https://doi.org/ 10.33472/AFJBS.6.7.2024.222-231

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

A NOVEL PERFORMANCE ANALYSIS APPROACH TO PREDICT CHRONIC KIDNEY DISEASE USING ANN

[1]V. Srikanth, [2]D.Hima Bindu, [3]B.Chakradhar, [4]S. Neeraja

[1]Assistant Professor, Department of Computer Science and Engineering, Raghu Engineering College (A), Visakhapatnam, Andhra Pradesh, India

[2]Assistant Professor, Department of Computer Science and Engineering (A), Raghu Engineering College (A), Visakhapatnam, Andhra Pradesh, India

[3]Assistant Professor, Department of Computer Science and Engineering, Raghu Engineering College (A), Visakhapatnam, Andhra Pradesh, India

[4]Assistant Professor, Department of Computer Science and Engineering, Lendi Institute of Engineering and Technology (A), Vizianagaram, Andhra Pradesh, India

Email ID : [1]santhasri99@gmail.com, [2]abindumtech@gmail.com, [3]chakri.589@gmail.com, [4]neerajasreerama@gmail.com

Article History
Volume 6,Issue 7, 2024
Received: 22 Feb 2024
Accepted : 26 Mar 2024
doi: 10.33472/AF5BS.6.7.2024.222-231

ABSTRACT: A condition that can be life-threatening is Chronic kidney disease (CKD), it can be described as a gradual decrease in renal function over a period of months to years without any major symptoms. Six phases relate to the different levels of severity. Depending on the Glomerular filtration rate (GFR), which considers several factors including age, sex, race, and serum creatinine, it is divided into different phases. Chronic renal failure is a dangerous condition that can be deadly if not identified and treated in a timely manner. Accurate early detection techniques for CKD are becoming increasingly essential as the number of cases of CKD increases. In the field of medical science, algorithms for Machine learning (ML) have grown in importance as a disease prediction tool. Therefore, developing an accurate machine learning model that accurately predicts CKD is the main goal of this work. This paper describes a novel performance analysis approach for predicting CKD using an Artificial Neural Network (ANN). Accuracy, Precision, Sensitivity, F1-score, and Specificity are the metrics used to analyze the performance analysis of the presented approach.

KEYWORDS: Glomerular Filtration Rate (GFR), Chronic Kidney Disease (CKD), Artificial Neural Network (ANN) and Machine Learning.

I. INTRODUCTION

Among the most important organs in the body is the kidneys, which generate urine, filter out all waste products and water. The two kidneys that are present in the human body are essential organs that are required for its correct operation and are situated behind the peritoneal cavity. Regulating the proper balance of salt, water, and other ions and trace elements in the body, including calcium, phosphorus, magnesium, potassium, chlorine, and acids, is the kidneys' primary task. Secrete hormones including erythropoietin, vitamin D, and rennin are secreted by the kidneys [1].

Sometimes referred to as chronic kidney disease (CKD) or chronic renal failure is an illness that may be deadly caused by the kidneys' incapacity to carry out their normal functions [2]. The symptoms of CKD might appear gradually and are not unique to any one condition; In fact, some individuals may not have any symptoms. As a result, early disease detection becomes extremely challenging.

The condition of the kidneys gradually decreases is referred to as chronic kidney disease, also called chronic kidney failure. Filtering wastes and extra fluid from the circulation is the kidneys' primary function. If the kidney is not functioning normally and is unable to sufficiently filter blood, it is said to have CKD. Once chronic kidney disease reaches to its final stages, dangerous levels of fluid, electrolytes, and waste products can occur in the body [3].

A disease or other condition that affects kidney function can cause chronic kidney disease, kidney damage that gets worse over several months or years is the characteristic of this condition. The primary signs of advanced CKD include lethargy and low energy, difficulty paying attention, decreased appetite, difficulty sleeping, cramps in the muscles at night, dry skin with intense itching, swelling around the eyes and legs, and frequent midnight urination [4].

CKD increases of Cardiovascular disease (CVD), including metabolic syndrome, hypertension, dyslipidemia, and diabetes mellitus. Moreover, End stage renal disease (ESRD), which has no cure, is also a result of CKD [5]. Early identification of CKD can help to guide the patient care and lower the death rate by stopping the disease from progressing to its final point, it is frequently linked to several health conditions, including high blood pressure, loss of bone, and heart disease [6]. To save their patients lives, doctors must be able to identify this illness at an early stage. Researchers have used a range of techniques to identify this condition early on [7].

The diagnosis and assessment of the severity of CKD have traditionally involved the use of a number of biomarkers and computational techniques. The GFR, which is most frequently used to evaluate kidney function, is insufficient to identify CKD [8]. Urine albumin testing and creatinine measurement, often known as creatinine clearance tests are required for routine CKD disease screening in order to determine GFR, but they are not costly or practical enough. Some patients with diabetes and cardiovascular illness may have a high GFR, which might mask the existence of true CKD [9].

Image saliency detection involves computer vision techniques like image segmentation that resemble video surveillance [10]. Nowadays, ML methods have been used in many applications related to health, numerous advantages have resulted from developments in sensor networks, data science, statistical processing, communication technologies, and data science, include the early diagnosis of a number of chronic illnesses and the implementation of pervasive, smart homes and assisted living facilities controlled by the Internet of Things (IoT) can identify deaths in older people [11].

Medical diagnosis of CKD can be improved by ML approaches [12]. Machine learning is becoming more important in the diagnosis of medical conditions because it makes complex analysis possible, which reduces human error and improves prediction accuracy. These days, the most dependable methods for predicting cancers, diabetes, liver disease, and heart disease are considered to be machine learning algorithms and classifiers [13].

In the medical domain, yet regression was done using logistic regression, random forest, and linear regression for prediction purposes [14], several machine learning algorithms used Naive Bayes, SVM, and decision trees with the goal of classifying data [15]. Due to early diagnosis and timely treatment, the death rate can be reduced with the effective application of these algorithms [16]. As a result, a unique performance analysis method for ANN-based chronic renal disease prediction is presented in this work.

Following is the arrangement of the remaining content: In Section II, the literature survey is explained. Section III presents a novel performance analysis approach for predicting CKD using ANN. Section IV discusses the result of the analysis of the approach that was described. Section V concludes with a conclusion.

II. LITERATURE SURVEY

S Reshma, Ajina, S R Salma Shaji, S R Vishnu Priya, Janisha et. al., [17] integrates the using of ML to predict CKD. This method offers the application of ML methods, such as Ant Colony Optimization (ACO) technique, for Chronic Kidney Disease. Using only the smallest possible features, the final output makes a prediction about whether or not that person has CKD. An optimization algorithm known as ACO is a meta-heuristic. 12 of the 24 available features are chosen as the best ones for prediction. Machine learning is the method used for prediction. Two classes, one with CKD and the other without are identified in the output of this classification challenge by the machine learning algorithm. This study's primary goal was to predict CKD patients with greater accuracy while utilizing smaller variables.

Qin J., Liu Y., Feng C., Liu C., Chen L., and Chen B. et. al., [18] explains an ML diagnostic approach for CKD. With a significant amount of missing values, the CKD data set was taken from the machine learning repository at the University of California, Irvine (UCI). The missing values were filled an using KNN imputation, which chooses a few full samples with the most comparable measurements to process each incomplete sample's missing data. Models were established using six ML algorithms: logistic regression, naive Bayes classifier, feed forward neural network, random forest, k-nearest neighbor, and support vector machine. The machine learning model that performed the best among them was random forest.

Zhao Jing, Gu Shaopeng, Adam Mc Dermaid et. al.,[19] explains using Random Forest Regression to predict CKD results from EMR (Electronic Medical Records) data. In this work, they created and verified an eGFR prediction model using data is taken from a local health system. Information from primary care clinics' lab, clinical, and demographic components are included in this dataset. Random Forest regression was used to develop the model, while discrimination metrics and goodness-of-fit statistics were used to evaluate its performance. Following preprocessing of the data, 61740 patients made up the patient cohort used for model building and validation. The final model, which had a mean coefficient of determination of 0.95, included Body mass index (BMI), gender, obesity, diabetes, eGFR, and age.

Babu Satukumati Suresh, Raghu Kogila, Shivaprasad Satla, et. al., [20] explains CKD feature extraction techniques. Investigating whether they can reasonably precisely predict CKD or non-CKD with fewer features is the aim of this work. The present investigation is using an intellectual framework advancement methodology. The informational index's reduced features are found using an

essential feature determination system, which is introduced. The reduced set of capabilities has been validated by two perceptive paired grouping techniques. As suggested by the results, they should concentrate more on the characteristics that are less common in the identification of CKD, which will decrease vulnerability, save time and money. The suggested method for CKD prediction makes use of less characteristics.

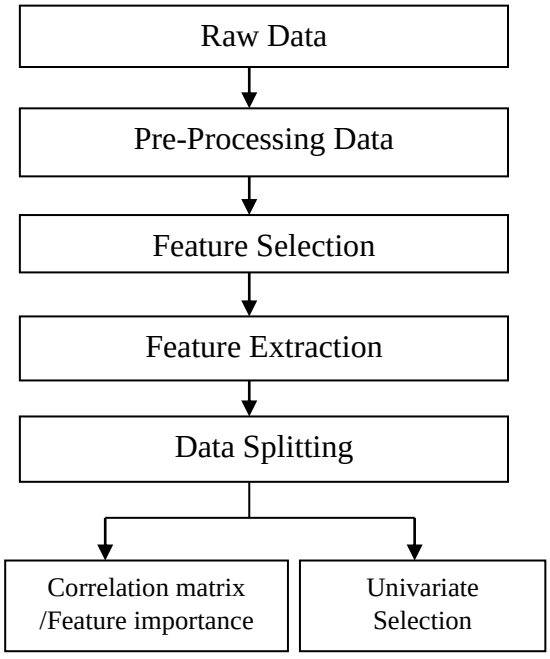
Tekale Siddheshwar, Shingavi Pranjal, Wandhekar Sukanya, Chatorikar Ankit et. al., [21] gives a presentation about using a ML algorithm to predict CKD. The authors examined 14 different features related to patients with CKD and estimated the accuracy of several ML methods, such as the decision tree and support vector machine. 91.75% accuracy is obtained by the decision tree algorithms, according to this results analysis. When examining the decision tree algorithm, all of the dataset's features are used to form the tree, which is based on the full dataset. The primary advantage of this system is that it requires less time to make predictions. It can help doctors in diagnosing more patients in less amount of time and in initiating treatment for CKD patients earlier. This study has certain limitations, including missing attribute values and a smaller data set that prevents the data strength from being greater.

Wang Zixian, Won Chung Jae, Muning Wang, Xilin Jiang, Yantong Cui, Anqi Zheng et. al.,[22] The associative classification technique-based machine learning prediction system for CKD is described. Developing the use of the Apriori association method, with 10-fold cross-validation testing, 400 cases of chronic kidney failure patients are recognized as having CKD. The results are compared against several classification algorithms, such as k-nearest-neighbor, J48, ZeroR, OneR, and naive Bayes. By filling and normalizing missing data, the dataset is preprocessed. For better accuracy and reduced training times, the dataset's most important features are chosen. According to Apriori, the results show that 99% of the characteristics in the dataset can be used to detect CKD. Four patient data samples are used to test the identified technique is further in predicting the patients CKD.

N Uma Dulhare; Ayesha Mohammad et. al., [23] discusses using the Naïve Bayes classifier to extract action rules for CKD. This Naive Bayes with OneR has been proposed by authors. Using the OneR approach, the amount of attributes in the dataset is additionally decreased by 80%, and the accuracy is enhanced by 12.5% when compared to the previous system. To stop the CKD from moving on to the next stage, the action rules for each stage are extracted by the suggested system and the appropriate therapies can be implemented in accordance with those people.

III. A NOVEL PERFORMANCE ANALYSIS APPROACH TO PREDICT THE CHRONIC KIDNEY DISEASE

A novel performance analysis approach for predicting chronic kidney disease using ANN is describes in this section. Fig. 1 demonstrates the architecture of the presented method.



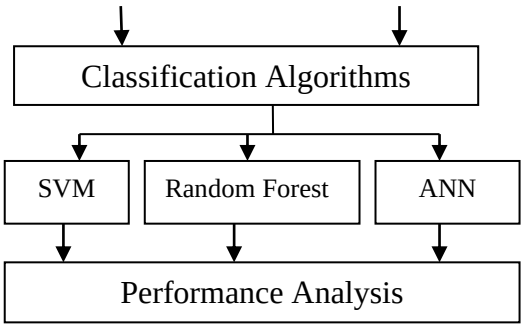


Fig. 1: The Architecture to predict the Chronic Kidney Disease using ANN

This ML repository uses the CKD dataset from UCI. It has 400 patients records. There are 25 attributes in the data. In this data collection, there are 400 instances; 62.5% of them have CKD, whereas 37.5% do'nt have CKD. There are twenty-five qualities in overall. The dataset needs to be classified into CKD and Not-CKD because there are values lacking that need to be provided. Through it has NA (Not Available) or null values for several attributes for a specific instance, it is necessary to clean the dataset that was obtained from the internet source. Data with missing values can be difficult to classify. Prior to data analysis, the missing values are used in the dataset must be eliminated as they lower the dataset's efficiency. There are two ways to identify the missing values: cases (records) or attributes. The degree of missing values from a cases (record) point of view can be simple, medium, or complex. If there is only one missing value in any one attribute for the case (record), then the degree is simple. From an attributes perspective, the absent values could be randomly chosen, monotone, or univariate. If every attribute (feature) has a missing value, then it is univariate. For every three attributes that have missing values, the attribute is considered monotone; for every other attribute, the missing values are considered arbitrary.

The Pandas libraries "dropna" and "fillna" are used to eliminate missing values from the dataset, such as NAs or blank values. They drop the column or row containing the missing data and replace NAs with the attribute's mean values, respectively. Every row or record with a null value is eliminated. The binary numbers 0 or 1 are used to replace all of the dataset's categorical values. The dataset's visualization demonstrates that individuals with high blood pressure and the elderly are more likely to get CKD. The process of choosing the most significant predictive features to use as model input is known as feature selection. Preprocessing is an essential stage in solving the high dimensionality issue.

Therefore, choosing the subset of characteristics that are relevant and unrelated to one another is the primary goal of feature selection while training a model. In the same way, feature selection is essential for creating a prediction model of chronic renal disease. As a result, the data's dimensionality and complexity are decreased, and the model becomes quicker, more accurate, and more efficient. Thus, following the creation of the dataset, relevant characteristics have been chosen using a feature selection technique. By identifying the most specific characteristics in signals, feature extraction makes them easier for a machine learning algorithm to process and understand. Due to information redundancy and high data rate, directly training machine learning or deep learning with raw signals frequently produces poor results. In this method, the features are extracted using PCA. Consequently, the major purpose of feature selection during model training is to gather the subset of attributes that are relevant and unrelated to one another. Similarly, feature selection is necessary to build a chronic renal disease prediction model. Consequently, this reduces the dimensionality and complexity of the data makes the model faster, more accurate, and

more efficient. Thus, a feature selection technique has been used to identify relevant features are once the dataset was created. The most differentiating features of signals are located by feature extraction, which improves processing of the signals by a machine learning algorithm. When machine learning or deep learning is trained directly using raw signals, it often performs poorly due to the high data velocity and information redundancy. The following are the parameters of the confusion metric:

True Positive (TP): The condition known as TP occurs when the expected value and the actual value are both positive.

True Negative (TN): When a data point's presented value and actual value are both negative, this is known as TN.

False Positive (FP): When a data point's actual value is positive but its actual value is negative, these are the cases.

False Negative (FN): represent the cases where a data point's actual value is negative but its actual value is positive.

Classification algorithms use the results of the confusion matrix and univariate selection. Mapping dataset data into pre-established groups or classes is referred to as classification (identifying the unclassified data's class). This classification is a categorization-related procedure that divides a set of data into various groups. Here, two classes are used in binary categorization. So, a person's test results for CKD would be either positive or negative. Three different algorithms Random forest, support vector machine, and Artificial Neural Network are used in this method and are explained as follows:

Random Forest (RF): Among of all the machine learning algorithms, RF is the most efficient in the ensemble and stacking classification technique. The RF technique has been applied to prediction and probability calculations. Numerous decision trees made up the Random Forest (RF) classifier. In 1995, Bell Labs Tin Kam Ho presented the idea of random forest, in which each decision tree votes are identify the class of an object. The combination of bagging and random attribute selection is known as the RF approach. Three hyperparameter tuning options are available for the random forest classifier. The three hyperparameter tuning settings are present in the random forest classifier: (i) the number of decision trees (n trees); (ii) the size of the smallest node in the trees; and (iii) the number of attributes utilized to break each node into its various parts trees (mtry). The number of attributes in this case is m.

Support Vector Machine: SVM, which may be used to address both linear and non-linear problems, this is a linear model for regression and classification. The technique uses a hyperplane to classify the data. This application will represent each data item as a point in n-dimensional space, where n is the number of features, and each feature's value corresponds to a certain position. The correct hyper-plane that can effectively separate the two classes will be found in order to carry out the classification process.

Within the field of machine learning models, artificial neural networks are a subset that are constructed utilizing the principles of neuronal architecture seen in the biological neural networks found in animal brains through connection. A computational network based on biological neural networks, which create the structure of the human brain, is called an artificial neural network. Neural networks in artificial intelligence are made up of neurons that are connected to one another at different layers, just as neurons in a human brain. Nodes are the name for these neurons. They refer to these neurons as nodes. An ANNs are primarily used to process and learn from data in order to identify patterns, predict results, and resolve challenging issues. It is made up of layers of connected nodes, or neurons. These nodes process data, changing the

weights of connections to identify trends and predict results. These categorization algorithms predict whether CKD would be present or Non-CKD. These three algorithms' performances are compared in terms of F1-score, Specificity, Accuracy, Precision, and Sensitivity.

IV. RESULT ANALYSIS

In this section, a novel performance analysis approach to predict the CKD using ANN is implemented. In this approach, CKD dataset is used. The CKD dataset attributes and type of attributes are tabulated as follows:

Table 1: Dataset Attribute Description

S. No	Attribute	Type
1	Age	Numerical
2	Blood Pressure	Numerical
3	Blood Glucose random	Numerical
4	Blood urea	Numerical
5	Serum creatinin	Numerical
6	Potassium	Numerical
7	Red Blood cell count	Numerical
8	Packed cell count	Numerical
9	White Blood cell count	Nominal
10	Haemoglobin	Nominal
11	Sodium	Nominal
12	Sugar	Nominal
13	Hypertension	Nominal
14	Anemia	Nominal
15	Class	Nominal
16	Pus Cell	Nominal
17	Pus Cell Clumps	Nominal
18	Red Blood Cells	Nominal
19	Diabetes Mellitus	Nominal
20	Albumin	Nominal
21	Cornary Artery Disease	Nominal
22	Bacteria	Nominal
23	Specific Gravity	Nominal
24	Appetite	Nominal
25	Pedal edema	Nominal

A wide range of performance evaluation metrics have been calculated, including as sensitivity, specificity, f1-score, accuracy, and precision.

The capacity of the classification algorithm to correctly predict the classes in the dataset is determined by accuracy. It is a measurement of the degree to which the real or theoretical value and the expected value are equal. In general, the proportion of correct forecasts to all instances is known as accuracy.

Precision: The true values that were accurately predicted from all of the expected values in the actual class are measured for precision. The capacity of the classifiers is not categorize a negative example as positive is measured by precision.

Sensitivity: The rate of positively classified values that are correctly identified is determined by this metric, which is sometimes referred to as recall or True positive rate (TPR). The percentage of real positives that are accurately classified is known as recall.

Specificity: The ability of the model or algorithm to forecast an actual negative for every accessible category is the defined. The ratio of accurately identified negative cases to the total number of negative instances is defined.

An accuracy metric for machine learning evaluation is the F1 score. It combines a model's precision and sensitivity scores.

The table 1 represents the evaluation of performance analysis.

Table 2: Performance Metrics Evaluation

Metrics/Algorithms	RF	SVM	ANN
Precision (%)	90.34	92.56	96.34
Sensitivity (%)	91.54	92.67	95.26
Specificity (%)	92.34	93.42	96.78
F1-score (%)	91.23	92.36	97.22
Accuracy (%)	92.53	94.21	97.43

From table 1, it is observed that, the results of ANN algorithms are better to those of RF and SVM methods. The Fig. 2 shows the comparison graph for Specificity.

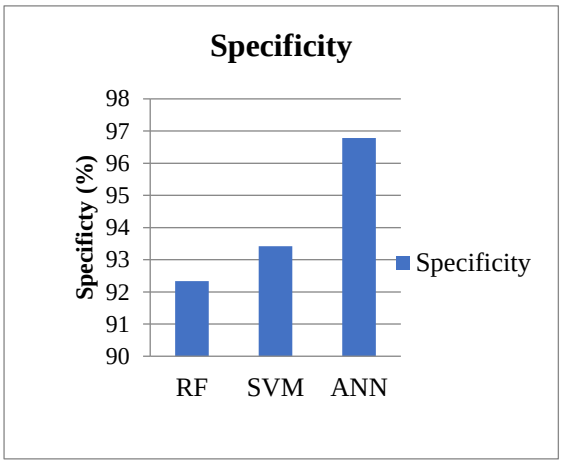


Fig. 2: Specificity Comparison

In Fig.2, the y-axis indicates the specificity values in terms of percentage and x-axis indicates different classification algorithms. ANN has better and high specificity than SVM and RF. The Fig. 3 shows the performance metrics comparison in terms of precision and sensitivity.

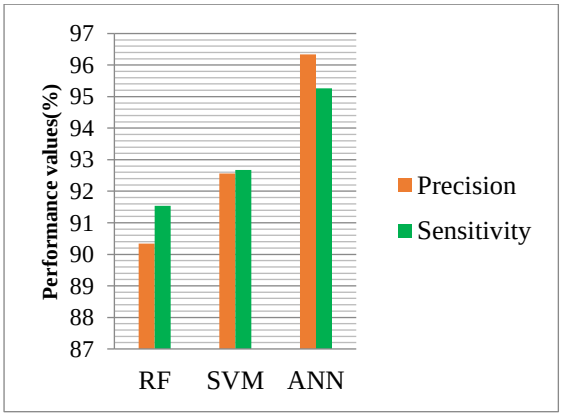


Fig. 3: Sensitivity and Specificity Comparison

The ML classification techniques are shown on the x-axis in Figure, while performance results are shown on the y-axis. ANNs are more sensitive and accurate than SVMs and RF classifiers. The accuracy and f1-score performance comparison is presented in Figure 4.

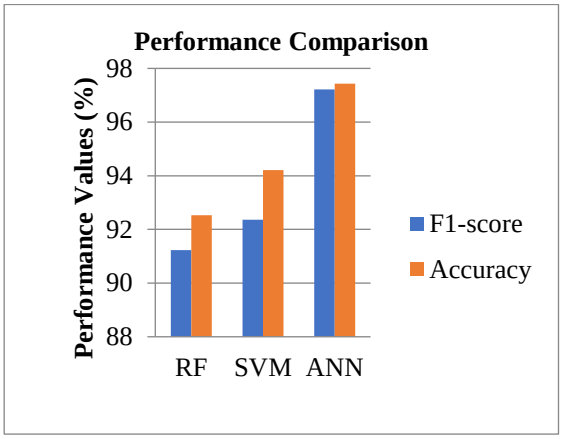


Fig. 4: Performance Comparison

Figure 4 shows performance values on the y-axis and ML classification algorithms on the x-axis. Compared to SVM and RF classifiers, the ANN has better accuracy and F1-score. Compared to RF algorithm, SVM has better performance. From the results, it is cleared that, ANN has better performance than SVM.

V. CONCLUSION

This section describes a novel performance analysis approach for predicting chronic kidney disease using ANN. Three different ML classifiers are used in this approach, to predict the development of CKD, such as SVM, RF, and ANN. A dataset including 400 patient records is used, the machine learning repository at UCI provides CKD. The collected data is preprocessed and necessary features are extracted. Univariate feature selection utilizes univariate statistical tests to each feature-outcome pair and selects features. The elements of the confusion matrix are utilized to predict whether CKD would be present or absent. Three classification algorithms are tested for performance using the following metrics: F1-score, Accuracy, Precision, Sensitivity, and Specificity. In terms of Precision, Sensitivity, F1-score, Accuracy, and

Specificity, ANN performs better than RF and SVM algorithms. This approach has effectively predicted the presence and absence of CKD. Hence, this approach will be useful for real time to predict the CKD and as a result, accurate diagnosis will be provided.

VI. REFERENCES

[1] Kumar, K.; Pradeepa, M.; Mahdal, M.; Verma, S.; RajaRao, M.V.L.N.; Ramesh, J.V.N, "A Deep Learning Approach for Kidney Disease Recognition and Prediction through Image Processing," Appl. Sci. 2023, vol. 13, no. 6, 2023, doi:10.3390/app13063621

[2] Shahinda Elkholy, Amira Rezk,Ahmed Abo El Fetoh Saleh, "Enhanced Optimized Classification Model of Chronic Kidney Disease", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 14, No. 2, 2023, doi: 10.14569/IJACSA.2023.0140239

[3] Dibaba Adeba Debal and Tilahun Melak Sitote, "Chronic kidney disease prediction using machine learning techniques", Journal of Big Data, vol. 9, no.109, 2022, doi: 10.1186/s40537-022-00657-5

[4] Mondol, Chaity, F. M. Javed Mehedi Shamrat, Md. Robiul Hasan, Saidul Alam, Pronab Ghosh, Zarrin Tasnim, Kawsar Ahmed, Francis M. Bui, and Sobhy M. Ibrahim, "Early Prediction of Chronic Kidney Disease: A Comprehensive Performance Analysis of Deep Learning Models" Algorithms, vol. 15, no. 9, 2022, doi:10.3390/a15090308

[5] Shahinda Mohamed Mostafa Elkholy, Amira Rezk, Ahmed Abo El Fetoh Saleh, "Early Prediction of Chronic Kidney Disease Using Deep Belief Network", IEEE Access, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3114306

[6] Ebrahime Mohammed Senan, Mosleh Hmoud Al-Adhaileh, Fawaz Waselallah Alsaade, Theyazn H. H. Aldhyani, Ahmed Abdullah Alqarni, Nizar Alsharif, M. Irfan Uddin, Ahmed H. Alahmadi, Mukti E Jadhav, and Mohammed Y. Alzahrani, "Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques", Journal of Healthcare Engineering, vol. 2021, pp. 1-10, 2021, doi:10.1155/2021/1004767

[7] Poonia RC, Gupta MK, Abunadi I, Albraikan AA, Al-Wesabi FN, Hamza MA, B T. Intelligent Diagnostic Prediction and Classification Models for Detection of Kidney Disease. Healthcare (Basel), vol. 10, no. 2, 2022 Feb, doi: 10.3390/healthcare10020371

[8] P. Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawat, Tulika Chakrabarti, Zbigniew Leonowicz, Michał Jasiński, Łukasz Jasiński, Radomir Gono, Elżbieta Jasińska, Vadim Bolshev, "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," in *IEEE Access*, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763

[9] Saurabh Pal, "Prediction for chronic kidney disease by categorical and non_categorical attributes using different machine learning algorithms", Multimedia Tools and Applications, 2023, doi:10.1007/s11042-023-15188-1,

[10] Elias Dritsas and Maria Trigka, "Machine Learning Techniques for Chronic Kidney Disease Risk Prediction", Big Data Cogn. Comput, MDPI Journal, 2022, 6, 98, doi:10.3390/bdcc6030098

[12] Md. Rashed-Al-Mahfuz, Abedul Haque, Akm Azad, Salem A. Alyami, Julian M. W. Quinn, And Mohammad Ali Moni, "Clinically Applicable Machine Learning Approaches to Identify Attributes of Chronic Kidney Disease (CKD) for Use in Low-Cost Diagnostic Screening", IEEE Journal of Translational Engineering in Health and Medicine, volume 9, 2021, doi: 10.1109/JTEHM.2021.3073629

[13] Swetha, A. ., Lakshmi, M. S. ., & Kumar, M. R, "Chronic Kidney Disease Diagnostic Approaches using Efficient Artificial Intelligence methods", International Journal of Intelligent Systems and Applications in

Engineering, vol. 10, no. 1, 2022, doi: https://ijisae.org/index.php/IJISAE/article/view/2289

[14] Pankaj Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawat, Tulika Chakrabarti, Zbigniew Leonowicz, Michał Jasiński, "Prediction of Chronic Kidney Disease - A Machine Learning Perspective", IEEE Access (Volume: 9), 2021, doi: 10.1109/ACCESS.2021.3053763

[15] Hamida Ilyas, Sajid Ali, Mahvish Ponum, Osman Hasan, Muhammad Tahir Mahmood, Mehwish Iftikhar and Mubasher Hussain Malik, "Chronic kidney disease diagnosis using decision tree algorithms", BMC Nephrology, vol. 22, no. 273, 2021, doi:10.1186/s12882-021-02474-z

[16] Rahul Gupta, Nidhi Koli, Niharika Mahor, N Tejashri, "Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease", 2020 International Conference for Emerging Technology (INCET), Belgaum, India. Jun 5-7, 2020

[17] Reshma S, Salma Shaji, S R Ajina, Vishnu Priya S R, Janisha, "Chronic Kidney Disease Prediction using Machine Learning", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 9, no. 07, July-2020, doi: 10.17577/IJERTV9IS070092

[18] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng and B. Chen, "A Machine Learning Methodology for Diagnosing Chronic Kidney Disease," in *IEEE Access*, vol. 8, pp. 20991-21002, 2020, doi: 10.1109/ACCESS.2019.2963053.

[19] Jing Zhao, Shaopeng Gu, Adam McDermaid, "Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression", Mathematical Bio-sciences (2019), doi:10.1016/j.mbs.2019.02.001

[20] Suresh Babu Satukumati, Shivaprasad Satla, Raghu Kogila, "Feature Extraction Techniques for Chronic Kidney Disease", Ingenierie des Systemesd' Information, Volume 24, issue No.1, 2019, pp:95-99, doi: 10.18280/isi.240114

[21] Siddheshwar Tekale, Pranjal Shingavi, Sukanya Wandhekar, Ankit Chatorikar, "Prediction of Chronic Kidney Disease Using Machine Learning Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 7, Issue 10, October 2018, ISSN (Online) 2278-1021

[22] Zixian Wang, Jae Won Chung, Xilin Jiang, Yantong Cui, Muning Wang, Anqi Zheng, "Machine Learning-Based Prediction System For Chronic Kidney Disease Using Associative Classification Technique", Vol.7, 2018, special issue 36, doi: 10.14419/ijet.v7i4.36.25377

[23] Uma N Dulhare; Mohammad Ayesha, "Extraction of action rules for chronic kidney disease using Naïve bayes classifier", 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), DOI: 10.1109/ICCIC.2016.7919649