

Risk Assessment of Chronic Kidney Disease

Dr. Rajeshwari B S¹, Swetha Patil², Vaghdevi Praveen T³, Yashwanth Kiran S⁴ and Sushant⁵

¹Assistant Professor Dept. of CSE B.M.S. College of Engineering Bangalore, India

²⁻⁵Dept. of CSE B.M.S. College of Engineering Bangalore, India

E mail: rajeshwari.cse@bmsce.ac.in, swethapatil.cs19@bmsce.ac.in, vaghdevit.cs19@bmsce.ac.in, yashwanths.cs19@bmsce.ac.in, sushant.cs19@bmsce.ac.in

Abstract -The pervasiveness of chronic kidney disease (CKD), an international health concern is rising primarily as a result of rising rates of hypertension and diabetes. Furthermore, people who are suffering from CKD may develop end-stage renal failure if their condition is not managed. Finding those who have undiagnosed CKD or who have a higher chance of getting CKD or progressing to End-Stage Renal Kidney Disease (ESKD) is thus a significant concern. For generating risk-predicting models that can precisely predict CKD in adults or end-stage renal failure, some appropriate methods have to be accompanied and if inappropriate methods are used, it can affect the predictive capability of the models. In this paper, an attempt is made to high-point the significance of disease related parameters in the neuropsychological assessment of individuals with chronic kidney disease (CKD) and discuss the summary of the research done in this area.

I. INTRODUCTION

Untreated chronic kidney disease (CKD) can have serious consequences, including end-stage kidney failure, which is a significant global health issue. However, even though end-stage kidney disease is not as common, the treatments available such as dialysis or transplantation which is expensive [1]. To aid in the clinical management of patients, the use of risk prediction models can be highly beneficial. The techniques are typically multivariable and combine various patient risk parameters to analyze a person's future health status.

Chronic kidney disease is a condition characterized by the gradual decrease of functioning of the kidney as time passes, resulting in the accumulation of potentially harmful levels of wastes, electrolytes, and fluids in the body [2]. This has been a prevalent condition affecting many people worldwide. CKD can be caused by various risk parameters like dietary habits, environmental factors, and living conditions. Often, people may not be much aware of their illness until it progresses to an advanced stage. Machine learning (ML) is emerging as a promising approach for predicting CKD, offering more precise analysis of data than traditional methods.

In the initial condition of CKD, symptoms or warning signs may be minimal [3]. The primary objective of treatment is to address the underlying cause and decelerate the advancement of kidney damage. However, simply addressing the root cause might not be sufficient to prevent further damage. Therefore, assessing the risk of renal disease is crucial to prevent additional harm.

II. RELATED WORK

Bilal Khan et al., [4] discusses many techniques and models to identify the CKD in early-stage. Machine Learning (ML) methods have been crucial in initial identification of various illnesses, and the dataset collected from UCI repository was utilized in this research. Seven machine learning (ML) techniques were utilized, including Naive Bayes Tree (NBTree), J48, Support Vector Machine, Multi-layer Perceptron (MLP), Naive Bayes (NB), Logistic

Regression (LR) and Composite Hypercube on Iterated Random Projection (CHIRP). Performance of these mentioned methods were evaluated utilizing different measures, such as Mean Absolute Error (MAE), Relative Absolute Error (RAE), Root Mean Squared Error (RMSE), Root Relative Squared Error (RRSE), Recall, Precision, Accuracy and F-Measure. The authors compared the overall accuracy metrics, recall, precision and F-measure, and found that the CHIRP had a performance of 0.998 for all three metrics. This suggests that CHIRP is the one that gave better results for predicting CKD, while the other six ML techniques are not as adequate as CHIRP for predicting CKD.

Mirza Muntasir Nishat et al., [5] applied different machine learning algorithms with the aim of assessing and comparing its accuracies and other performance parameters for the identification of the CKD. This study utilized the dataset which is present in UCI repository consisting of 400 rows and 25 attributes. Due to missing values in columns, the researchers created four data frames, which involved inserting missing values with the mean, median, mode values, or removing them completely. The final dataset consisted 158 instances that were employed to compare performance metrics of ML algorithms without and with the hyperparameter tuning. Random Forest (RF) demonstrated with the accuracy 99.75% which is highest after hyperparameter tuning, followed by Logistic Regression with the accuracy 99.36%, Decision Tree with the accuracy 99%, Support Vector Machine with the accuracy 99.36%, and Naïve Bayes with the accuracy 99%. The study investigated the utilization of supervised machine learning algorithms in bioinformatics and illustrates its suitability in early identification of various fatal diseases including CKD.

R.Parthiban et al.,[6] proposed filter embedded wrapper feature selection for the diagnosis of CKD. The researchers utilized a Support Vector Machine (SVM) based machine learning algorithm to detect CKD and employed three different feature selections which are wrapper, filter, and embedded methods to decrease the number of parameters in the CKD dataset. The filter approach used Relief, One-R, Gain Ratio (GR), and Gini Index (GI), while the wrapper approach utilized the Improved Bat Algorithm. The embedded technique, Support Vector Machine-t-statistics (SVM-t), was utilized to choose analytical features from the CKD dataset. The suggested SVM by hybrid filter, wrapper, and Embedded feature selection (SVM-HFWE-FS) algorithm showed higher accuracy of 93.33%, compared to Artificial Neural Network using hybrid filter, wrapper, and Embedded feature selection (ANN-HFWE-FS) which had an accuracy of 86.67%. The dataset had 400 patients, 250 of whom had early-stage CKD, and had some missing and noisy values. Authors recommended that future research can consider a classification algorithm capable of handling missing and noisy values.

Fuzhe Ma et al., [7] proposed Heterogeneous Modified Artificial Neural Network (HMANN) for the early detection, segmentation, and identification of renal failure on Internet of Medical Things (IoMT) platform. The suggested method HMANN utilized the combination including Support Vector Machine and also Multilayer Perceptron algorithms with Backpropagation to classify ultrasound images of the kidney. The region of interest in the kidney was segmented and utilized as the base for the algorithm's operation after a pre-processing step. The pre-processing of the kidney image involved three stages: Image Restoration, Sharpening and Smoothing, and Contrast Enhancement. SVM classifiers were used with different kernels to identify abnormal images, achieving a multilayer perceptron accuracy of 97.50%. The availability of this data was useful in recommending immediate disease prevention and monitoring of progression. The proposed HMANN method effectively reduces noise and aided in segmenting the kidney image, making it easier to locate kidney stones.

Charumathi Sabanayagam et al., [8] developed an Artificial Intelligence based deep learning algorithm that utilizes retinal images to identify CKD. The algorithm has the capacity to improve the current screening methods for this disease. The Singapore Epidemiology of Eye Diseases study (SEED) provided data and images from participants who attended the baseline visits. Chronic renal illness was identified based on glomerular filtration rate. Three models were developed and trained, namely Image Deep Learning Algorithm (DLA), Risk Factors (RF) model, and Hybrid Deep Learning Algorithm (DLA) that combined both image and risk factors. The models' performance was assessed by calculating the area under the receiver operating characteristic curve (AUC). The results indicate that the hybrid deep learning algorithm had the highest AUC of 0.938, followed by the risk factor model of 0.916, and Image Deep Learning Algorithm with 0.911. Deep learning algorithms for the retinal images demonstrated excellent performance in identifying chronic renal illness.

Weilun Wang, et al., [9] proposed a two-stage method to calculate the risk of CKD under the condition that creatinine dataset is not available. The value of creatinine is predicted using supervised data in the first stage using a machine learning regression model. The data was extremely unbalanced to prognosticate the target variable: creatinine. Thus, authors used an under-sampling technique and presented a cost-sensitive mean squared error (MSE) loss function to address the issue. Regarding model choice, proposed method considered three machine learning models in the work: The Random Forest model (RF), the XGBoost boosting tree model (XGB), and the ResNet Neural Network model. As part of their ensemble learning, the proposed method

averaged the data from eight predictors to better predict creatinine levels. Finally, a binary prediction of the risk of CKD was made using the predicted creatinine and the original 23 features. Authors utilized the open source information given by the National Health Insurance Sharing Service for the work. The best result, an R2 score of 0.5523, was attained using XGBoost. Therefore, in the subsequent tests, XGBoost is chosen as a standard predictor. When categorizing samples for CKD, the AUC of 0.76 is obtained with the anticipated value of creatinine.

Sarah A et al., [10] proposed a method that combines information-gain-based attribute selection with a cost-sensitive AdaBoost classifier to increase the pinpointing of chronic kidney illness. The CKD dataset created in 2015 by Apollo Hospitals in TamilNadu, India, is used in the work. For performance comparison, six additional machine learning classifiers were put into use as a baseline: Logistic Regression, Decision Trees, Random Forests, SVM, XGBoost, and the conventional AdaBoost. First, the relevance of the various traits was calculated and ranked using the Integrated Gradient (IG) approach. Second, both the complete and reduced feature sets were used to train the classifiers. The results demonstrate that some attributes improved the classifiers' performance. Furthermore, in comparison to other classifiers and algorithms, the suggested cost-sensitive AdaBoost performed better.

Authors concluded, combining the IG-based feature selection strategy with the cost-sensitive AdaBoost is a successful method for CKD detection and may be used for computer aided diagnosis of CKD in its early stages.

Pankaj Chittora, et al., [11] incorporated seven different machine learning classifiers on the dataset to determine whether or not someone has CKD. Both the entire features and the chosen features were active for each algorithm. Every outcome was recorded and oversampling was done using Synthetic Minority Oversampling Technique (SMOTE). For training the model, they used machine learning classifiers such as the Artificial Neural Network (ANN), Logistic Regression, Linear Support Vector Machine (LSVM), K-Nearest Neighbors (KNN), and Random Forest. First, the performance of different Machine Learning algorithm has been evaluated on all features. Second part, attribute selection algorithm such as Correlation based feature selection (CFS), forward wrapper and Least Absolute Shrinkage and Selection Operator (LASSO) have been applied on the dataset to select the important features. Third, the performance of all above-mentioned classification algorithm on important features was calculated. Fourth, Synthetic Minority Oversampling Technique (SMOTE) filter was also applied on the dataset and the result of classifiers were checked. Various tools were used. Weka tool was used for LASSO. IBM SPSS modeler was used for performance of classifiers. Deep Neural Network was built in IBM SPSS modeler. Artificial Neural Network (ANN) was used with two Hidden layers for building Deep Neural Network. The result demonstrates that among all models, the Deep Neural Network had the highest accuracy of 99.6%. In SMOTE with all features, LSVM with penalty L2 displayed the best performance and recorded the maximum accuracy of 98.46%.

Reshma S et al., [12] proposed machine learning approaches like ACO and SVM by decreasing the features and choosing the best features in order to increase prediction accuracy. The CKD dataset taken from the UCI repository. 400 samples from two separate classes are included. Here, the patient's information—including age, blood pressure, specific gravity, albumin, sugar, red blood cells, etc. is used to populate the dataset. ACO is a meta-heuristic optimization algorithm. The 12 top attributes are selected for prediction out of the 24 available attributes. The SVM machine learning approach is used for prediction. SVM divides the output of this classification task into two classes: those with and those without CKD. The major goal of this experiment was to accurately predict individuals with CKD while using fewer selected numerical attributes which helped in achieving an accuracy of 96%.

Alvaro Sobrinho et al., [13] discusses three main contributions: First part, the outcome of systematic literature review focused on CKD diagnosis assisted by machine learning. Second part, the consequence of an experiment based on machine learning techniques using a CKD dataset collected from developing country hospitals and analyzed by nephrologists. Third part, a discussion on adequacy of machine learning techniques in developing countries. The k-fold cross validation method with $k=1$ was used to calculate the classifiers based on 114 records including 60 real world and 54 augmentation data. Samples were augmented by duplicating real world medical records and carefully modifying the attributes; increasing the CKD biomarker by 0.5. The training set was conducted based on the CKD risk evaluations provided by most experienced nephrologists. The training and testing steps included k-fold cross validation method by applying the J48 decision tree, Random Forest, Naive Bayes, Support Vector Machine, Multi-Layer Perceptron and K-Nearest Neighbour algorithms with Weka software. The results were compared with the opinions of three nephrologists who evaluated the medical records and provided their conclusions concerning the health situations of the subjects considering their risk of developing CKD. Their results were collated to the machine learning classifiers using Cohen's Kappa statistic.

The results demonstrated that J48 decision tree and Random Forest are the most adequate machine learning techniques in the context of developing countries showing 95.00% and 93.33% accuracy.

Jiongming Qin et al., [14] proposed a machine learning model for detection of CKD. The CKD values were acquired from the University of California Irvine (UCI) repository, which has some null values. KNN algorithm was used to fill in the null data, which selects median data values accompanying calculations to fill the null values. By efficiently inserting the incomplete dataset, six machine intelligence models: Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbor, Naive Bayes classifier and Feed Forward Neural Network were applied to authorize models. With the models, Random Forest got high-quality efficiency with 99.75% disease veracity. By analyzing the results obtained from the machine learning algorithms, the authors made an algorithm that had both Logistic Regression and Random Forest by utilizing a perceptron that could get to an average veracity of 99.83% after 10 occasions of imitation. Hence, they speculated that the given methods may be used for more complex medical data for CKD.

Revathy et al., [15] presented a prediction algorithm to identify CKD at a prior stage. The models have training and testing models for the provided input parameters, and the dataset displays input parameters gathered from CKD patients. For the CKD diagnosis, Decision Tree, Random Forest, and Support Vector Machine training models are built. The prediction of the models is assessed depending on how accurately they make predictions. The study's findings demonstrated that, when compared with Decision Trees and Support Vector Machines, the Random Forest Classifier model is more accurate at predicting CKD. The accuracy of the Random Forest model is 99.16%.

The potential link between CKD and the severity of COVID-19 infection was investigated by Henry et al. [16]. The authors reported how a meta-analysis using retrievable data included the estimation of the odds ratio (OR) and its 95% confidence interval (95% CI) in people who had severe or non-severe COVID-19. Both local legislation and the Helsinki Declaration happened during the study's conduct. Based on electronic and reference searches, 84 articles were identified, of which, after screening by title, abstract, and full text, 80 were disregarded as not specifically relevant to COVID-19 (n = 25), review articles (n = 7), articles that lacked pertinent data (n = 28), articles that were edited (n = 10), articles that lacked information on severity and/or comorbidities (n = 8), and patients by mortality but not by severity (n = 1). A total of four trials that comprised 1389 COVID-19 patients of whom 273 (19.7%) were deemed to have severe disease could be incorporated into the meta-analysis. According to Ekanayake et al.'s workflow [17] for predicting CKD status based on clinical data, a collaborative filtering, and attribute selection are all included for addressing missing values. The additional tree classifier and random forest classifier are demonstrated to produce the highest accuracy and least amount of bias to the attributes out of the 11 machine learning techniques taken into consideration. The study shows the significance of using domain expertise when utilizing machine learning for CKD status prediction as well as the practical elements of data collection. Therefore, it would be worthwhile to investigate the adoption of a KNN computer-based strategy in the future to handle missing values in data sets connected to various diseases. Furthermore, the authors discussed how more insights into CKD can be gained by adding knowledge of genomics, water consumption patterns, and food types into the analysis.

R. Devika et al. [18] compared the efficiency of Naive Bayes, K-Nearest Neighbour (KNN), and Random Forest classifiers to predict the presence of CKD in terms of accuracy, precision, and execution time. Authors also developed a classifier model for forecasting the transitioning period of renal disease stages 3 to 5. The final finding of the model is that Random Forest performance surpasses Naive Bayes and KNN in terms of classifier performance. The Random Forest Classifier model is more accurate in predicting CKD with an accuracy of 99.32%.

Ghosh et al., [19] done the evaluation process of the different models using the 80:20 distinction, and this study also clarifies accuracy, error rate, execution time, and AUC. Four machine learning algorithms viz. Support Vector Machine, AdaBoost, Linear Discriminant Analysis and Gradient Boosting were put on training using training data. Once the algorithms were trained, they were implemented on test data to obtain prediction the accuracy and achievement of the prediction of all four algorithms were analyzed to figure out the most efficient algorithm to predict the CKD among the patients. Among all the four classification algorithms though Gradient Boosting classifier took more time compared to all others to give a prediction and highest predictable score in both Receiver Operating Characteristic (ROC) and AUC curves gave the accuracy rate of 99.80% while AdaBoost and Linear Discriminant Analysis gave 97.91% accuracy.

Bhavya Gudeti et al., [20] compared the effectiveness of several machine learning algorithms based on their accuracy. This study's primary goal was to classify cases of CKD and non-CKD leveraging the dataset for CKD. Support Vector Machine, Logistic Regression, K-Nearest Neighbor and Gradient Boosting were analyzed and then proceeded to perform the analysis of CKD. The outcomes illuminated that the Support Vector Machine

algorithm successfully predicted CKD than Logistic Regression and K-Nearest Neighbors within stipulated narrow limits as per the given medical scenario. The advantage of this method is that the process of prediction required significantly less time, enabling doctors to start treating CKD patients as soon as possible and to classify an extensive group of patients in a shorter period of time. The authors also discussed that the model can be experimented with very huge datasets in the near future to compare the outcomes of this dataset with any other dataset that includes exactly the same information because the dataset utilized in the work has only 400 cases. Additionally, using the proper information, authors attempted to determine whether any human with these criteria has a very high likelihood of developing chronic risk factors like hypertension, a hereditary history of renal failure and diabetes for the sake of reducing the future incidence of CKD. The simulation results obtained were, Logistic Regression with 77.25% accuracy, Support Vector Machine with 99.25% accuracy and K-Nearest Neighbor with 78.75 % accuracy.

M. U. Emon et al., [21] considered some attributes to measure the analysis of CKD, and these attributes were major occurrences of CKD. The authors have used eight different machine learning classifiers to measure the analysis using Weka tools, viz., Logistic Regression (LG), Naive Bayes (NB), Multilayer Perceptron (MLP), Stochastic Gradient Descent (SGD), Adaptive Boosting (Adaboost), Bagging, Decision Tree (DT), and Random Forest (RF) classifiers. They also conducted feature extraction of the attributes, leveraging Principal Component Analysis (PCA). In the principal component analysis, they were colored as red and green. The red color indicates for CKD and the green color indicates for Non CKD. Authors conducted cross validation which is one of the method for evaluating machine learning models that involves training different machine learning algorithms on the subsets of input data that are readily available and then comparing their performance on the correspondent data subset. They made use of k-fold cross validation. In K-Fold cross validation, the data are divided into k subsets. The held out solution was treated k times, with each iteration using the test the set/validation as one of the k subsets and combining the leftover k-1 subsets to create a training set. The accuracy of several types of algorithms was calculated using the degree of precision in the data set, and the best result was obtained by utilizing a Random Forest classification technique. The work gained the accuracy of 99% for the Random Forest which is the highest and ROC curve value obtained was also highest from other algorithm, the second-highest accuracy attained was 98% with Multi Level Perceptron, Stochastic Gradient Descent, and Decision Tree classifiers. The Naive Bayes classifier gave an accuracy as low as 95%. The classifier's second-lowest accuracy was 96% from Logistic Regression.

Md. Ashiqul Islam et al., [22] applied six algorithms: Naïve Bayes, Random Forest, Simple Logistic Regression, Decision Stump, Linear Regression model, Simple Linear Regression for forecasting the risk factors associated with CKD. Six algorithms provided a better and speedier characterization execution when taking into account the ordered execution and analysis of these strategies. In this investigation the authors have obtained outcomes from these three specific computations Naive Bayes, Random Forest, Simple Logistic which were regulated calculations. They examined the result and got a follow up of around 90%, and authors described these models as highly effective for the dataset. Additionally, the authors identified the possible causes for kidney problems by inspecting calculations that were made by Decision Stump, Linear Regression Model, and Simple Linear Regression. From the outcome analysis, authors ascertained that the Random Forest algorithm got the highest accuracy 98.88% and was superior to any other model.

E. Dritsas et al., [23] aimed to build efficient tools for predicting CKD occurrence, following an approach which exploits ML techniques. First, authors applied class balancing in order to tackle the non-uniform distribution of the instances in the two classes, then features ranking and analysis are performed, and finally, methodology based on supervised learning is described, which aims to create efficient models for determining the risk of CKD occurrence by mainly focusing on probabilistic, tree-based and ensemble learning-based models. Moreover, authors evaluated Support Vector Machine, Logistic Regression, Stochastic Gradient Descent, Artificial Neural Network and K-Nearest Neighbour. The obtained results highlighted the Random Forest model achieved greater performance compared to other different models with an AUC of 100%, Precision, Recall, F-Measure and Accuracy equal to 99.2%.

III. CONCLUSION

The effect of CKD has been steadily increasing, posing a significant public health challenge worldwide. The paper has highlighted the importance of early detection and intervention in managing CKD, as it has been shown to improve patient outcomes and reduce the burden on healthcare systems. The paper has also highlighted on the risk parameters associated with CKD, including hypertension, obesity, diabetes and smoking. Identifying and addressing these risk parameters through targeted interventions and lifestyle modifications can play a crucial role

in preventing the onset and progression of CKD. Importantly, the study has underscored the need for further research in several areas. These include exploring novel therapeutic targets, developing more accurate predictive models, and investigating the impact of lifestyle interventions and dietary modifications on CKD progression. By leveraging the findings presented helps to improve the prevention, diagnosis, and management of CKD, ultimately enhancing patient outcomes and reducing the global burden of this debilitating disease. In conclusion, research from many authors suggested models that gives good accuracy but need to be done for more clinical data and also should optimize the parameters.

REFERENCES

- [1] O'Connell, Philip J., Mark Brown, Tak Mao Chan, Rolando Claure-Del Granado, Simon J. Davies, Somchai Eiam-Ong, Mohamed H. Hassan. "The Role of Kidney Transplantation as a Component of Integrated Care for Chronic Kidney Disease" *Kidney International Supplements* 10, No. 1, 2020, PP. 78-85, DOI: 10.1016/j.kisu.2019.11.006.
- [2] Rasel, Nur-E., Md Alam, Mohammad Akther-uz Zahan, Md Hossain, and Shahi Emran, "Serum Creatinine can be Reduced by Applying Homeopathic Medicines according to the Symptom Similarity: Case study Analysis of Chronic Kidney Disease (CKD)", *Journal of Complementary and Alternative Medical Research* 12, No. 1, 2020, PP.1-10.
- [3] Bridoux, Frank, Nelson Leung, Mohamed Belmouaz, Virginie Royal, Pierre Ronco, Samih H. Nasr, Jean Paul Fermand, "Management of Acute Kidney Injury in Symptomatic Multiple Myeloma." *Kidney international* 99, International Kidney and Monoclonal Gammopathy Research Group, No. 3, 2021, PP. 570-580.
- [4] Khan, B., Naseem, R., Muhammad, F., Abbas, G. and Kim, S., "An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Prophecy". *IEEE Access*, 8, PP.55012-55022, 2020.
- [5] Nishat, M.M., Faisal, F., Dip, R.R., Nasrullah, S.M., Ahsan, R., Shikder, F., Asif, M.A.A.R. and Hoque, M.A., "A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms", *Endorsed Transactions on Pervasive Health and Technology*, Volume 7, No.29, .2021, DOI: <https://doi.org/10.4108/eai.13-8-2021.17067>.
- [6] Parthiban, R., Usharani, S., Saravanan, D., Jayakumar, D., Palani, D.U., StalinDavid, D.D. and Raghuraman, D., "Prognosis of Chronic Kidney Disease (CKD) using Hybrid Filter Wrapper Embedded Feature Selection Method", *European Journal of Molecular & Clinical Medicine*, Volume 7, No.9, PP.2511-2530, 2021.
- [7] Ma, F., Sun, T., Liu, L. and Jing, H., "Detection and Diagnosis of Chronic Kidney Disease using Deep Learning-Based Heterogeneous Modified Artificial Neural Network", *Future Generation Computer Systems*, 111, PP.17-26, 2020.
- [8] Sabanayagam, C., Xu, D., Ting, D.S., Nusinovi, S., Banu, R., Hamzah, H., Lim, C., Tham, Y.C., Cheung, C.Y., Tai, E.S. and Wang, Y.X., "A Deep Learning Algorithm to Detect Chronic Kidney Disease from Retinal Photographs in Community-Based Populations", *The Lancet Digital Health*, Volume 2, No. 6, PP.295-302, 2020.
- [9] Wang, W., Chakraborty, G. and Chakraborty, B., "Predicting the Risk of Chronic Kidney Disease using Machine Learning Algorithm" *Applied Sciences*, Volume 11, No.1, 2020.
- [10] Ebiaredoh-Mienye, S.A., Swart, T.G., Esenogho, E. and Mienye, I.D., "A Machine Learning Method with Filter-Based Feature Selection for Improved Prediction of Chronic Kidney Disease", *Bioengineering*, Volume 9, No.8, 2022.
- [11] Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasiński, M., Jasiński, Ł., Gono, R., Jasińska, E. and Bolshev, V., "Prediction of Chronic Kidney Disease-A Machine Learning Perspective", *IEEE Access*, Volume 9, PP.17312-17334, 2021.
- [12] Reshma, S. and Ajina, V., Janisha, "Chronic Kidney Disease Prediction using Machine Learning", *International Journal of Engineering Research & Technology*, Volume 09, Issue 07, July 2020.
- [13] Sobrinho, A., Queiroz, A.C.D.S., Da Silva, L.D., Costa, E.D.B., Pinheiro, M.E. and Perkusich, A., "Computer-Aided Diagnosis of Chronic Kidney Disease in Developing Countries: A Comparative Analysis of Machine Learning Techniques", *IEEE Access*, Volume 8, PP.25407-25419, 2020.
- [14] Qin, J., Chen, L., Liu, Y., Liu, C., Feng, C. and Chen, B., "A Machine Learning Methodology for Diagnosing Chronic Kidney Disease", *IEEE Access*, Volume 8, PP. 20991-21002, 2019.
- [15] Revathy, S., Bharathi, B., Jeyanthi, P. and Ramesh, M., "Chronic Kidney Disease Prediction using Machine Learning Models", *International Journal of Engineering and Advanced Technology*, Volume 9, NO 1, PP.6364-6367, 2019.
- [16] Henry, B.M. and Lippi, G., "Chronic Kidney Disease is Associated with Severe Coronavirus Disease 2019 (COVID-19) Infection", *International urology and nephrology*, Volume 52, No.6, Pp.1193-1194, 2020.
- [17] Ekanayake, I.U., & Herath, D., "Chronic Kidney Disease Prediction Using Machine Learning Methods", 2020 Moratuwa Engineering Research Conference, PP. 260-265, 2020.
- [18] R. Devika et al. 2019. Chronic Kidney Disease Prediction using Machine Learning Models. *International Conference on Inventive Computation Technologies (ICICT)*, PP.634-678, 2019,
- [19] Ghosh, P., Shamrat, F.J.M., Shultana, S., Afrin, S., Anjum, A.A. and Khan, A.A., "Optimization of Prediction Method of Chronic Kidney Disease using Machine Learning Algorithm", *15th IEEE International Joint Symposium on Artificial Intelligence and Natural Language Processing*, PP. 1-6, 2020.
- [20] B. Gudeti, S. Mishra, S. Malik, T. F. Fernandez, A. K. Tyagi and S. Kumari, "A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms," *4th International Conference on Electronics, Communication and Aerospace Technology*, PP. 1630-1635, 2020, DOI: 10.1109/ ICECA49313.2020.9297392.

- [21] M. U. Emon, A. M. Imran, R. Islam, M. S. Keya, R. Zannat and Ohidujjaman, "Performance Analysis of Chronic Kidney Disease through Machine Learning Approaches", 6th International Conference on Inventive Computation Technologies, PP. 713-719, 2021, DOI: 10.1109/ICICT 50816.2021.9358491.
- [22] M. A. Islam, S. Akter, M. S. Hossen, S. A. Keya, S. A. Tisha and S. Hossain, "Risk Factor Prediction of Chronic Kidney Disease based on Machine Learning Algorithms", 3rd International Conference on Intelligent Sustainable Systems, PP 952-957, 2020, DOI: 10.1109/ICISS49785.2020.9315878.
- [23] E. Dritsas and M. Trigka, "Machine Learning Techniques for Chronic Kidney Disease Risk Prediction," Big Data and Cognitive Computing, Volume.6, No.3, 2022, doi: 10.3390/bdcc6030098.