Explainable Chronic Kidney Disease (CKD) Prediction using Deep Learning and Shapley Additive Explanations (SHAP)

Khadiime Jhumka*
khadhiime@gmail.com
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Muhammad Muzzammil Auzine
mmuzzammil.auzine@gmail.com
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Maleika Heenaye-Mamode Khan
m.mamodekhan@uom.ac.mu
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Mohammad Shoaib Casseem
shoaibcasseem@gmail.com
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Swalay Aboo Fedally
Drswalay@yahoo.com
Department of Nephrology, Sir Seewoosagur Ramgoolam National (SSRN) Hospital
Pamplemousses, Mauritius

Zahra Mungloo-Dilmohamud
z.mungloo@uom.ac.mu
Department of Digital Technologies, University of Mauritius
Reduit, Mauritius

ABSTRACT

Over the years, machine learning and deep learning have shown tremendous potential and good accuracy in the detection of Chronic Kidney Disease (CKD). Yet, these models are being criticised for their opaque decision-making process. To date, few studies have been conducted to look into which variables are more relevant in determining CKD. In this study, we first used two (2) ensemble approaches: XGBoost, AdaBoost. for detecting CKD. Next, we propose a model that uses deep neural networks (DNN) to extract and represent the features for the prognosis of CKD. The DNN was enhanced by fine-tuning the hyperparameters and it outperformed the ensemble methods with an accuracy of 98.8%. Finally, SHapley Additive exPlanations (SHAP), which is a technique that has attracted attention in the field of explainable artificial intelligence (XAI), was used to investigate which features contribute to predicting CKD and the results confirmed the medical explanations.

CCS CONCEPTS

• **Computing methodologies → Causal reasoning and diagnostics**.

KEYWORDS

CKD, Neural network, Explainable AI, SHAP

ACM Reference Format:
Khadiime Jhumka, Muhammad Muzzammil Auzine, Maleika Heenaye-Mamode Khan, Mohammad Shoaib Casseem, Swalay Aboo Fedally, and Zahra Mungloo-Dilmohamud. 2023. Explainable Chronic Kidney Disease (CKD) Prediction using Deep Learning and Shapley Additive Explanations (SHAP). In *2023 The 7th International Conference on Advances in Artificial Intelligence (ICAAI) (ICAAI 2023), October 13–15, 2023, Istanbul, Turkiye*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3633598.3633604

ICAAI 2023, October 13–15, 2023, Istanbul, Turkiye

ACM ISBN 979-8-4007-0898-5/23/10…$15.00
https://doi.org/10.1145/3633598.3633604

1 INTRODUCTION

Chronic Kidney Disease (CKD) is a major health concern in many countries. According to [15] the prevalence of CKD is 13.4% worldwide and patients requiring renal replacement for final stages of kidney disease is estimated to be between 4.9 and 7.1 million. CKD also impacts on the risk of cardiovascular diseases and therefore directly influences morbidity and mortality globally [9]. Hence, early diagnosis of CKD is an essential step in hindering the progress of kidney damage and failure. Artificial Intelligence (AI) and big data have been found to address many unresolved nephrology issues and aid in the diagnosis, prognosis and decision making of renal disease [27]. It is therefore crucial to leverage on these technologies. Much research has been done on the use of AI to predict or diagnose CKD using textual data but most of these work have focused on using Machine Learning (ML) algorithms [4, 19] with few cases focusing on neural networks [26]. In this research, two(2) ML algorithms and one (1) Deep Learning (DL) algorithm have been investigated. Furthermore, since DL algorithms are often considered as black boxes and their inner workings are not always transparent, explainable artificial intelligence has been investigated through the use of Shapley Additive Explanations (SHAP). Explainable artificial intelligence (XAI) englobes all the means that humans can use to better understand the results produced by AI algorithms and SHAP is the state of the art in XAI [10]. The aim of this work is two-fold: firstly, to find out which algorithm performs better with textual data and secondly, to try to explain how the proposed DL model has come to its decision through the use of SHAP. The rest of the paper is structured as follows: section 2 summarises the state of the art in the field, section 3 provides details of the materials and methods, section 4 presents the results and discussions and section 5 concludes the paper.

2 RELATED WORKS

Researchers have explored different techniques to detect abnormalities and artefacts in the kidney. Earlier, hand-crafted techniques were used to identify abnormalities in the kidney. The dataset used in the different works cited in this review section is the CKD dataset that was released by Apollo Hospitals, Tamil Nadu, India, which is now publicly available on the UCI machine learning repository. The dataset consists of data from 400 patients with 24 attributes. In the research done by Pinto et al. [18], the J48 algorithm, a ML model using decision tree as the classifier, was trained using the selected chronic kidney disease (CKD) dataset. This model achieved the highest test results compared to the other ML models such as Artificial Neural Networks (ANN), Naive Bayes (NB), k-Nearest Neighbors (kNN), Support Vector Machine (SVM). On the oversampled CKD dataset, the J48 algorithm achieved the following results: 97.66% of accuracy, 96.13% of sensitivity, 98.78% of specificity, and 98.31% of precision. Meanwhile Saroja et al. [23] proposed a Neural network classifier called Novel Weight Convolution Neural Network (NWCNN) classifier together with Tensor factorization and Adaptive Neuro-Fuzzy Inference System (ANFIS) for data imputation of missing values and Adaptive Weight Dynamic Butterfly Optimization Algorithm (AWDBOA) as a feature selection tool. Training this hybrid classifier on the UCI Chronic kidney disease dataset achieved a classification accuracy of 99.048% which is higher than that achieved by classic ML Models. Our review of existing literature shows that machine learning has tremendous potential in classification. One drawback of machine learning techniques is that they depend on hand-crafted features, which are extracted and selected by algorithms, which usually lack generality. On the other side, deep learning makes use of the convolution layers to extract the different features but the algorithm behaves like a black box. There is a need to determine the vital parameters that determine the decision, leading to the development of explainable AI. This can be seen in the work by Rashed-Al-Mahfuz et al. [21]. Firstly, the data were pre-processed and machine learning algorithms were used for feature selection and extraction. Classifiers such as random forest (RF), gradient boosting (GB), XGBoost (XGB), Logistic Regression (LR), and support vector machine (SVM) were adopted. Eventually, the SHAP technique was used to identify the principal features that most influence the classification decision. However in this study, a feature selection was already used before the classification, thus reducing the probabilities of the model discovering new patterns from the less important features. From some of the work conducted so far, there is scope in the development of interpretable AI in the medical field using custom-made deep learning models with no feature selection tools. The contributions of this paper are as follows:

3 MATERIALS AND METHODS

3.1 Workflow

Figure 1 shows the workflow for the proposed system. For this work, the open dataset obtained from the UCI repository of datasets was used. The workflow is as follows (1) downloading and analysing the dataset (2) data pre-processing (3) training the classifiers (4) evaluating the models and retraining the model, if needed (5) analysing the results.

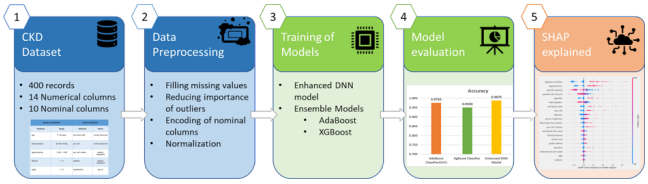

Figure 1: Workflow of our proposed system

3.2 Dataset

The dataset utilised in this study for the diagnosis and prediction of chronic renal disease, was from the UCI dataset repository (https://archive.ics.uci.edu/ml/datasets/chronic kidney disease) and it was gathered from a variety of patients in India. The collection is made up of records from 400 individual patients, each with their own set of 25 features, out of which, 13 are nominal, 11 are numerical and the 25th is the class. The 24 attributes are age, blood pressure, specific gravity, albumin, sugar, red blood cells, pus cell, pus cell clumps, bacteria, blood glucose random, blood urea, serum creatinine, sodium, potassium, haemoglobin, packed cell volume, white blood cell count, red blood cell count, hypertension, diabetes mellitus, coronary artery disease, appetite, pedal edema, and anaemia. The attribute "class," which is a measurable field, is the output data with the value "ckd" denoting a person with CKD, while "notckd" indicates an individual without CKD. There are 250 "ckd" and 150 "notckd" occurrences in the dataset.

3.3 Data Preprocessing

In this work, the preprocessing steps adopted are as follows: dealing with noise such as outliers, estimating missing values and normalisation. The first process involves dealing with outliers. In some studies, outliers are described as noise and they can alter and confuse the machine learning algorithm's learning process. As a result, they are removed from the dataset. In this work, instead of removing the outliers, we have reduced their importance in the model training by using transformation techniques to obtain a bell-shaped distribution as used by [16]. First, we determined the skewness of each number column, which informs us about the direction of outliers. Following that, we performed log transformation to the features with substantially positively skewed values in comparison to the others. In the case of significantly negatively skewed data, the Box-Cox transformation was applied. Using this method, the skewness of each column is quite comparable to the skewness of the other numerical columns, minimising the errors' normalcy [3]. Missing values are typical in real-world data, and 1012 missing values were discovered in this dataset. Removing rows containing the missing value can lead to bias in the learning process and the loss of valuable knowledge [14]. Thus the missing values in the nominal columns were replaced with the mode value for that attribute. As per[11], in the case of numerical categories, missing values were replaced using the values from the same column by a kNN imputer (k-Nearest Neighbours). For numerical columns, [8, 17] have shown that kNN imputation is one of the most suitable imputation techniques used to replace missing values. For the encoder part, the data which includes the words, 'normal', 'notpresent', 'no', 'good' and 'notckd' was converted to 0 whereas 1 was attributed to the data

labelled ‘abnormal’, ‘present’, ‘yes’, ‘poor’ and ‘ckd’. The dataset was then normalised using the MinMaxScaler normalisation approach. According to [20], normalisation before classification leads to considerably more accurate performance, and also MinMaxScaler outperformed all the other normalisation techniques in the case of Ensemble models and DNN models. Before using the preprocessed dataset for classification, it was divided randomly into the training set and the testing set in the ratio of 80:20. In the instance of training the DNN Model, we further divided the training set in a 75:25 ratio, with one-quarter of the training set utilised as validation data to enhance the model.

3.4 Ensemble models

Ensemble models proves to be an accurate classifiers in previous works [1, 25]. The ensemble approach is a technique for increasing classifier accuracy [6]. The ensemble approach combines models to generate a better composite model that improves performance. The ensemble technique’s fundamental principle is to join numerous "weak learners" to create a "strong learner." The first model is the AdaBoost classifier. AdaBoost iteratively trains weak learners and assigns each one a weight, which indicates the robustness of the weak learner. The final model is made up of the final categorization of the weak learners as well as all of their updated weights. In this study, Support Vector Classifier (SVC) was employed as the weak learner for the AdaBoost classifier. For the SVC AdaBoost classifier, the parameters were adjusted as follows: the number of estimators is set to 12 and the learning rate is set to 1. The second model, XGBoost, employs gradient descent on decision trees to produce numerous models that are sequentially integrated while correcting the prior models to generate a final optimal model. Based on the work by Murty and Kumar [17], the following parameters were tuned as follows for the XGBoost classifier: the learning rate is set to 0.05, the max depth is set to 3, the number of estimators (Decision Trees) is set to 50, and gamma is set to 0.2.

3.5 Enhanced DNN model

A deep neural network is a complex neural network structure with several hidden layers between the input and output layers. A neural structure is intended to predict outcomes and uncover hidden structures and patterns in a dataset. Except for the class property, the neural network in this research uses all remaining 24 attributes as input. Figure 2 shows the proposed DNN model used in this work. For the concealed layers, 5 dense layers of varying sizes were employed. The first dense layer has 24 neurons which is the same as the number of attributes. As discussed in [12], the larger the number of units in the hidden layer, the better is the performance of the classification. So, for the next three dense layers, each of them contain 576 neurons, which is a factor of 24. According to [5], the activation function introduces nonlinearity into our model and enhances pattern learning for our complicated data. The activation function employed in this work for the hidden layers is ReLU, which is the first-choice activation function proposed by [24]. The output layer is then linked to the final hidden layer. Because the categorization is binary (CKD or notCKD), the Sigmoid function is employed as the activation function for the output layer. A deep neural network is seen as a black box, and the model can learn a large number of hidden patterns that may or may not be beneficial. To evaluate and validate the model’s performance, the model applies the learned patterns to the validation data.

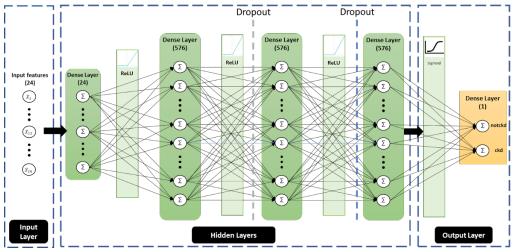

Figure 2: Proposed DNN Model

For the model to be used for the classification, several hyperparameters were tuned. The loss function applied to the output layer is the binary cross entropy and the optimizer used for this model is the adam optimizer with a learning rate of 1x10-4. Since the hidden layers have a lot of neurons, it can lead to model overfitting. To prevent this from happening, an early stopping function is inserted to the model. As a result, the model stops the training when the validation loss reaches its minimum point. Eventually, the model is set to train with the training data and the final model is used to predict the test data and the results were compared to that predicted from the other classifiers.

3.6 Explainability and SHAP library

There is an ever-increasing need to have transparency of how different AI models have come to a decision but this is very challenging given the high dimensionality and complexity of the data being handled as well as the model used. However, this understanding is crucial in healthcare where critical decision making is being entrusted to AI models. Currently, AI models are unable to explain how they have reached a decision and humans are unable to understand how the models have reached a solution. SHapley Additive exPlanations (SHAP) has proved to be successful in explaining different learning models [2]. In this work, SHAP has been used to calculate Shapely values which are measures of the contributions each feature has in the trained model. The SHAP Summary plot which gives an overview of all the features that were involved in the classification process has been used.

4 RESULTS AND DISCUSSIONS

To evaluate the classifiers, we applied them to a blind test data set, and the experimental results are shown below. To begin, we evaluate the three models using several assessment criteria which includes overall accuracy, recall, and the F1-score. By making the CKD class value positive and the notCKD class value negative, the model accuracy was determined. The ratio of True Positive (TP) and True Negative (TN) instances to total observations is the overall accuracy. The recall measure specifies the proportion of TP instances to the sum of TP and TN cases. Precision is defined as the ratio of TP cases to expected CKD patients overall. The F1-score is the harmonic mean of precision and recall.

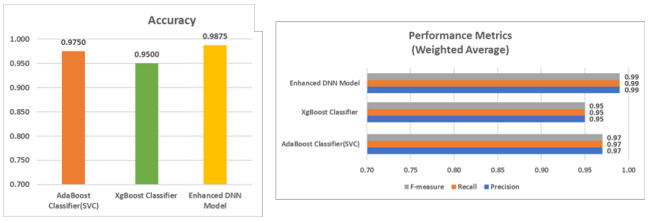

Figure 3: Performance of the 3 models

Figure 3 show the performance achieved by the 3 models. Our proposed DNN model has the maximum testing accuracy of 98.8 percent. The enhanced DNN Model outperforms the other ensemble models. The metrics in Figure 3 emphasise that the proposed DNN model performed better for the classification task than the other classifiers, having higher accuracy, higher recall, and higher F1-score. The enhanced DNN model has achieved a relatively smaller misclassification rate compared to the other classifier model as it has misclassified only one nonCKD case. This can be observed by the 0.01 error in the three classification metrics as shown in Figure 3. However, as all of the classifier assessment metrics are higher than 95 percent, we can clearly see that the models are reliable models for predicting CKD and notCKD cases with the DNN better.

Although we obtained very good overall testing accuracy using our enhanced DNN model, we do not know how the model came to this result. Hence SHAP was selected for an in depth analysis of the model. SHAP, as described above, is a method that aids in the explanation of the output of our proposed DNN model. Using SHAP, we can also investigate the important factors that the model has used to predict CKD. The SHAP summary is generated to determine which variables have a strong predictive importance. The characteristics/parameters in Figure 4 are organised in order of relevance, with the features having the largest influence on CKD at the top and the features having the least impact at the bottom. As shown in Figure 4, specific gravity has the biggest influence on categorization, followed by diabetes, albumin, and so on. Similarly, potassium and white blood cell count are towards the bottom of the SHAP summary plot, indicating that the model deducts that they have little effect on CKD classification. Red blood cell count, blood urea, blood pressure, and age have the least influence on CDK prediction and were therefore excluded from the SHAP summary figure.

The SHAP summary plot also demonstrates how those individual parameters affect CKD (negatively or positively). The data points in Figure 4 that are directly on the right side of zero on the horizontal axis help in estimating how those features result in high CKD, while the data points on the left show how those features result in low CKD. For example, when we consider features such as albumin, we deduce that the higher the albumin level, the higher the risk of getting CKD. Nonetheless, in specific gravity, it appears that higher specific gravity results in low risk of CKD. Hence albumin is a positive impact feature whereas specific gravity is a negative impact feature. In Figure 4, the positive impact features include diabetes mellitus, albumin, pedal edema, pus cell, appetite, hypertension, serum creatinine, red blood cells, blood glucose random, anaemia, pus cell clumps, coronary artery disease, bacteria, potassium and white blood cell count while specific gravity, packed cell volume, haemoglobin, sodium, sugar are the negative impact features. Summary plots are rich with information and offer a general overview of SHAP values for several attributes at once. However, in order to fully comprehend the link between a feature's values and the model's expected outcomes, dependence plots must be examined. Figure 5 represent the dependence plots of the top 5 numerical variables. The x-axis represents the percentile of the values while on the y-axis, the line y=0 represents the line when the model switches prediction from CKD to notCKD. It is also known as the cutoff region. With the MinMaxscaler function, the minimum value of each attribute was 0 and the largest value was 1. The features selected for numerical attributes are specific gravity, albumin, serum creatinine, packed cell volume and haemoglobin.

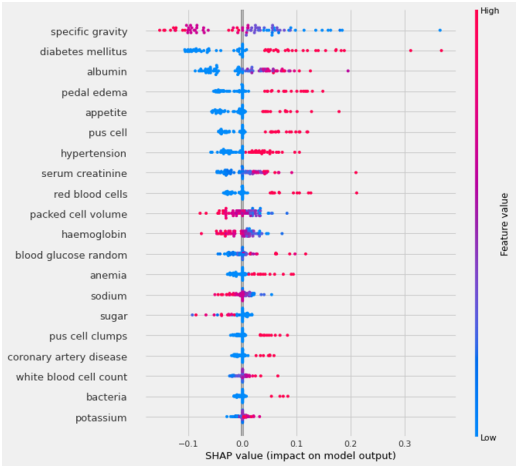

Figure 4: Numerical attributes individual SHAP graph

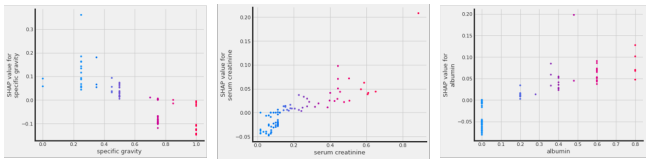

Figure 5: Numerical attributes individual SHAP graph

Specific gravity, often known as urine specific gravity (USG), is the most important characteristic in Figure 5 for determining CKD classification. The enhanced DNN Model predicts a cutoff zone around the 0.7 percentile. The 0.7 percentile corresponds to a urine specific gravity of 1.018. According to Sajadi et al [22], a USG value of less than 1.010 may indicate CKD. The model cutoff area for albumin is 0.2 percentile, which corresponds to an albumin level of 1. As noted in Lang et al. [13], an albumin level greater than one increases the risk of renal failure. When the dependency graphs of specific gravity and albumin are compared, larger values of specific gravity have negative SHAP values whereas higher values of albumin have positive SHAP values, demonstrating that specific gravity is a negative impact feature while albumin is a positive

impact feature. From the serum creatinine dependence plot, it can be observed that the higher the value, the more likely the patient will suffer CKD. Serum creatinine is a very useful component together with age, gender and ethnicity in calculating estimated glomerular filtration rate (eGFR) which indicates CKD stage. A high value of eGFR is an indication that the patient is suffering from CKD [7].

This research is subject to one limitation related to the dataset used. The size of the dataset is considered to be small (400 instances), which may influence the reliability of the results, for example the age variable. On the other hand, we faced difficulty acquiring another dataset that has the same features to compare the results of the current dataset.

5 CONCLUSION

We have explored deep neural networks (DNN) and ensemble learning to determine CKD, since these techniques have spurred much interest in their application in the medical problem. So far, researchers have mostly focussed on handcrafted techniques. In this study, we have used the chronic renal disease datasets available from the UCI dataset repository, whereby data was collected from a variety of patients in India. The data was first pre-processed to cater for outliers and missing values. Initially, we investigated two ensemble models namely: XGBoost and AdaBoost which have brought equally promising results in the medical domain. In order to check for overfitting or underfitting, the DNN was further enhanced. All the 24 attributes listed in the dataset were taken as input in the DNN. The enhanced DNN achieved a performance accuracy of 98.8% which was higher than the accuracy achieved by the ensemble learning methods. However, there is a lack of transparency as to which of the 24 attributes predicted CKD. To bridge this gap, SHAP, which is a recent tool that allows the study of per-class feature significance, has been adopted to show the impact of each feature on the decision making. SHAP was able to identify the features that are the most relevant for the prognosis of CKD when using the DNN. In this work, we have discussed the SHAP features that have been able to provide a clear explanation on the interpretability of the features that determine CKD. Considering, the above mentioned limitations, as a future work, a larger local dataset with the same features can be created and this can be used for more accurate interpretable results.

REFERENCES

[1] Deema Mohammed Alsekait, Hager Saleh, Lubna Abdelkareim Gabralla, Khaled Alnowaiser, Shaker El-Sappagh, Radhya Sahal, and Nora El-Rashidy. 2023. Toward Comprehensive Chronic Kidney Disease Prediction Based on Ensemble Deep Learning Models. *Applied Sciences* 13, 6 (2023), 3937.

[2] Liat Antwarg, Ronnie Mindlin Miller, Bracha Shapira, and Lior Rokach. 2021. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. *Expert systems with applications* 186 (2021), 115736.

[3] Anthony C Atkinson, Marco Riani, and Aldo Corbellini. 2021. The box–cox transformation: Review and extensions. (2021).

[4] Anusorn Charleonnan, Thipwan Fufaung, Tippawan Niyomwong, Wandee Chokchueypattanakit, Sathit Suwannawach, and Nitat Ninchawee. 2016. Predictive analytics for chronic kidney disease using machine learning techniques. In *2016 management and innovation technology international conference (MITicon)*. IEEE, MIT–80.

[5] Hoon Chung, Sung Joo Lee, and Jeon Gue Park. 2016. Deep neural network using trainable activation functions. In *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 348–352.

[6] Thomas G Dietterich. 2000. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine learning* 40 (2000), 139–157.

[7] George Drosos, Fotini Ampatzidou, Pantelis Sarafidis, Theodoros Karaiskos, Athanasios Madesis, and Afroditi K Boutou. 2018. Serum creatinine and Chronic Kidney Disease-Epidemiology Estimated Glomerular Filtration Rate: independent predictors of renal replacement therapy following cardiac surgery. *American Journal of Nephrology* 48, 2 (2018), 108–117.

[8] Tlamelo Emmanuel, Thabiso Maupong, Dimane Mpoeleng, Thabo Semong, Banyatsang Mphago, and Oteng Tabona. 2021. A survey on missing data in machine learning. *Journal of Big Data* 8, 1 (2021), 1–37.

[9] Minhaz Uddin Emon, Rakibul Islam, Maria Sultana Keya, Raihana Zannat, et al. 2021. Performance analysis of chronic kidney disease through machine learning approaches. In *2021 6th International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 713–719.

[10] Alex Gramegna and Paolo Giudici. 2021. SHAP and LIME: an evaluation of discriminative power in credit risk. *Frontiers in Artificial Intelligence* 4 (2021), 752558.

[11] Sebastian Jäger, Arndt Allhorn, and Felix Bießmann. 2021. A benchmark for data imputation methods. *Frontiers in big Data* 4 (2021), 693674.

[12] VL Helen Josephine, AP Nirmala, and Vijaya Lakshmi Alluri. 2021. Impact of hidden dense layers in convolutional neural network to enhance performance of classification model. In *IOP Conference Series: Materials Science and Engineering*, Vol. 1131. IOP Publishing, 012007.

[13] Joshua Lang, Ronit Katz, Joachim H Ix, Orlando M Gutierrez, Carmen A Peralta, Chirag R Parikh, Suzanne Satterfield, Snezana Petrovic, Prasad Devarajan, Michael Bennett, et al. 2018. Association of serum albumin levels with kidney function decline and incident chronic kidney disease in elders. *Nephrology Dialysis Transplantation* 33, 6 (2018), 986–992.

[14] Roderick JA Little and Donald B Rubin. 2019. *Statistical analysis with missing data*. Vol. 793. John Wiley & Sons.

[15] Ji-Cheng Lv and Lu-Xia Zhang. 2019. Prevalence and disease burden of chronic kidney disease. *Renal fibrosis: mechanisms and therapies* (2019), 3–15.

[16] Leigh Metcalf and William Casey. 2016. Introduction to data analysis. *Cybersecurity and applied mathematics* (2016), 43–65.

[17] Sivala Vishnu Murty and R Kiran Kumar. 2019. Accurate liver disease prediction with extreme gradient boosting. *Int. J. Eng. Adv. Technol* 8, 6 (2019), 2249–8958.

[18] Ana Pinto, Diana Ferreira, Cristiana Neto, António Abelha, and José Machado. 2020. Data mining to predict early stage chronic kidney disease. *Procedia Computer Science* 177 (2020), 562–567.

[19] Jiongming Qin, Lin Chen, Yuhua Liu, Chuanjun Liu, Changhao Feng, and Bin Chen. 2019. A machine learning methodology for diagnosing chronic kidney disease. *IEEE Access* 8 (2019), 20991–21002.

[20] VN Ganapathi Raju, K Prasanna Lakshmi, Vinod Mahesh Jain, Archana Kalidindi, and V Padma. 2020. Study the influence of normalization/transformation process on the accuracy of supervised classification. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 729–735.

[21] Md Rashed-Al-Mahfuz, Abedul Haque, Akm Azad, Salem A Alyami, Julian MW Quinn, and Mohammad Ali Moni. 2021. 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* 9 (2021), 1–11.

[22] Sahar Sajadi, Ching Yu, Jean-Daniel Sylvestre, Karl J Looper, Marilyn Segal, and Soham Rej. 2016. Does lower urine-specific gravity predict decline in renal function and hypernatremia in older adults exposed to psychotropic medications? An exploratory analysis. *Clinical Kidney Journal* 9, 2 (2016), 268–272.

[23] T Saroja and Y Kalpana. 2023. Hybrid missing data imputation and novel weight convolution neural network classifier for chronic kidney disease diagnosis. *Measurement: Sensors* 27 (2023), 100715.

[24] Tomasz Szandała. 2021. Review and comparison of commonly used activation functions for deep neural networks. *Bio-inspired neurocomputing* (2021), 203–224.

[25] Vinoth Kumar Venkatesan, Mahesh Thyluru Ramakrishna, Ivan Izonin, Roman Tkachenko, and Myroslav Havryliuk. 2023. Efficient Data Preprocessing with Ensemble Machine Learning Technique for the Early Detection of Chronic Kidney Disease. *Applied Sciences* 13, 5 (2023), 2885.

[26] Piervincenzo Ventrella, Giovanni Delgrossi, Gianmichele Ferrario, Marco Righetti, and Marco Masseroli. 2021. Supervised machine learning for the assessment of chronic kidney disease advancement. *Computer Methods and Programs in Biomedicine* 209 (2021), 106329.

[27] Lijing Yao, Hengyuan Zhang, Mengqin Zhang, Xing Chen, Jun Zhang, Jiyi Huang, and Lu Zhang. 2021. Application of artificial intelligence in renal disease. *Clinical eHealth* 4 (2021), 54–61.