Sai Ram Yadla

by OfficialTurnitin

Submission date: 28-Apr-2024 05:25PM (UTC+0900)
Submission ID: 2364050084
File name: final-paper.pdf (1.14M)
Word count: 3952
Character count: 22594

A machine learning-based Predictive Framework was employed for Chronic Kidney Disease Prognosis.

Yadla Sai Ram[1], Ramakrishna Badiguntla[2], Dulam Balu[3], Mundlapati Ramanjaneyulu[4]

[1,3,4]Department of Computer Science and Business System, RVR & JC College of Engineering, Guntur, Andhra Pradesh 522019, India

[2]Assistant Professor, Department of CSE (Data Science), RVR & JC College of Engineering, Guntur, Andhra Pradesh 522019, India

Corresponding author: Yadla Sai Ram (e-mail: sairamyadla88@gmail.com).

ABSTRACT Chronic Kidney Disease (CKD) or Long-term renal disease has turned into a serious issue with the constant rise in the number of patients. Without kidneys, one can live only for 18 days on average, thus kidney disease treatment is very much in demand. There is a need to have effective ways of early foreseeing CKD. This research intends the creation and verification of a CKD prediction model. Medical professionals frequently use machine learning algorithms to diagnose and categorize patients. Gradient Boosting algorithms are used within the proposed system. This proposed model will be useful for predicting future CKD and non-CKD patients according to several criteria with an impressive accuracy level of 98.75%. One example would be designing a machine learning algorithm that divides CKD into stages, using eGFR as the dominant metric for staging CKD progressions.

INDEX TERMS Chronic Kidney Disease (CKD); Predictive model; Machine learning algorithms; Gradient Boosting; CKD progression; Early prediction; Severity stages

I. INTRODUCTION

Currently, machine learning techniques and algorithms are being developed by engineers and healthcare professionals in an effort to identify chronic kidney illness early. This is proving difficult because healthcare data is widely spread out and highly complex necessitating complicated analysis of the same. However, by making use of the technology of data mining, this data can be formatted in the form of a machine learning algorithm. This condition has led to many deaths globally, claiming lives each year as well as resulting in kidney failure [1]. A study done by Study of the Worldwide burden of Diseases (GBDS) 2010 found persistent renal disease moved from its rank at position 27th in 1990 to an alarming rate of 18th most common cause of global death [2]. People with this condition have an elevated risk of developing advanced renal disease, requiring expensive therapeutic interventions such as hemodialysis and surgery. [3]. Kidney illness, which affects hundreds of thousands of people, thus stands out as a serious public health concern. across Ethiopia irrespective of age or sex [4]. This is because chronic kidney disease has no symptoms most of the time until a decline in normal kidney function to about 15-20% [5]. In patients with advanced CKD, key clinical features include fatigue and lack of energy, difficulty concentrating, loss of appetite, insomnia, night leg cramps, leg and ankle swelling (edema), the appearance of puffiness under the eyes, itchy, dry skin, and nighttime urine [6]. Despite over two million individuals requiring dialysis or kidney transplants due to renal failure, which causes more deaths than breast and prostate cancer combined, a hormone that is produced by kidneys, controls several physiological processes like erythropoiesis, regulation of blood pressure as well as calcium metabolism too.

II. RELATED WORK

A neural network framework was built by Vasquez-Morales and his co-authors [7] to estimate the likelihood of developing chronic renal disease using 40,000 instances in their dataset. Their model indicated a level of accuracy at 95%.

Machine Learning classifier methods have been proposed by Padmanaban and Parthiban [8] as a means of assisting

diabetic patients in identifying persistent kidney disease at earlier stages. A diabetic research center in Madras supplied the data, which was subsequently processed using the Naive Bayes (NB) and Decision Tree (DT) algorithms. The Weka tool yielded an accuracy rating of 91% for the NB classifier, which was the most accurate

In a study by S. Drall, G. S. Drall, S. Singh, B. B. Naib, and their team [9], a CKD dataset from UCI was examined that consisted of 400 instances including 25 attributes. The authors employed data preprocessing wherein missing values were found and substituted with zero; before then transforming and applying it into the data set. Afterward, they identified the top five features using an algorithm that discriminated significant attributes and later applied classification algorithms; namely Naïve Bayes KNN. Analysis showed that the highest accuracy was obtained applying the K-Nearest Neighbour approach.

Deepika et al. [10] developed a project where they used the dataset specific to CKD. The dataset was historical, CHD-related consisting 24 features and one intended variable. During their model building process, they employed two supervised Specifically, KNN and NB, are ML algorithms. It is worth mentioning that the peak accuracy of KNN was 97% whereas A 91% accuracy rate was obtained with Naïve Bayes.

In the study of comparison, Charleonnan and her colleagues [11] used predictive models which included LR, SVM, KNN and DT. These Indian CKD-specific datasets were made to find the best classifier for CKD prediction. SVM was found to be the most accurate model with an accuracy of 98.3% and sensitivity of 0.99.

In Tekale et al. [12] A group of 400 examples and 14 features made up the information set. DT and SVM are the approaches applied here. Following preprocessing, the data set had just 14 characteristics, as opposed to the initial 25. The most accurate model was Support Vector Machine, that had a success rate of 96.75%.

Yashfi et al. [13] introduced a methodology where Utilising ML techniques, they forecasted the probability of Long-term renal disease, and data from CKD patients. The analysis was done using RF and ANN. From 25 features in the beginning, 20 for each were chosen and both RF and ANN were applied. RF showed maximum accuracy which amounted to 97.12%.

Pal, Saurabh et al. [14] attempted to discuss machine learning approaches towards predicting chronic kidney disease thereby providing meaningful insights into how biomedical predictive tools are used. Numerous algorithmic methods, including decision tree classifiers, LR, SVM, and others, were used among others but it is observed that decision trees have the highest accurate diagnosis.

Debal, Adeba, and Sitote, the authors, decided to explore the utilization of ML methods. in prognosticating Long-term renal disease. Among its techniques were logistic regression, random forests, and gradient boosting [15]. When it came to accurately predict the possibility of long-term renal damage, the decision tree was the most accurate method, followed by random forests and neural networks in that order (Dritsas & Trigka; [16]).

The forecast of chronic kidney disease was investigated by Almustafa [17] using different classification algorithms thus enhancing medical informatics. DT, KNN, and SVM were part of the analysis to determine which algorithm performed better in predicting the disease, indicating that support vector machines had a higher accuracy rate.

According to Ifraz et al [18], their research compared intelligent approaches like artificial neural networks to predict kidney disease with other techniques that are described in computational healthcare literature. This study looked at different algorithms such as DT, LR, and ANN showing that increased accuracy was obtained by artificial neural networks in disease prediction.

Chittora et al. [19] gave their perspective on how biomedical informatics has been influenced by the machine learning aspect of CKD prediction. The study used various algorithms such as DT, RF and Gradient Boosting (GB) where it was found that GB possessed the greatest degree of sensitivity in identifying the illness.

Walse et al. [20] also contributed to this area by exploring suitable applications of methods such as NB, RF while analyzing CKD thereby promoting machine learning for information in medical. In this regard, among these techniques which randomly select subsets from the original dataset for classification problems; a better performance was noted for random forest during disease analysis.

To further medical image analysis, Nithya et al. [21] also explored multi-kernel k-means clustering methods and neural networks for kidney disease classification and recognition in ultrasound pictures. With artificial neural networks being more accurate than previous techniques for illness identification and segmentation, they employed multi-kernel k-mean clustering techniques in conjunction with neural networks.

Navaneeth and Suchetha [22] introduced a method that was new based on dynamic pooling in convolutional neural networks for identifying persistent renal failure and advancing deep learning's boundaries in biomedical signal analysis. For

example, the study employed convolutional neural networks where the proposed dynamic pooling methodology proved more efficient than other methods when detecting diseases.

Aqlan et al [23] utilized data mining strategies as a way to predict chronic kidney disease which offered some insights into how industrial and systems engineering could be applied in healthcare. The investigation involved the use of diverse algorithms that may embrace neural networks besides artificial neural network factors among others. In addition to this, probably this research also looked at the potential marker of persistent kidney damage is the estimated glomerular filtration rate, or eGFR.

III. METHODOLOGY

The research involves the usage gradient boosting, one type of machine learning algorithm and voting classifiers Utilizing a dataset sourced from the University of California, Irvine Machine Learning Repository focusing on chronic renal disorders, the study allocated 80% of the data for training and 20% for testing the model. Our research methodology is centered on the creation of a machine learning model specifically designed for classifying the stages of severity of chronic renal disease with eGFR serving as the prime metric for stage classification in CKD progression. Clinical parameters and eGFR measurements are included in a large dataset to train and validate our model. We used machine learning algorithms, such as Gradient Boosting, to find the most effective one for CKD stage classification. When this research is complete, it will produce a powerful, accurate machine learning approach that distinguishes patients' GFR values by their CKD severity stages to enhance healthcare decisions and patient management.

The Figure 1. is a structured method for predicting kidney disease which starts with the acquisition of the "Kidney_disease.csv" dataset expected to have relevant data about kidney disease parameters. After acquiring this data, pre-processing activities will be conducted; these include handling missing values and categorical value encoding to guarantee that the data is in line with machine learning techniques and that its confidentiality is not threatened. The next step involves dividing the dataset into separate subgroups meant for training and testing purposes. The model training phase employs Gradient Boosting as a robust machine learning technique in developing GB Classifier with the ultimate goal of estimating the likelihood of developing long-term kidney disease or not based on user inputs.

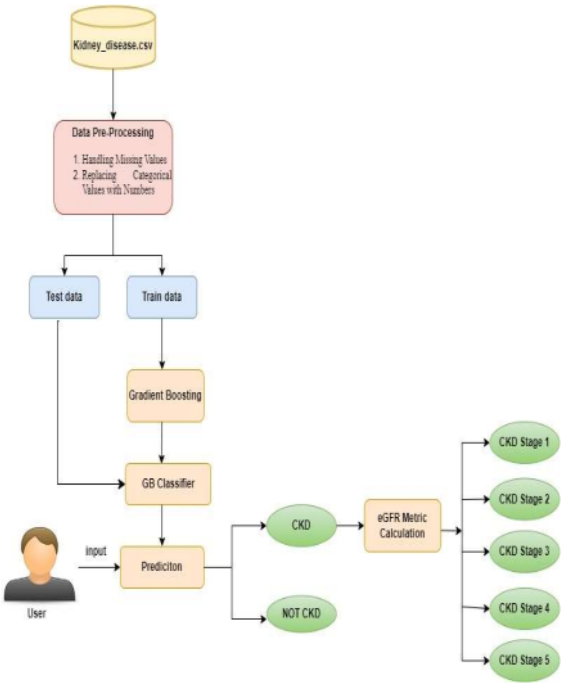

Figure 1. Block Diagram

CKD prediction involves an assessment phase where the estimated Glomerular Filtration Rate (eGFR) is calculated, categorizing the five stages from CKD Stage 1 to CKD Stage 5.

Stage 1: Kidney injury with a GFR that is either normal or elevated (>90 mL/min/1.73 m²)

Stage 2: GFR decline to a mild level (60-89 mL/min/1.73 m²)

Stage 3a: GFR declines moderately (45–59 mL/min/1.73 m²)

Stage 3b: GFR declines moderately (30–44 mL/min/1.73 m²)

Stage 4: Significant decline in GFR (15–29 mL/min/1.73 m²)

Stage 5: Kidney failure (GFR < 15 mL/min/1.73 m²)

Figure 2. CKD Stages

A. DATASET

We gathered the CKD dataset through the ML Repository at University of California, Irvine in order to assess our recommended model. This data set consists of 400 samples that were used in this study to evaluate and validate our model

[24]. Each sample has 24 predictors which includes 11 continuous and categorical attributes. Furthermore, the category respond variable 'class' exists. indicating whether CKD is present or not in the dataset. The variable 'class' contains two possible values: "CKD" stands for cases of individuals who have been diagnosed with chronic kidney disease; "notckd" is for people who do not. In order to offer more clarification, Table 1 presents a descriptive summary of the qualities that were included in our thorough investigation.

S. No	Attribute	Description	Type
1	Age	Age of the patient	Numerical
2	bp	Blood Pressure (mm/Hg)	Numerical
3	sg	Specific Gravity (range 1.005 to 1.030)	Numerical
4	al	Albumin (0-5)	Numerical
5	su	Sugar (0-5)	Numerical
6	rbc	Red Blood Cells (Normal or Abnormal)	Categorical
7	pc	Pus Cell (Normal or Abnormal)	Categorical
8	pcc	Pus Cell Clumps (Present or Not Present)	Categorical
9	ba	Bacteria (Present or Not Present)	Categorical
10	bgr	Blood Glucose Random (mgs/dL)	Numerical
11	bu	Blood Urea (mgs/dL)	Numerical
12	sc	Serum Creatinine(mgs/dL)	Numerical
13	sod	Sodium (mEq/L)	Numerical
14	pot	Potassium (mEq/L)	Numerical
15	hemo	Hemoglobin (gms)	Numerical
16	pcv	Packed Cell Volume (percentage)	Numerical
17	wc	White Blood Cell Count (cells/cubic mm)	Numerical
18	Rc	Red Blood Cell Count (millions/cubic mm)	Numerical
19	htn	Hypertension (Yes or No)	Categorical
20	dm	Diabetes Mellitus (Yes or No)	Categorical
21	cad	Coronary Artery Disease (Yes or No)	Categorical
22	appet	Appetite (Good or Poor)	Categorical
23	pe	Pedal Edema (Present or Not Present)	Categorical
24	ane	Anemia (Present or Not Present)	Categorical
25	class	Class (target variable)	Categorical

Figure 3: Data Description

B. PREPROCESSING

Data preprocessing is a very important method of refining unprocessed data to be used in machine learning classifiers. This process entails filling in the missing data, changing categorical variables into numerical equivalents, making sure that attributes are uniform by rescaling, transforming data into binary format, and standardization. Rescaling is very vital in handling varying scales across attributes whereas binary transformation converts data into binary format [25]. It is advantageous to replace missing values with feature medians particularly when considering large datasets. Conversion of categorical values into their numeric counterparts enhances interpretability and enables machine learning algorithms to analyze such information which boosts performance within various applications.

C. TECHNIQUES FOR CLASSIFICATION

Supervised learning relies heavily on the process of classification. The training dataset provides details that classifiers use to identify the target characteristic on the testing dataset. Here are several research-related categorization strategies. Using a combination of conventional machine learning methods, ensemble learning serves to augment the predictive prowess of a model. Among these approaches, boosting algorithms stand out as particularly potent in the machine learning domain. In this experimental endeavor, we employ gradient boosting, specifically tailored for chronic kidney disease (CKD) prediction.

Gradient Boosting:

Gradient Boosting (GB) classifiers are machine learning methods that combine several weak learners in order to create a powerful predictor, often by using decision trees. The general idea behind it is that combining subsequent models with the earlier ones will reduce the sum of prediction errors. However, a key idea within this approach is about specifying what is expected from subsequent models to minimize these errors. Moreover, the gradient of error in prediction provides targets for each instance. Consequently, at every iteration, models attempt to correct their predictions on individual training cases and decrease predictive mistakes as much as possible [26].

In this research paper, the Gradient Boosting approach is formulated as follows:

1. Model Initialization:

a) To initialize the model, commence with a first prediction that is usually set at the average of target values. This model is then saved as an “initial_model”.

2. Boosting Method:
To increase First execute an iterative process that covers every step of the boosting process and does the following:

a) Determine residuals for the current model.
b) Train a weak learner only on residuals.
c) Sum all previous predictions multiplying them by their learning rates from weak learners.
d) Incorporate weak learners that have been trained for the ensemble model.

3. The Last Model-Based Predictions:

a) It starts with predictions of the first models and goes through each weak learner in the ensemble one after another until
b) Adjusted Present Forecasts for a Specific Weak Learner.

4. Last Output of Prediction:

The final prediction is made after an iterative boosting procedure involving all the weak learners in the ensemble.

D.KIDNEY SEVERITY STAGE CLASSIFICATION

One aspect of renal function is eGFR. It is an approximation of the volume of blood that flows via glomeruli, tiny kidney units, in a minute. Milli litres per minute (mL/min/1.73 m²) is the unit of measurement for eGFR [27].

While there are alternative formulae available, the MDRD and CKD-EPI equations are frequently used to estimate GFR. These algorithms occasionally take into account variables including age, gender, serum creatinine level, and race.

The following is the simplified MDRD equation which can be often seen in clinical practice:

$$eGFR = 175 \times Age^{(-0.203)} \times SerumCreatinine^{(-1.154)} \quad \text{Eq (1)}$$

Where:

- Serum Creatinine: Blood creatinine levels, which are usually measured in milligrammes per deciliter, or mg/dL.
- Age: A person's age expressed in years.

GFR here is estimated through this equation for milliliters per minute and 1.73 m². However, it should be noted that the MDRD equation is commonly used but may not be accurate among different groups like those with extreme muscle mass, elderly people, or some other illnesses.

In order to diagnose and stage chronic kidney disease (CKD), the eGFR must be ascertained. Stages of chronic kidney disease are determined by the eGFR measurement.

eGFR >= 90 mL/min/1.73 m² in Stage 1.
89 >= eGFR >= 60 mL/min/1.73 m² in Stage 2.
59 >=eGFR >= 45 mL/min/1.73 m² in Stage 3a.
44 >= eGFR >= 30 mL/min/1.73 m² in Stage 3b.
29 >= eGFR >= 15 mL/min/1.73 m² in Stage 4.
eGFR less than 15 mL/min/1.73 m² in Stage 5.

Figure 4. CKD Stage Classification

It is of great importance to establish the extent and progression of renal damage so that these patients may live longer and healthier lives. Bear in mind, however, that eGFR could give information on renal function but it is just a part of the full picture. Other factors for assessing kidney state are urine albumin levels, clinical symptoms, and medical history among others. Therefore, interpretation of eGFR should involve other clinical outcomes.

IV. RESULTS AND DISCUSSION

The study entailed testing on the UCI CKD dataset, where missing values were imputed using a variety of imputation strategies and categorical descriptors were encoded. In order to forecast instances of persistent kidney disease, boosting algorithms are used in the experimental approach, as described in the paper that accompanies. While gradient boosting algorithms were trained using 80% of the datasets, their effectiveness was assessed and confirmed using the remaining 20%. The performance indicators that are carefully examined in their future evaluations include F1-score, accuracy, recall, and precision [28].
Performance metrics are:

Confusion Matrix:

A table that displays a model's performance in classification is called a confusion matrix. By grouping the predictions into four groups—true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN)—it summarises the proportion of accurate and inaccurate forecasts. This is how a confusion matrix is structured:
In simpler terms;

- True Positive (TP): the amount of accurately anticipated positive cases.
- True Negative (TN): the amount of accurately anticipated negative situations.
- False Positive (FP): The quantity of cases that were incorrectly assumed to be positive (Type I mistake).
- False Negative (FN): Count of cases where a Type II mistake led to an incorrect negative prediction.

Accuracy: In the framework of CKD, accuracy measures how well instances categorized as CKD-positive or CKD-negative are done.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \qquad \text{Eq (2)}$$

Recall: Recall measures CKD positive cases that have been correctly labeled as such with respect to all CKD cases in the data set.

$$Recall = \frac{TP}{TP+FN} \times 100 \qquad \text{Eq (3)}$$

Precision: Precision will look at what percentage of people with CKD are truly diagnosed as being so.

$$Precision = \frac{TP}{TP+FP} \times 100 \qquad \text{Eq (4)}$$

F1-Score: A statistic called the F1 score is used to assess how well a classifier performs in detecting instances that test positive for CKD while taking accuracy and recall into account.

$$F1-score = 2 \times \frac{precision \times recall}{precision+recall} \times 100 \quad \text{Eq (5)}$$

Classification Report:

```
Accuracy: 0.9875
Confusion Matrix:
 [[40  1]
 [ 0 39]]
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.98      0.99        41
           1       0.97      1.00      0.99        39

    accuracy                           0.99        80
   macro avg       0.99      0.99      0.99        80
weighted avg       0.99      0.99      0.99        80
```

Figure 5. Classification Report

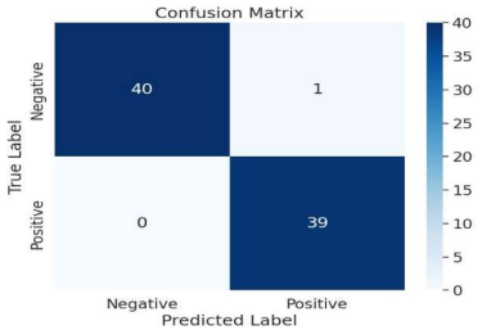

Figure 6. Confusion Matrix

Figure 5 illustrates how diverse configurations done on individual model parameters lead to different outcomes. This presents the results derived from experiments performed on each model were represented based on metrics.

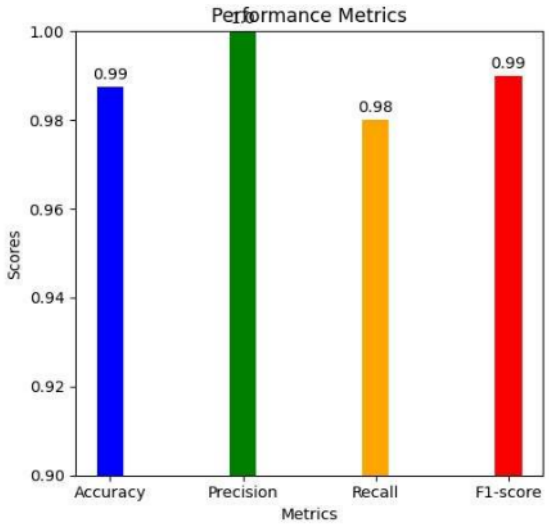

Figure 7: Results

The gradient boosting algorithm demonstrated strong performance with an accuracy of 98.75%, precision of 100%, recall of 98%, and F1 score of 99%.

Reference	Accuracy	Precision	Recall	F1-Score
Pal Saurabh et al. [16]	97%	0.99	0.98	0.98
Debal, Adeba, and Sitote et al. [17]	96%	0.97	0.94	0.96
Dritsas & Trigka et al. [18]	97.4%	0.97	0.97	0.97
Almustafa et al. [19]	95.75%	0.96	0.95	0.95
Chittora et al. [21]	96.4%	0.981	0.913	0.964
Proposed Model	98.75%	1.00	0.98	0.99

Table 1: Comparison with Other Models

V. CONCLUSION AND FUTURE WORK

A CKD prognosis model was developed using a gradient boosting approach. Recognizing it in its early phases is one of its skills. When measured using criteria like recall, precision, F1 score, and accuracy, the model performs well. It can grow increasingly adept at identifying people who could be vulnerable when both category and non-categorical features are included. With immediate action and preventive steps, this approach may benefit patients as well as medical professionals. To improve this system's clinical utility, feature selection approaches must be used to further develop it. In order to choose the ideal subset of characteristics for model building, this works combined feature selection approaches with a guided machine-learning technique. Using unsupervised or deep learning techniques might be helpful in identifying performance variations. Because the suggested model makes expert conclusions more quickly, it may be possible to use it on mobile devices for self-assessment and real-time tracking of patients.

REFERENCES

[1] A. S. Levey, R. Atkins, J. Coresh, et al., "Chronic kidney disease as a global public health problem: approaches and initiatives–a position statement from kidney disease improving global outcomes," Kidney International, vol. 72, no. 3, pp. 247–259, 2007.

[2] V. Jha, G. Garcia-Garcia, K. Iseki, et al., "Chronic kidney disease: global dimension and perspectives," The Lancet, vol. 382, no. 9888, pp. 260–272, 2013.

[3] M. J. Lysaght, "Maintenance dialysis population dynamics: current trends and long-term implications," Journal American Society Nephrology, vol. 13, suppl 1, pp. S37–S40, 2002.

[4] Stanifer JW, et al. The epidemiology of chronic kidney disease in sub-Saharan Africa: A systematic review and meta-analysis. Lancet Glob Heal. 2014;2(3):e174–81.

[5] CKD. Available online: https://www.urologyhealth.org/urology-a-z/k/kidney-(renal)-failure (accessed on 27 June 2022).

[6] Abdel-Kader, K. Symptoms with or because of Kidney Failure? Clin. J. Am. Soc. Nephrol. 2022, 17, 475–477.

[7] 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.

[8] 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.

[9] 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.

[10] B. Deepika, "Early prediction of chronic kidney disease by using machine learning techniques", Amer. J. Comput. Sci. Eng. Survey, vol. 8, no. 2, pp. 7, 2020

[11] Charleonnan A, Fufaung T, Niyomwong T, Chokchueypattanakit W, Suwannawach S, Ninchawee N. Predictive analytics for chronic kidney disease using machine learning techniques. Manag Innov Technol Int Conf MITiCON. 2016;80–83:2017.

[12] Tekale S, Shingavi P, Wandhekar S, Chatorikar A. Prediction of chronic kidney disease using a machine learning algorithm. Disease. 2018;7(10):92–6.

[13] Yashfi SY. Risk Prediction Of Chronic Kidney Disease Using Machine Learning Algorithms. 2020.

[14] Pal, Saurabh. "Chronic kidney disease prediction using machine learning techniques." Biomedical Materials & Devices 1.1 (2023): 534-540.

[15] Debal, Dibaba Adeba, and Tilahun Melak Sitote. "Chronic kidney disease prediction using machine learning techniques." Journal of Big Data 9.1 (2022): 109.

[16] Dritsas, Elias, and Maria Trigka. "Machine learning techniques for chronic kidney disease risk prediction." Big Data and Cognitive Computing 6.3 (2022): 98.

[17] Almustafa, Khaled Mohamad. "Prediction of chronic kidney disease using different classification algorithms." Informatics in Medicine Unlocked 24 (2021): 100631.

[18] Ifraz, Gazi Mohammed, et al. "Comparative analysis for prediction of kidney disease using intelligent machine learning methods." Computational and Mathematical Methods in Medicine 2021 (2021).

[19] Chittora, Pankaj, et al. "Prediction of chronic kidney disease-a machine learning perspective." IEEE access 9 (2021): 17312-17334.

[20] Walse, Rajesh S., et al. "Effective use of naïve Bayes, decision tree, and random forest techniques for analysis of chronic kidney disease." Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020, Volume 1. Springer Singapore, 2021.

[21] Nithya, A., et al. "Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images." Measurement 149 (2020): 106952.

[22] Navaneeth, Bhaskar, and M. Suchetha. "A dynamic pooling based convolutional neural network approach to detect chronic kidney disease." Biomedical Signal Processing and Control 62 (2020): 102068.

[23] Aqlan, Faisal, Ryan Markle, and Abdulrahman Shamsan. "Data mining for chronic kidney disease prediction." IIE Annual Conference. Proceedings. Institute of Industrial and Systems Engineers (IISE), 2017.

[24] Rubini, L.; Soundarapandian, P.; Eswaran, P. Chronic Kidney Disease. UCI Machine Learning Repository. 2015. Available online: https://archive.ics.uci.edu/dataset/336/chronic+kidney+disease (accessed on 10 June 2023).

[25] García, S.; Luengo, J.; Herrera, F. Data preprocessing in data mining. CA Cancer J. Clin. 2015, 72, 59–139.

[26] Fatima M., Pasha M. Survey of machine learning algorithms for disease diagnostic. J Intel Learn Syst Appl. 2017;9(01)

[27] 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.

[28] Moccia, S.; De Momi, E.; El Hadji, S.; Mattos, L.S. Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics. Comput. Methods Programs Biomed. 2018, 158, 71–91.

Sai Ram Yadla

ORIGINALITY REPORT

7%	5%	5%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ratna Kumari Anantha, Sobha Kota. "Bio-composites for the sorption of copper from aqueous solution: A comparative study", Groundwater for Sustainable Development, 2018 Publication	1%
2	Md. Ariful Islam, Md. Ziaul Hasan Majumder, Md. Alomgeer Hussein. "Chronic kidney disease prediction based on machine learning algorithms", Journal of Pathology Informatics, 2023 Publication	1%
3	Submitted to Coventry University Student Paper	1%
4	www.ijraset.com Internet Source	1%
5	Submitted to University of West London Student Paper	<1%
6	link.springer.com Internet Source	<1%

7	Youness Khourdifi, Mohamed Bahaj. "Applying Best Machine Learning Algorithms for Breast Cancer Prediction and Classification", 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 2018 Publication	<1 %
8	Bryan Chandra, Grace Oktaviani Kindy, Kelvin Susanto Gunawan, Gavin Putra Satria, Ivan Sebastian Edbert, Derwin Suhartono. "Deep Learning Architectures for Facial Emotion Analysis", Proceedings of the 8th International Conference on Sustainable Information Engineering and Technology, 2023 Publication	<1 %
9	dokumen.pub Internet Source	<1 %
10	ijsrcseit.com Internet Source	<1 %
11	Md. Mustafizur Rahman, Md. Al-Amin, Jahangir Hossain. "Machine learning models for chronic kidney disease diagnosis and prediction", Biomedical Signal Processing and Control, 2024 Publication	<1 %
12	www.ijfmr.com Internet Source	

<1 %

13 Daravath Anil, Shaik Naimudden, Aujugari Santosh Reddy, A Lavanya. "Prediction of Chronic Kidney Disease Using Various Machine Learning Algorithms", 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), 2023
Publication
<1 %

14 S L Parvathi Sitanaboina, Satyanarayana Reddy Beeram, Harikiran Jonnadula, Lakshmikanth Paleti. "Attention 3D-CU-Net: Enhancing Kidney Tumor Segmentation Accuracy Through Selective Feature Emphasis", IEEE Access, 2023
Publication
<1 %

15 etd.astu.edu.et
Internet Source
<1 %

16 researchinventy.com
Internet Source
<1 %

17 www.medicalnewstoday.com
Internet Source
<1 %

18 Amir Rehman, Huanlai Xing, Mehboob Hussain, Nighat Gulzar, Muhammad Adnan Khan, Abid Hussain, Sajid Mahmood. "HCDP-DELM: Heterogeneous Chronic Disease
<1 %

Prediction with Temporal Perspective Enabled Deep Extreme Learning Machine", Knowledge-Based Systems, 2023

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Sai Ram Yadla

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7