Severity Prediction and Multi Classification of Chronic Kidney Disease Based on Machine Learning Techniques

Moataz Mohamed El Sherbiny
Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Egypt, moatazelsherbiny@mans.edu.eg

Eman Abdelhalim
Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Egypt

Hossam El-Din Mostafa
Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Egypt

Abstract

One of the major causes of morbidity and mortality from non-communicable diseases is chronic kidney disease (CKD). Symptoms do not appear till kidneys lose most of its functionality. Hence, early and precise CKD stage detection can minimize the impact on the health of patients. Moreover, Further complications such as hypertension, anemia, brittle bones and nerve damage can be reduced. Recently, machine learning techniques are widely employed for the prediction and classification of diseases in healthcare system. This study focuses on the use of machine learning techniques for specific stage prediction and detection of CKD. The proposed model involves applying a set of seven distinct ML based classification models such as Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Extreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost). Several experiments were conducted in this study including different data imputation techniques and feature selection methods. The assessment of these models has done based on four performance metrics including accuracy, precision, f-measure, and recall. Results had indicated that XGBoost and RF outperformed other techniques.

Keywords: *Chronic Kidney Disease (CKD), Machine Learning, Multi-classification, Performance Measures, Feature Selection.*

1. INTRODUCTION

Kidneys are vital organs that purify the blood by eliminating extra waste, which is then expelled from the body as urine. The steady deterioration of kidney function is referred to as kidney disease. Chronic kidney disease is one of the prevalent causes of death and suffering in the twenty-first century. Chronic kidney disease affects nearly 10% of the world's population [1]. Furthermore, the number of patients with CKD has been rising, with an estimated 843.6 million people diagnosed globally in 2017. CKD is the 11th deadliest cause of mortality worldwide with 1.2 million deaths annually [2]. Currently, it is the sixth most rapidly growing reason of death globally. Hence, the prediction of CKD in early stages is considered a quite essential process, because it could enable patients to

receive a timely effective treatment to ameliorate the progression of the disease. Since CKD is a progressive and irreversible pathologic syndrome[3]Besides, the treatment and medication are neither accessible nor affordable in the majority of developing countries.

Chronic Kidney Disease is classified into five distinct stages depending on the deterioration of kidney functionality and reduced Glomerular Filtration Rate (GFR) according to national kidney foundation. GFR measures a level of kidney function. Stage five is considered as end-stage renal disease (ESRD). Stage one and two are considered mildest stages known with only few symptoms.

Machine learning (ML) calculates and deduces the information related to the task. In addition to obtaining the properties of the corresponding pattern [4]. It has been utilized to detect numerous diseases and cancers [5]. Therefore, ML is considered a promising method for diagnosis of CKD.

Most of previous studies employed the CKD data set that was obtained from the University of California Irvine (UCI) machine learning repository. The UCI contains 400 sample records and 24 features. Which is considered quite a lot of features for a relatively small-size dataset. The number of complete instances without any missing values is 158. Additionally, the target classification was binary either considered a ckd patient or not. Therefore, No severity or stage prediction. A summary of related work is presented in Table 1.

Charleonnan et al. [6] employed the use of Decision Tree, Support Vector Machine and Logistic Regression on the Indians CKD dataset as predictive models for classification and prediction of CKD. Their results indicated that SVM reached the highest classification accuracy. Salekin et al. [7] preprocessed the UCI dataset and reduced the number of features used in prediction to 14 attributes. SVM was stated as the best model scoring and accuracy of 96.75%. Xiao et al. [8] used a different dataset with 551 patient and 18 features. The outcome of classification was mild, moderate and severe. Several machine learning techniques were utilized including logistic regression, random forest, support vector machine and neural network. Logistic regression achieved the best performance with 0.83 sensitivity and 0.82 specificity. Yashfi et al. [9] employed feature selection techniques to reduce dimensionality of UCI dataset. They have extracted the top twenty relevant features for classification. Their proposal for risk prediction of CKD indicated that Random Forest reached the highest accuracy of 97.12%. Vinoid [10] conducted seven machine learning techniques including Naïve Bayes, Support Vector Machine, Logistic Regression, Decision Tree, Neural Network, K-Nearest Neighbor, and Random Forest. KNN was indicated as the best performer based on different evaluation method for reaching 97% accuracy. Debal et al. [11] included Random Forest, Support Vector Machine and Random Forest achieving an accuracy of 78.3%, 63% and 77.5% respectively. Random forest reached 79% when half of the total number of features were selected. Rady et al. [12] showed that Neural Network (NN) outscored Support Vector Machine in terms of prediction performance achieving accuracy of 96.7%. Mohsin et al. [13] tested K-Nearest Neighbor (KNN), Decision Tree (DT), Neural Network (NN) and Naïve Bayes (NB). In comparison to previous methodologies, their prototype claimed that Nave Bayes had a superior accuracy of 94.6%.

Table 1. Summary of related work

Author	Technique	Target Classification	Accuracy
Charleonnan et. al	SVM	Binary	0.98
Salekin et al.	SVM with feature selection	Binary	0.96
Xiao et al.	LR	Multi	0.87
Yashfi et al.	RF	Binary	0.97
Vinoid et al.	KNN	Binary	0.97
Debal et al.	RF with feature selection	Multi	0.78
Rady et al.	Neural Network	Binary	0.96
Mohsen et al.	Naïve Bayes	Binary	0.94

This study aims to improve the prediction model's accuracy for CKD The proposed methodology is covered in detail in Section 2. Section 3 delves into the evaluation and discussion of experimental results. Finally, work conclusions are discused in Section 4. The contributions of the proposed work are as follows:

1) Different imputation techniques including mean approach and KNN imputation for handling missing values in the dataset.

2) Feature selection techniques for removal of irrelevant attributes and involving the most crucial features.

3) Use of various ML models for severity prediction of chronic kidney disease and determination of specific CKD stage.

2. METHODS

This section contains dataset description, data preprocessing, machine learning algorithms and evaluation methods as shown in Figure 1.

A. Dataset Description

The data source used in this paper is obtained from St. Paulo's Hospital. It is regarded as Ethiopia's second-largest public hospital and treats a significant number of patients with renal diseases. The dataset contains 1718 sample records with 18 features in addition to target class. The attributes include 12 numerical features and 7 nominal ones. The multi stage target class distribution as follows: 276 instances are considered as not CKD patients or at normal stage (Stage I), 248 patients at mild stage (Stage II), 354 samples at moderate stage (stage III), 399 patients at severe level (Stage IV), and 441 instances of end-stage renal disease (V). Dataset description is summarized in Figure 2.

Figure 1. Block diagram of the proposed system

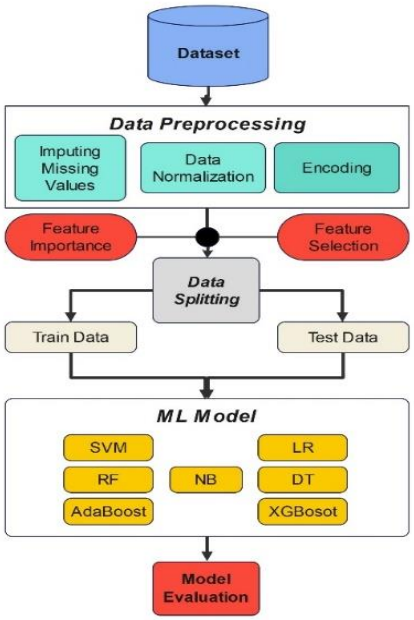
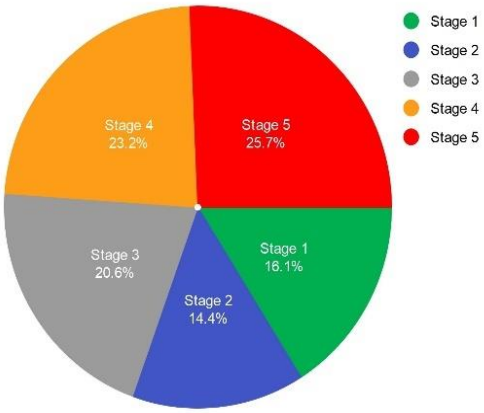


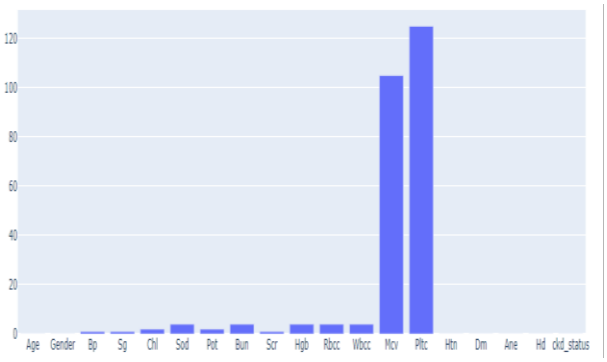
Figure 2. The multi class distribution percentage



B. Data Preprocessing

Preprocessing is a very crucial component of developing the prediction model in our study. Since the inconsistent data alter the accuracy of the model, considering that the gathered data includes missing values and nominal variables. Therefore, cleaning noisy data must be performed. The dataset contains a number of missing values as shown in Figure 3. Patients frequently overlook several measures for several reasons. As a result, missing values will show up in the data if the diagnostic categories of the samples are unidentified, necessitating the use of an appropriate imputation method. Data preprocessing is implemented in four different stages which are imputation, encoding, scaling and feature selection.

Figure 3. Missing values count in the dataset



1) Imputation:

There are various common and widely used strategies for dealing with non-existing values such as replacing them with a constant, mean, median or the most frequency value. Theoretically, replacing missing feature values with zero has no effective biasing in prediction. However, this assumption is practically impossible in medical dataset. In this paper, two methods were conducted. First one is simple imputation technique which involves replacing null values with the mean value, while second one is K-Nearest Neighbor imputer with different k values. When the values of the numerical variables in the K complete samples are sorted by numerical value, K is optimally set as an odd number since the middle value in this scenario is clearly the median.

2) Encoding:

Data must be transformed into required format to ease processing. Therefore, Nominal values need to be converted into numbers to make machine learning algorithm able to understand data it receives. Handling Categorical variables is conducting through encoding process. Ordinal encoder and label encoder were implemented for nominal features and target class respectively.

3) Scaling:

Before fitting any models, it is usually vital to scale numeric descriptive features since several significant classes of techniques require it such as SVM and other ML Algorithms since scaling facilitates the ability of model to learn and comprehend the problem [14]. Standard scaling method which adjusts the attribute to 0 mean and 1 standard deviation, was implemented in this work. Normalization and standardization are the two most effective scaling techniques.

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

Where; z is Z-score, x is feature, μ is the mean, and σ is the standard deviation.

4) Feature Selection:

It is a mechanism for minimizing the number of irrelevant input variables that do not have a significant contribution on the target variable [15]. As it identifies a subset of relevant predictive features which is crucial for better relative accuracy. Consequently, the issue of high dimensionality is minimized as possible. There are several types of feature selection including filter and wrapper. The filter method is a commonly used approach since it is independent of the learning algorithm. The wrapper method uses classification to choose relevant features. In terms of accuracy, it is superior to Filter method. Regrettably, it requires more time to process [16]. Uni-variate Feature Selection (UFS) represent the most common, straightforward, and quick method used in medical dataset, where each feature is being taken into account separately to evaluate how strongly a feature is related to the dependent variable. And furthermore, it is classifier independent. Different options are available for univariate algorithms including information gain, Pearson correlation, Chi-square and Analysis of variance (ANOVA). The ANOVA test was implemented in this study for continuous numerical features while Chi-Square was used for categorical nominal ones.

$$F = \frac{MST}{MSE} \qquad (2)$$

Where; F is the ANOVA coefficient, MST is the Mean Squares of Treatments, and MSE is the Mean Squares of Errors.

$$X^2 = \sum \frac{(Observed\ Value - Expected\ Value)^2}{Expected\ Value} \quad (3)$$

Where; X^2 is Chi-Square Test.

C. Machine Learning Algorithms

Machine learning algorithms classify or predict data without explicit programming after going through the training phase. Seven machine learning techniques had been utilized in this study. In order to determine the best machine learning technique that provides the highest classification performance thorough a comparative analysis of the tested algorithms. Methodologies that have been tested includes Random Forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), Extreme Gradient Boosting (XGB), Adaptive boosting (ADA).

Random Forest: A learning algorithm that develops numerous decision trees during the training phase and pro- vides output class of those individual trees. Regression and classification can both be employed [17]. This model makes a slight adjustment that makes use of the de-correlated tree by bagging, which is the development of numerous decision trees from training data using bootstrapped samples. A specified number of feature columns are removed from the total number of feature columns during bootstrapping. Bootstrap modelling increases bias while minimizing variance.

Naive Bayes: A probability-based model is a supervised algorithm that necessitates feature independence for classifying data. This model works well for datasets with a large number of input attributes. It encompasses every feature that is provided, even some that have minor effect on the outcome of the prediction [18].

Support Vector Machine: A Decision plane-based model is one of the most robust statistical learning framework-based algorithms that provides a solution for both regression and classification problems as well as both linear and non-linear datasets [19]. Every data point is regarded as an n-dimensional vector, and a (n-1) hyper plane

divides the datasets. A hyper plane is a line that splits a plane into two halves in a two-dimensional space.

Decision Tree: A supervised learning approach, whose purpose is to comprehend basic chained decision rules from prior input variables in order to train a model to classify a target variable [20] A set of impurity criteria is applied to recursively separate the variables until a set of stopping requirements are met. Gini impurity is chosen for the model from a variety of impurity measuring techniques.

Logistic Regression: In the healthcare system, logistic regression is a well-known supervised learning algorithm [21]. Logistic regression predicts the probability of the class output using a set of independent features. Assuming that p is the probability of a subject belongs to the CKD class, therefore 1-p is the probability of a subject belongs to the non-CKD class. Decision boundary is the threshold set to determine which data belongs to certain class This classification probability is calculated using the logistic sigmoid function.

XGBoost: An Extreme gradient boosting is a tree-based sequential decision trees algorithms [22]. It is regarded as one of the most efficient methods for performing classification and predictions on small to medium-sized structured or tabular datasets. It uses a gradient descent architecture to accurately estimate a target variable or feature, through integrating relatively weaker and simpler models. One of XGBoost's most significant aspects is scalability, where it directs abrupt learning through parallel and distributed computing as well as provides well-structured memory usage [23].

AdaBoost: Adaptive boosting is an iterative machine learning algorithm that is less prone to over-fitting of data. Where dataset is split into two partitions for each iteration, the features used in the first iteration will be given less weight, and the incorrectly classified data are given more weight in the next iteration. When all iterations are finally completed, they are merged with appropriate weights to yield a powerful and effective classifier that predicts the classes of the unseen data [24].

D. Evaluation Methods

The most prominent performance measurements are precision, F1- score, sensitivity (recall), and accuracy. True positives (TP), false positives (FP), true negatives (TN), and false negatives are the four variables needed by the evaluation methods (FN).

- Accuracy: This is the percentage of cases that were correctly identified out of all the cases

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

- Precision: It is the ratio of correctly predicted positive outcomes to all positive outcomes.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

- Recall: It is the proportion of correctly predicted events among the foreseen data.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

- F1-Score: It is the weighted average of precision and recall.

$$F1 - Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (7)$$

- Sensitivity: It is the mean proportion of actual true positives that are correctly identified.

$$Sensitivity = \frac{TP}{TN+FP} \quad (8)$$

• Specificity: It is used to measure the fraction of negative values that are correctly classified.

$$Specificty = \frac{TN}{TN+FP} \quad (9)$$

3. RESULTS AND DISCUSSION

Dataset is split into 70% train size and 30% test size. Training and testing have been applied using Kaggle. Three experiments were conducted on the CKD dataset and discussed.

3.1 Experiment 1

Simple imputation approach with mean strategy is used to handle with null values in the CKD dataset. Results are summarized in Table 2.

Table 2. Results of experiment 1

Model Name	Accuracy	Precision	F1-score	Recall
LR	0.73	0.742	0.729	0.730
SVM	0.627	0.607	0.605	0.627
DT	0.773	0.814	0.767	0.773
RF	0.812	0.827	0.813	0.812
NB	0.585	0.6	0.562	0.585
XGB	0.835	0.844	0.834	0.835
ADA	0.633	0.633	0.626	0.633

3.2 Experiment 2

People with similar physical conditions should have comparable physiological data, which is the rationale behind adopting the KNN-based technique to fill in the missing values. Particularly, for individuals in comparable circumstances, physiological measurement data variances should not be significant. The selection of K should neither be too large nor too small. The inconspicuous mode may not be taken into account if the K value is too high, which could be crucial. On the other hand, an excessively small K value results in noise, and the irregular data has a significant negative impact on filling in the missing values [25]. Hence, the K values implemented in this work were as 3, 5, 7, 9. Performance results are illustrated in table 3.

Table 3. Results of experiment 2

K Value	Model Name	Accuracy	Precision	F1-score	Recall
K = 3	LR	0.728	0.74	0.727	0.728
	SVM	0.631	0.612	0.609	0.631
	DT	0.773	0.814	0.767	0.773
	RF	0.81	0.828	0.811	0.81
	NB	0.593	0.607	0.57	0.593
	XGB	0.829	0.837	0.828	0.829
	ADA	0.536	0.567	0.506	0.536
K =5	LR	0.722	0.733	0.721	0.722
	SVM	0.631	0.611	0.608	0.631
	DT	0.771	0.813	0.765	0.771
	RF	0.812	0.830	0.813	0.812
	NB	0.594	0.609	0.572	0.594
	XGB	0.837	0.848	0.836	0.837
	ADA	0.536	0.567	0.506	0.536
K = 7	LR	0.722	0.733	0.721	0.722
	SVM	0.629	0.609	0.607	0.629
	DT	0.771	0.813	0.765	0.771
	RF	0.804	0.821	0.805	0.804
	NB	0.594	0.609	0.571	0.594
	XGB	0.829	0.84	0.827	0.829
	ADA	0.536	0.567	0.506	0.536

K=9	LR	0.724	0.736	0.723	0.724
	SVM	0.633	0.614	0.611	0.633
	DT	0.773	0.814	0.767	0.773
	RF	0.808	0.826	0.808	0.808
	NB	0.594	0.609	0.571	0.594
	XGB	0.839	0.852	0.838	0.839
	ADA	0.532	0.562	0.503	0.532

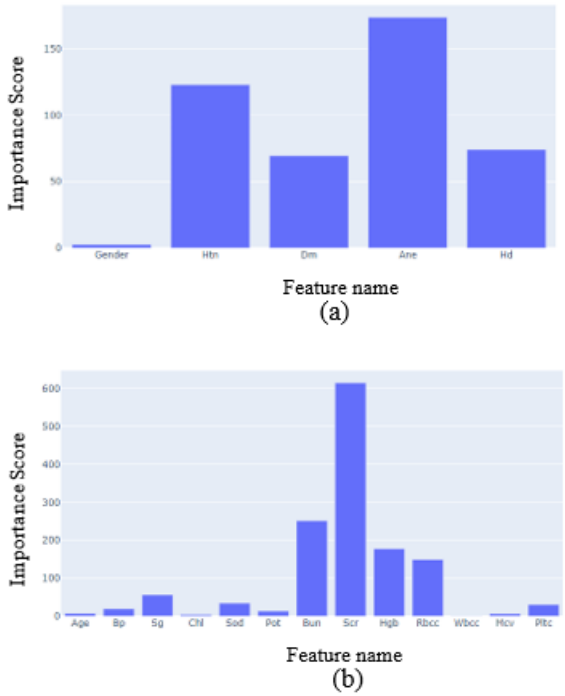
3.3 Experiment 3

Feature selection is a crucial component of ML so it is needed to select the most relevant features to build the model [26]. The Chi-square and ANOVA tests examine the connection between the characteristics. The feature importance score for both continuous and categorical variables is shown in Figure 4. Results due to different number of attributes selected for prediction of target class are illustrated in table 4.

Table 4. Results of experiment 3

Model Name	Number of features			Accuracy	Precision	F1-score	Recall
	Total	Numeric	Nominal				
LR	17	13	4	0.707	0.716	0.705	0.707
SVM				0.624	0.602	0.599	0.624
DT				0.707	0.711	0.706	0.707
RF				0.817	0.834	0.818	0.817
NB				0.579	0.591	0.556	0.579
XGB				0.817	0.832	0.817	0.817
ADA				0.633	0.633	0.626	0.633
LR	15	11	4	0.732	0.738	0.727	0.732
SVM				0.62	0.592	0.595	0.62
DT				0.707	0.716	0.708	0.707
RF				0.817	0.833	0.817	0.817
NB				0.602	0.626	0.577	0.602
XGB				0.8	0.809	0.8	0.8
ADA				0.662	0.655	0.655	0.662
LR	12	8	4	0.732	0.742	0.73	0.732
SVM				0.604	0.578	0.575	0.604
DT				0.705	0.713	0.707	0.705
RF				0.781	0.799	0.78	0.781
NB				0.624	0.649	0.603	0.624
XGB				0.755	0.766	0.755	0.755
ADA				0.6	0.593	0.589	0.6
LR	8	6	2	0.73	0.739	0.726	0.73
SVM				0.602	0.58	0.573	0.602
DT				0.672	0.679	0.673	0.672
RF				0.757	0.772	0.756	0.757
NB				0.631	0.654	0.607	0.631
XGB				0.732	0.74	0.731	0.732
ADA				0.575	0.597	0.557	0.575
LR	8	4	4	0.73	0.748	0.72	0.73
SVM				0.581	0.521	0.524	0.581
DT				0.686	0.689	0.685	0.686
RF				0.748	0.759	0.745	0.748
NB				0.593	0.622	0.564	0.593
XGB				0.736	0.741	0.732	0.736
ADA				0.658	0.668	0.654	0.658

Figure 4. Feature importance scores for (a) Numeric features and (b) Nominal features



4. CONCLUSION

The severity prediction of chronic kidney disease with high accuracy is considered to be one of the challenging biomedical research topics nowadays. This research has resulted in the development of a ML-based pipeline to successfully identify chronic kidney disease using a dataset of 1718 sample instants with 18 features and a five-class target prediction. Consequently, our goal was met by utilizing and analyzing various ML algorithms such as Random Forest, Decision Tree, AdaBoost, XGBoost, Naive Bayes in addition to artificial neural network, then compared the performance of these algorithms.

The proposed model reached 83.53% and 84.45% for accuracy and precision respectively using XGBoost without feature selection technique. When KNN imputation was implemented, XGBoost achieved an accuracy of 83.9 % and precision of 85.27%. Random forest scored 81.08% in terms of accuracy when using feature selection method. Validation and testing were performed.

In future work, more advanced ML and DL algorithms will be applied on different datasets either statistical or medical images so that the efficiency and effectiveness of CKD prediction can be boosted at earlier stages.

Reference

[1] A. Subasi, E. Alickovic, and J. Kevric, “Diagnosis of Chronic Kidney Disease by Using Random Forest,” in CMBEBIH 2017, 2017, pp. 589–594.

[2] K. J. Jager, C. Kovesdy, R. Langham, M. Rosenberg, V. Jha, and C. Zoccali, “A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases,” Nephrology Dialysis Transplantation, vol. 34, no. 11. 2019. doi: 10.1093/ndt/gfz174.

[3] M. M. Hossain et al., “Mechanical Anisotropy Assessment in Kidney Cortex Using ARFI Peak Displacement: Preclinical Validation and Pilot in Vivo Clinical Results in Kidney Allografts,” IEEE Trans Ultrason Ferroelectr Freq Control, vol. 66, no. 3, 2019, doi: 10.1109/TUFFC.2018.2865203.

[4] L. Bravo, N. Nieves-Pimiento, and K. Gonzalez-Guerrero, “Prediction of University-Level Academic Performance through Machine Learning Mechanisms and Supervised Methods,” Ingeniería, vol. 28, p. e19514, Nov. 2022, doi: 10.14483/23448393.19514.

[5] N. Park et al., “Predicting acute kidney injury in cancer patients using heterogeneous and irregular data,” PLoS One, vol. 13, no. 7, 2018, doi: 10.1371/journal.pone.0199839.

[6] A. Charleonnan, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwannawach, and N. Ninchawee,

"Predictive analytics for chronic kidney disease using machine learning techniques," in 2016 Management and Innovation Technology International Conference, MITiCON 2016, 2017. doi: 10.1109/MITICON.2016.8025242.

[7] A. Salekin and J. Stankovic, "Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes," in Proceedings - 2016 IEEE International Conference on Healthcare Informatics, ICHI 2016, 2016. doi: 10.1109/ICHI.2016.36.

[8] J. Xiao et al., "Comparison and development of machine learning tools in the prediction of chronic kidney disease progression," J Transl Med, vol. 17, no. 1, 2019, doi: 10.1186/s12967-019-1860-0.

[9] S. Y. Yashfi et al., "Risk Prediction of Chronic Kidney Disease Using Machine Learning Algorithms," in 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020, 2020. doi: 10.1109/ICCCNT49239.2020.9225548.

[10] V. Kumar, "Evaluation of computationally intelligent techniques for breast cancer diagnosis," Neural Comput Appl, vol. 33, no. 8, 2021, doi: 10.1007/s00521-020-05204-y.

[11] D. A. Debal and T. M. Sitote, "Chronic kidney disease prediction using machine learning techniques," J Big Data, vol. 9, no. 1, p. 109, 2022, doi: 10.1186/s40537-022-00657-5.

[12] E. H. A. Rady and A. S. Anwar, "Prediction of kidney disease stages using data mining algorithms," Informatics in Medicine Unlocked, vol. 15. 2019. doi: 10.1016/j.imu.2019.100178.

[13] I. Ibrahim and A. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," Journal of Applied Science and Technology Trends, vol. 2, no. 01, 2021, doi: 10.38094/jastt20179.

[14] M. Saar-Tsechansky and F. Provost, "Handling missing values when applying classification models," Journal of Machine Learning Research, vol. 8, 2007.

[15] E. Spyromitros-Xioufis, G. Tsoumakas, W. Groves, and I. Vlahavas, "Multi-target regression via input space expansion: treating targets as inputs," Mach Learn, vol. 104, no. 1, 2016, doi: 10.1007/s10994-016-5546-z.

[16] C. W. Chen, Y. H. Tsai, F. R. Chang, and W. C. Lin, "Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results," Expert Syst, vol. 37, no. 5, 2020, doi: 10.1111/exsy.12553.

[17] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, "A comparison of random forest variable selection methods for classification prediction modeling," Expert Systems with Applications, vol. 134. 2019. doi: 10.1016/j.eswa.2019.05.028.

[18] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," Journal of Supercomputing, vol. 77, no. 5, 2021, doi: 10.1007/s11227-020-03481-x.

[19] M. Wang and H. Chen, "Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis," Applied Soft Computing Journal, vol. 88, 2020, doi: 10.1016/j.asoc.2019.105946.

[20] A. B. Møller, B. V. Iversen, A. Beucher, and M. H. Greve, "Prediction of soil drainage classes in Denmark by means of decision tree classification," Geoderma,

vol. 352, 2019, doi: 10.1016/j.geoderma.2017.10.015.

[21] S. Nusinovici et al., "Logistic regression was as good as machine learning for predicting major chronic diseases," J Clin Epidemiol, vol. 122, 2020, doi: 10.1016/j.jclinepi.2020.03.002.

[22] A. Ogunleye and Q. G. Wang, "XGBoost Model for Chronic Kidney Disease Diagnosis," IEEE/ACM Trans Comput Biol Bioinform, vol. 17, no. 6, 2020, doi: 10.1109/TCBB.2019.2911071.

[23] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, vol. 13-17-August-2016. doi: 10.1145/2939672.2939785.

[24] Q. Huang, Y. Chen, L. Liu, D. Tao, and X. Li, "On Combining Biclustering Mining and AdaBoost for Breast Tumor Classification," IEEE Trans Knowl Data Eng, vol. 32, no. 4, 2020, doi: 10.1109/TKDE.2019.2891622.

[25] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease," IEEE Access, vol. 8, 2020, doi: 10.1109/ACCESS.2019.2963053.

[26] M. Almasoud and T. E. Ward, "Detection of chronic kidney disease using machine learning algorithms with least number of predictors," International Journal of Advanced Computer Science and Applications, vol. 10, no. 8, 2019, doi: 10.14569/ijacsa.2019.0100813.