Bayes classifier had the highest accuracy (91%). De Almeida et al.'s [18] study utilised Support Vector Machines (SVM) using linear, polynomial, sigmoid, and RBF functions in addition to Decision Trees, Random Forests, and SVM. They made use of the MIMIC-II database for their investigation. They came to the conclusion that decision tree and random forest produced the greatest results, with respective prediction accuracy of 80% and 87%.

In order to determine which machine learning classifier method is most appropriate for the dataset, Gunarathne et al. [19] created a model of multiple classifier algorithms. They made use of a UCI-provided dataset with 400 instances and 14 characteristics. They came to the conclusion that the Multiclass Decision Forest method, with an accuracy of 99.1%, was best matched for the CKD dataset. SVM technique was utilised by Polat et al. [20] to predict CKD. They focused on a crucial component in order to get the right outcome. They used the two-approach Wrapper and slter method with the SVM algorithm to choose the correct feature. Sujata Drall, Gurdeep Singh Drall, Sugandha Singh, Bharat Drall, and others [21] worked with the 400 case, 25 attribute CKD dataset provided by UCI. Data was first preprocessed, missing data was located, filled in with 0, after which a transformation was done to the dataset. After preprocessing, authors utilised an algorithm for significant characteristics and identified the top five features before using Naive Bayes and K-Nearest Neighbour as their classification algorithms. KNN obtained the highest degree of accuracy. 400 occurrences and 25 characteristics from the CKD dataset were used by Almasoud and Ward [22]. Haemoglobin, albumin, and specific gravity were discovered to be feature attributes in the CKD dataset when they used the filter feature selection approach to attributes. Following the selection of the features, they trained the dataset and performed 10-fold cross-validation. The approach that attained the best accuracy, 99.1%, was gradient boosting.

Three processes were used by Shankar et al. [23] on the same UCI dataset: (i) data preprocessing & feature selection (ii), (iii) determining the correctness of the algorithms, and (iv) suggesting a diet. Two methods were used in the feature selection technique: the Wrapper method and the LASSO method. Four classification techniques were used after the feature selection method: Logistic Regression, Random Forest Tree K-Nearest Neighbours, Neural Network, and Wide and Deep Learning. The blood potassium level was used to advise a diet. Depending on its value, the blood potassium level was split into three categories: the safe zone, the caution zone, and the danger zone.

Kidney function test (KFT) dataset was gathered by Vijayarani and Dhayanand [24] from medical labs, research facilities, and hospitals. The dataset included 584 occurrences, 6 characteristics, and the support vector machine (SVM) and artificial neural network (ANN) classification techniques. It was discovered that ANN had the highest accuracy, coming in at 87.7%. With the use of 9 machine learning algorithms, including XGBoost, logistic regression, lasso regression, support vector machine, random forest, ridge regression, neural network, Elastic Net, and K-nearest neighbour, Xiao et al. [25] utilised the data from 551 patients. The linear model had the maximum accuracy, according to their evaluation of accuracy, ROC curve, precision, and recall. On the CKD Dataset, Reshma et al.'s [31] feature selection approach was applied. ACO approach was used to choose the features. The feature selection meta heuristic algorithm is called ACO. It is a Wrapper method type. There were a total of 24 characteristics in their dataset. Twelve features were utilised to create the model after the feature selection technique was used. The model was created using the Support Vector Machine Classifiers technique. Based on an outdated dataset of CKD, Deepika et al.

[32] developed a project for the prediction of chronic kidney disease. 24 characteristics and 1 target variable were included in the dataset. They used the KNN and Naive Bayes supervised machine learning algorithms to develop the model. KNN and Nave Bayes both obtained accuracy levels of 91% and 97%, respectively. Ma et al.'s [33] deep learning system was suggested for early Chronic Kidney Disease prediction. Heterogeneous Modified Artificial Neural Network Algorithm was used to create the deep neural network. The model was created using ultrasound pictures. Three distinct classifiers—the Support Vector Machine, artificial neural network, and multilayer perceptron—were used to compare the results. UI The machine learning model for early diabetic illness prediction was proposed by Haq et al. [34]. They came to the conclusion that machine learning can be very important in the medical field. Amin et

al.'s [35] machine learning model was presented for the early Parkinson's disease prediction. SVM classifier was employed in the model's construction. In order to extract the crucial features, feature selection algorithms like Relief and ACO were also used. The main goal of this study is to determine whether or not someone has chronic kidney disease. Seven different machine learning classifiers were used on the dataset for this perception. Both the entire features and the chosen features were active for each algorithm. All of the findings from the oversampling using SMOTE were recorded. One deep neural network technique was used to compare the outcomes of every machine learning model. Two hidden layers of a deep learning neural network were employed. In order to do computations, IBM SPSS Modeller was used. Applying deep neural networks to the dataset yields an accuracy estimate of 99.6%, according to the contribution.

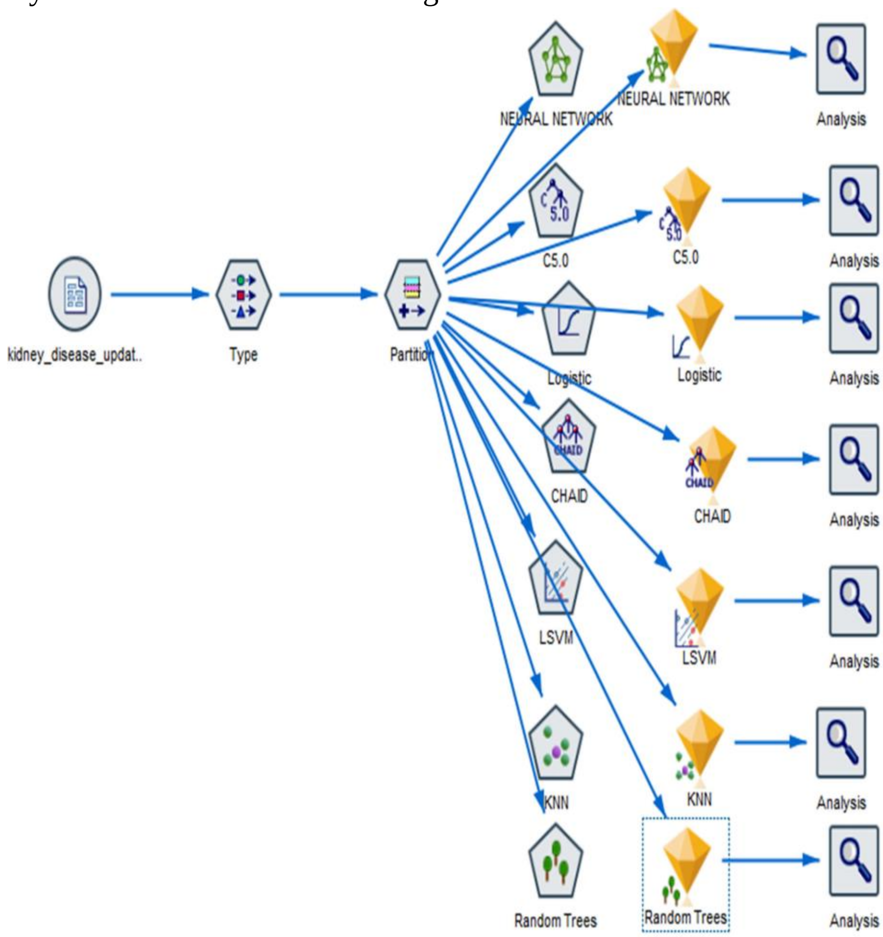
Sr. No	Attribute Name	Description
1	Age	Patient age (It is in years)
2	Bp	Patient blood pressure (It is in mm/HG)
3	Sg	Patient urine specific gravity
4	Al	Patient albumin ranges from 0-5
5	Su	Patient sugar ranges from 0-5
6	Rbc	Patient red blood cells two value normal and abnormal
7	Pc	Patient pus cell two value normal and abnormal
8	Pcc	Patient pus cell clumps two values present and not present
9	Ba	Patient bacteria two values present and not present
10	Bgr	Patient blood glucose random in mg/dl
11	Bu	Patient blood urea in mg/dl
12	Sc	Patient serum creatinine
13	Sod	Patient sodium
14	Pot	Patient potassium
15	Hemo	Patient hemoglobin (protein molecule in red blood cells)
16	Pcv	Patient packed cell volume % of red blood cells in circulating blood
17	Wc	Patient white blood cell counts in per microliter
18	Rc	Patient red blood cell count in million cells per microliter
19	Htn	Patient hypertension two value Yes and No
20	Dm	Patient diabetes mellitus two value Yes and No
21	Cad	Patient coronary artery disease two value Yes and No
22	Appet	Patient appetite two value good and poor
23	Pe	Patient pedal edema two value Yes and No
24	Ane	Patient anemia two value Yes and No
25	Class	Target Variable (CKD or Not)

III. RESULT OF WRAPPER FORWARD FEATURE SELECTION AND CLASSIFICATION

The Wrapper forward feature selection technique was used in this portion to choose the critical characteristics that would be passed to the classifier algorithms for result prediction. The six most crucial features—hemo, htn, dm, cad, pe, and al—were employed to identify outcomes. Hemo and htn are considered to be the most crucial variables for predicting chronic kidney disease by the CFS algorithm. Figure 9 displays the Wrapper algorithm's output. Table 6 summarises the outcome of each classification algorithm's performance. With the Wrapper algorithm, the C5.0 had the highest accuracy, with 96.1% accuracy, 98.55% precision, and 90.67% recall. The results from ANN, CHAID, and the random tree were all favourable. 94.63% accuracy, 90% precision, and 96% recall were attained with the ANN algorithm. 94.63 percent accuracy, 93.24% precision, and 92% recall were attained with the CHAID algorithm. 92% recall, 93.24% precision, and 94.63% accuracy were attained via the random tree algorithm. The accuracy, precision, and recall of the logistic regression approach were 78.54%, 98.55%, and 100% respectively. The accuracy, precision, and recall of the LSVM with Penalty L1 and Lambda 0.5 were 94.15%, 88.89%, and 96%, respectively. The accuracy, precision, and recall of the LSVM with Penalty L2 and Lambda 0.5 were 93.66%, 87.80%, and 96%, respectively. With a K value of 5 76, the KNN produced the poorest results for this dataset: 10% accuracy, 95.58% precision, and 95.58% recall. According to the outcome, LSVM had the highest AUC. Figure describes the contrast of precision, recall, and accuracy. The chart below displays a comparison of the GINI index.

IV. RESULT OF LASSO FEATURE SELECTION

The LASSO feature selection method was used in this portion to choose the critical features that would be accepted by the classifier algorithms for predicting the results. The six most crucial features—rbc, pc, al, ba, su, and pcc—were employed to find the results. According to the LASSO FS algorithm, the most crucial variables for predicting chronic kidney disease are rbc and pc. Figure 13 displays the outcome of the LASSO FS algorithm. Table 7 summarises the outcome of algorithm performance for each of the seven classifiers. The maximum accuracy was attained with LSVM and CHAID, 97.07%. The accuracy, precision, and recall of the LSVM with both penalty L1 and L2 were 97.07%, 98.59%, and 93.33%, respectively. The results of the CHAID algorithm.



92% recall, 100% precision, and 97.07% accuracy. 94.63% accuracy, 90% precision, and 96% recall were attained with the ANN algorithm. The accuracy, precision, and recall of the random tree method were each 90.24%, 80.90%, and 96%. The accuracy, precision, and recall of the logistic regression approach were 74.15%, 80.23%, and 100% respectively. The random tree method has a 96% recall rate, 78.26% precision, and 88.78% accuracy. With a K value of 5, the KNN produced the lowest results for this dataset: 56.59% accuracy, 92% precision, and 100% recall. According to the outcome, LSVM had the highest AUC. Fig. illustrates the contrast of precision, recall, and accuracy. Figure displays a comparison of the GINI index. The comparison of AUC is shown in the image.

V. RESULT OF SMOTE

As a consequence of the aforementioned finding, it was determined that the LASSO feature selection approach provided the maximum accuracy on the selected features. As a result, the LASSO regression method's chosen features and the entire features were both subjected to the SMOTE methodology. SMOTE evaluated the performance of ANN, CHAID, LSVM, and Random Tree. In every trial, these classification algorithms worked quite well. Because of the aforementioned finding, the performance of KNN and logistic regression on this dataset was not evaluated using the SMOTE approach.

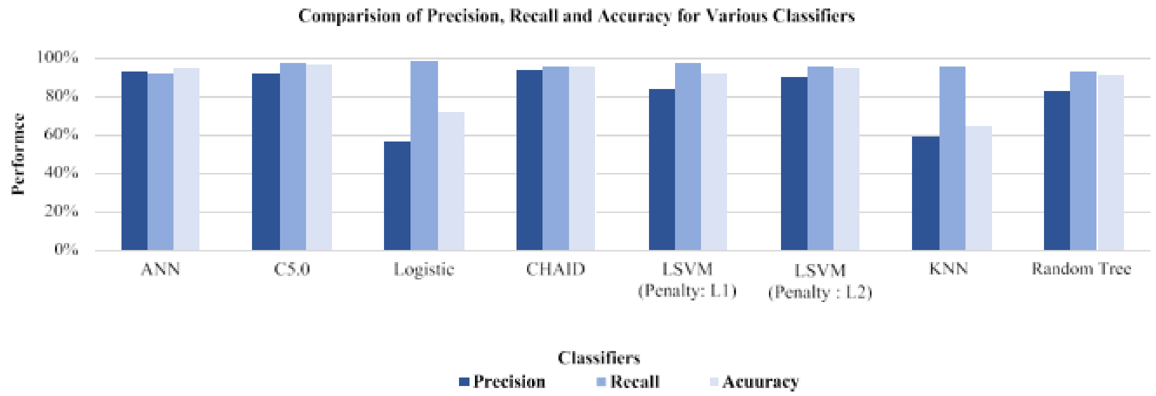


FIGURE 2. Comparison of precision, recall and accuracy for all classifiers without feature selections.

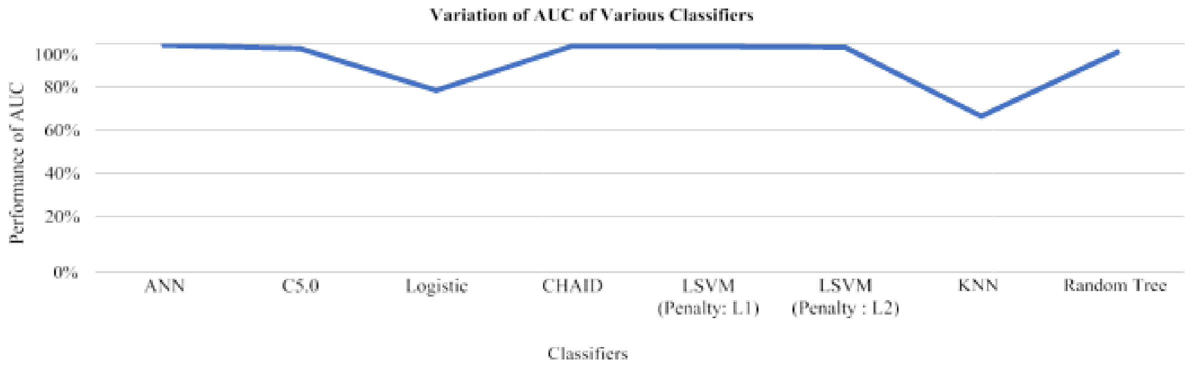
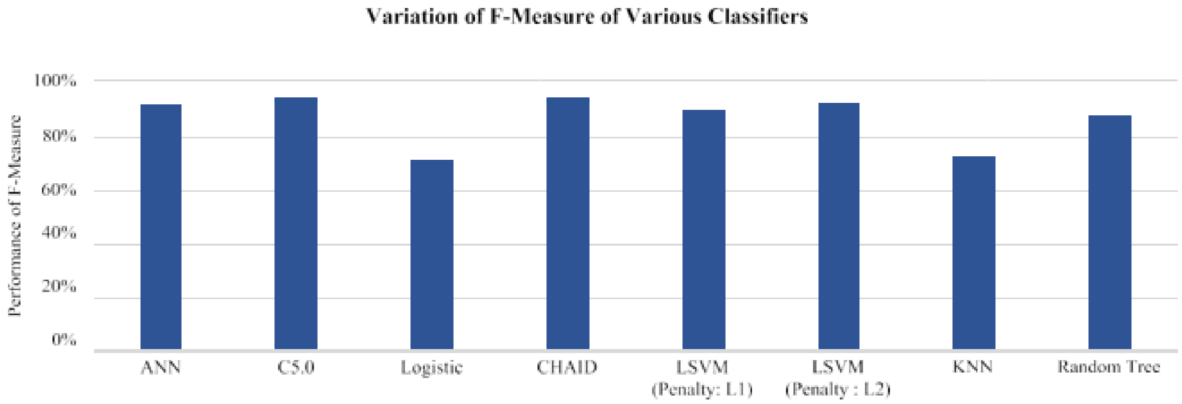


FIGURE 3. Performance of Area under curve for all classifiers without feature selections.



VI. CONCLUSION

This article criticises the use of the complete features and significant features of the CKD dataset to predict chronic kidney disease. The Wrapper approach, LASSO regression, and correlation-based feature selection have all been used for feature selection. Seven classifier algorithms, including the artificial neural network, C5.0, logistic regression, CHAID, linear support vector machine (LSVM), K-Nearest Neighbours, and random tree, were used to categorise this experience. Results for each classifier were calculated using complete features, CFS-selected features, Wrap-per-selected features, LASSO-regression-selected features, SMOTE with LASSO-selected features, and SMOTE with full features. In SMOTE with complete features, it was found that LSVM had the maximum accuracy of 98.86%. In tests using features chosen using LASSO regression both with and without SMOTE, all classifier methods showed good results. For all 5 classifiers, SMOTE with all characteristics produced the best results. Seven classifiers in all were utilised in this study. However, Logistic and KNN were not employed in SMOTE since they did not produce the desired results. According to the findings, SMOTE is the best strategy for balancing a dataset. It should be noted that the LASSO regression model performed better with SMOTE than it did without SMOTE when certain characteristics were included. In all studies, LSVM outperformed other classification algorithms in terms of accuracy.

REFERENCES

[1] Q.-L. Zhang and D. Rothenbacher, ‘‘Prevalence of chronic kidney disease in population-based studies: Systematic review,’’ BMC Public Health, vol. 8, no. 1, p. 117, Dec. 2008.

[2] W. M. McClellan, D. G. Warnock, S. Judd, P. Muntner, R. Kewalramani, M. Cushman, L. A. McClure, B. B. Newsome, and G. Howard, ‘‘Albumin- uria and racial disparities in the risk for ESRD,’’ J. Amer. Soc. Nephrol., vol. 22, no. 9, pp. 1721–1728, Aug. 2011.

[3] M. K. Haroun, ‘‘Risk factors for chronic kidney disease: A prospective study of 23,534 men and women in Washington County, Maryland,’’ J. Amer. Soc. Nephrol., vol. 14, no. 11, pp. 2934–2941, Nov. 2003.

[4] W. D. Souza, L. C. D. Abreu, L. G. D. SilvaI, and I. M. P. Bezerra, ‘‘Incidence of chronic kidney disease hospitalisations and mortality in Espírito Santo between 1996 to 2017,’’ Wisit Cheungpasitporn, Univ. Mississippi Medical Center, Rochester, MN, USA, Tech. Rep., 2019, doi: 10.1371/journal.pone.0224889.

[5] W. Mula-Abed, K. A. Rasadi, and D. Al-Riyami, ‘‘Estimated glomerular filtration rate (eGFR): A serum creatinine-based test for the detection of chronic kidney disease and its impact on clinical practice,’’ Oman Med. J., vol. 27, no. 4, pp. 339–340, 2012.

[6] A. S. Levey, D. Cattran, A. Friedman, W. G. Miller, J. Sedor, K. Tuttle, B. Kasiske, and T. Hostetter, ‘‘Proteinuria as a surrogate outcome in CKD: Report of a scientific workshop sponsored by the national kidney founda- tion and the US food and drug administration,’’ Amer. J. Kidney Diseases, vol. 54, no. 2, pp. 205–226, Aug. 2009.

[7] S. Gerogianni, ‘‘Concerns of patients on dialysis: A research study,’’Health Sci. J., vol. 8, no. 4, pp. 423–437, 2014.
[8] J. R. Chapman, ‘‘What are the key challenges we face in kidney transplan- tation today?’’ Transplantation Res., vol. 2, no. S1, pp. 1–7, Nov. 2013.
[9] T. Xiuyi and G. Yuxia, ‘‘Research on application of machine learn- ing in data mining,’’ in Proc. IOP Conf., Mater. Sci. Eng., 2018, doi: 10.1088/1757-899X/392/6/06220.
[10] B. Zupan, A. J. Halter, and M. Bohanec, ‘‘Qualitative model approach to computer assisted reasoning in physiology,’’ in Proc. Intell. Data Anal. Med. Pharmacol. (IDAMAP), Brighton, U.K., 2018, pp. 1–7.
[11] A. Dhillon and A. Singh, ‘‘Machine learning in healthcare data analysis: A survey,’’ J. Biol. Today’s World, vol. 8, no. 2, pp. 1–10, Jan. 2018.
[12] M. Ghassemi, T. Naumann, P. Schulam, A. L. Beam, I. Y. Chen, and R. Ranganath, ‘‘A review of challenges and opportunities in machine learn- ing for health,’’ in Proc. AMIA Joint Summits Transl. Sci., 2020, p. 191.
[13] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, ‘‘A machine learning methodology for diagnosing chronic kidney disease,’’ IEEE Access, vol. 8, pp. 20991–21002, 2020.
[14] G. R. Vasquez-Morales, S. M. Martinez-Monterrubio, P. Moreno-Ger, and J. A. Recio-Garcia, ‘‘Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning,’’ IEEE Access, vol. 7, pp. 152900–152910, 2019.
[15] Z. Chen, X. Zhang, and Z. Zhang, ‘‘Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate mod- els,’’ Int. Urol. Nephrol., vol. 48, no. 12, pp. 2069–2075, Jun. 2016.
[16] Y. Amirgaliyev, S. Shamiluulu, and A. Serek, ‘‘Analysis of chronic kidney disease dataset by applying machine learning methods,’’ in Proc. IEEE 12th Int. Conf. Appl. Inf. Commun. Technol. (AICT), Oct. 2018, pp. 1–4.
[17] K. R. A. Padmanaban and G. Parthiban, ‘‘Applying machine learning techniques for predicting the risk of chronic kidney disease,’’ Indian J. Sci. Technol., vol. 9, no. 29, Aug. 2016.
[18] L. Kilvia De Almeida, L. Lessa, A. Peixoto, R. Gomes, and J. Celestino, ‘‘Kidney failure detection using machine learning techniques,’’ in Proc. 8th Int. Workshop ADVANCEs ICT Infrastructures Services, 2020, pp. 1–8.
[19] W. Gunarathne, K. D. M Perera, and K. A. D. C. P Kahandawaarachchi, ‘‘Performance evaluation on machine learning classification techniques for disease classification and forecasting through data analytics for chronic kidney disease (CKD),’’ in Proc. IEEE 17th Int. Conf. Bioinf. Bioeng. (BIBE), Oct. 2017, pp. 291–296.
[20] H. Polat, H. D. Mehr, and A. Cetin, ‘‘Huseyin polat1 & homay danaei mehr1 & aydin cetin,’’ J. Med. Syst., vol. 41, no. 4, pp. 1–11, Apr. 2017, doi: 10.1007/s10916-017-0703-x.
[21] S. Drall, G. S. Drall, S. Singh, and B. B. Naib, ‘‘Chronic kidney disease prediction using machine learning: A new approach,’’ Int. J. Manage., Technol. Eng., vol. 8, pp. 278–287, May 2018.
[22] M. Almasoud and T. E. Ward, ‘‘Detection of chronic kidney disease using machine learning algorithms with least number of predictors,’’ Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 8, pp. 89–96, 2019.
[23] S. Shankar, S. Verma, S. Elavarthy, T. Kiran, and P. Ghuli, ‘‘Analysis and prediction of chronic kidney disease,’’ Int. Res. J. Eng. Technol., vol. 7, no. 5, May 2020, pp. 4536–4541.
[24] S. Vijayarani1 and S. Dhayanand, ‘‘Kidney disease prediction using SVM and ANN algorithms,’’ Int. J. Comput. Bus. Res., vol. 6, no. 2, pp. 1–12, Mar. 2015.
[25] J. Xiao, R. Ding, X. Xu, H. Guan, X. Feng, T. Sun, S. Zhu, and Z. Ye, ‘‘Comparison and development of machine learning tools in the prediction of chronic kidney disease progression,’’ J. Transl. Med., vol. 17, p. 119, Dec. 2019.

[26] M. S. Gharibdousti, K. Azimi, S. Hathikal, and D. H. Won, ‘‘Prediction of chronic kidney disease using data mining techniques,’’ in Proc. Ind. Syst. Eng. Conf., K. Coperich, E. Cudney, H. Nembhard, Eds., 2017, pp. 2135–2140.
[27] E. M. Karabulut, S. A. Ozel, and T. Ibrikci, ‘‘A comparative study on the effect of feature selection on classification accuracy,’’ Procedia Technol., vol. 1, pp. 323–327, Jan. 2012.
[28] A. Wosiak and D. Zakrzewska, ‘‘Integrating correlation-based feature selection and clustering for improved cardiovascular disease diagnosis,’’ Complexity, vol. 2018, Oct. 2018, Art. no. 2520706.
[29] N. A. Nnamoko, F. N. Arshad, D. England, J. Vora, and J. Norman, ‘‘Eval- uation of filter and wrapper methods for feature selection in supervised machine learning,’’ in Proc. 15th Annu. Postgraduate Symp. Converg. Telecommun., Netw. Broadcast., Liverpool, U.K., 2014, pp. 2–33.
[30] J. M. Pereira, M. Basto, and A. F. D. Silva, ‘‘The logistic lasso and ridge regression in predicting corporate failure,’’ Procedia Econ. Finance, vol. 39, pp. 634–641, Jan. 2016.
[31] P. G. Scholar, ‘‘Chronic kidney disease prediction using machine learn- ing,’’ Int. J. Eng. Res. Technol., vol. 9, no. 7, pp. 137–140, 2020.
[32] B. Deepika, ‘‘Early prediction of chronic kidney disease by using machine learning techniques,’’ Amer. J. Comput. Sci. Eng. Survey, vol. 8, no. 2, p. 7, 2020.
[33] F. Ma, T. Sun, L. Liu, and H. Jing, ‘‘Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artifi- cial neural network,’’ Future Gener. Comput. Syst., vol. 111, pp. 17–26, Oct. 2020.
[34] A. U. Haq, J. P. Li, J. Khan, M. H. Memon, S. Nazir, S. Ahmad, G. A. Khan, and A. Aliss, ‘‘Intelligent machine learning approach for effective recogni- tion of diabetes in E-healthcare using clinical data,’’ Sensors, vol. 20, no. 9, p. 2649, May 2020.
[35] U. H. Amin, J. Li, Z. Ali, M. H. Memon, M. Abbas, and S. Nazir, ‘‘Recognition of the Parkinson’s disease using a hybrid feature selection approach,’’ J. Intell. Fuzzy Syst., vol. 39, no. 1, pp. 1–21, Jul. 2020