Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MACHINE LEARNING APPLIED TO THE EARLY DETECTION OF HEART ATTACKS IN PATIENTS WITH TERMINAL CHRONIC KIDNEY DISEASE

BRYAN WALTER MEJIA MANZANARES[1], DIEGO RICHARD RIVERA DEMANUEL[2], YASIEL PÉREZ VERA[3]

[1,2,3] Universidad Nacional de San Agustín de Arequipa, Perú

E-mail: [1]bmejia@unsa.edu.pe, [2]driverad@unsa.edu.pe, [3]yperezv@unsa.edu.pe.

ABSTRACT

Currently, worldwide one in ten adults suffers from chronic kidney disease. The presence of renal lesions shows this disease. Causes the death of at least 2.4 million people a year; it is therefore appropriate to study and develop solutions to reduce the possibility of death. This study aims to develop a predictive model to aid in detecting heart attacks in patients suffering from this disease. To achieve this, six algorithms, Random Forest, XgBoost, Adaboost, Decision Tree, Support Vector Machine, and Gradient Boosting, were applied to construct the model. Statistical comparison was then performed using F1-score, Accuracy, Precision, Area Under the Curve, Recall, MCC and Kappa metrics to detect the best model. Adaboost was obtained as the best algorithm for the construction of models of the same nature. As a result, a model was developed to help predict a heart attack in people with chronic renal failure. This model allows classifications or predictions of this forecast to be made with good results and helps to reduce the risk of death in patients due to its high percentage of effectiveness. It could also be a starting point for future models that treat the same disease.

Keywords: *Machine Learning, Predictive Models, Adaboost, Random Forest, SVM, Decision Tree, Chronic Kidney Disease, Heart Attack.*

1. INTRODUCTION

Currently, 850 million people are affected by different types of kidney disorders. It is estimated that one in ten adults in the world suffers from Chronic Kidney Disease (CKD); in Latin America, this pathology reaches a prevalence of 715 per million inhabitants [1]. Projections of the National Institute of Statistics and Informatics in Peru indicate 19197 patients in the terminal stage nationwide [2].

In the United States, stage 1 to 4 CKD affects 10% of adults [3]. In 2007, more than 45,000 people in Spain, i.e., around 1,000 per million population, were undergoing renal replacement therapy. This figure is expected to almost double in the next 10 years due to the progressive aging of the population [4].

CKD is defined by a decrease in glomerular filtration rate and/or kidney lesions that have been present for more than 3 months. The diagnosis of renal failure is declared at the time when the glomerular filtration rate (GFR) is less than 60 ml/min/1.73 m2. In the presence of renal failure, it is essential to study the complications and progression factors of CKD [3].

There are populations considered to be at high risk for developing CKD. Physicians should be particularly attentive to this risk when evaluating five groups of patients: 1) Patients with high blood pressure, 2) Diabetic patients, 3) Patients older than 60 years of age, 4) Patients with cardiovascular disease, and 5) Family members of patients on dialysis or who have received a kidney transplant [5]. Patients with CKD, regardless of their disease status, have an increased cardiovascular risk relative to the general population. The two main factors for measuring the severity of the disease are arterial hypertension (AHT), which appears early, and microalbuminuria. These factors measure the severity of a patient's susceptibility to a heart attack [3].

Early detection of CKD is based on the determination of creatinine, GFR estimation, and microalbuminuria in patients with risk factors for

developing CKD [3]. To achieve maximum knowledge, autonomy, and good adherence to treatment, multidisciplinary therapeutic education programs are offered to patients [3]. Early identification of cardiovascular impairment is a feasible and implementable study even with small volume data samples. The integration of artificial intelligence (AI) in medical diagnostics offers a multitude of advantages. Some advantages of AI integration can potentially transform healthcare delivery for patients and providers alike [6].

Accurate and timely diagnosis is the cornerstone of adequate health care. Traditional methods can be limited by human subjectivity and the large volume of available data. Artificial Intelligence provides new opportunities to increase the accuracy and effectiveness of diagnosis [6]. In response to this problem, we propose machine-learning techniques to predict these critical attacks. According to [7], using artificial intelligence techniques in heart attack prediction represents a transformative approach to medical care that offers early detection, preventive intervention, and efficient patient care. This will help to reduce the likelihood of heart attacks in patients with heart disease.

The main objective of this research is to obtain the best model for predicting heart attacks in patients with CKD. We also have a secondary objective: to obtain the algorithm with the best impact on the model and the most influential variables for the construction of future models that deal with problems similar to those detailed in this work. This article contains the following sections: An analysis of related work is initially available, followed by the materials and methods section. The results and discussion section are presented below. Finally, the following conclusions are presented.

2. RELATED WORKS

Next, related work on applying AI algorithms to detect diseases or medical complications early is presented. A focused search has been made for articles that address medical conditions in the same category and are to be solved like that proposed in this work.

The paper presented by Halder has two objectives around CKD. The first objective is to improve the preprocessing of datasets for CKD classification. The second objective is to develop a web-based application for CKD prediction. The seven classifiers were used to predict the CKDs. The seven classifiers are AdaBoost (AdaB), Random Forest (RF), XgBoost (XgB), Support Vector Machine (SVM), Gradient Boosting (GB), Naive Bayes (NB), and Decision Tree (DT). The study resulted in a web application that enabled machine learning. The successful web application provides an accessible and affordable approach to CKD testing. The implications of the research are profound in the field of medical diagnostics. The study also supports healthcare accessibility. The study represents an important innovation in the medical field of CKD taking into account that the costs of treatment are very high [8].

The purpose of the research presented by Hamatani was to develop and verify a prediction model for machine learning (ML). The developed model will anticipate hospitalization for heart failure (HF) in people who suffer from atrial fibrillation (AF). Supervised ML was used. The classifier algorithms of the model used in the study are as follows: Elastic Net, RF, Neural Network, Light Gradient Boosting Machine, SVM, and NB Model. The result of the research was to achieve a machine learning model that could stratify the risk of hospitalization for HF in people who have AF. The model achieved provides opportunities to implement strategies to prevent HF among AF patients. The research concluded that the algorithms had comparable high predictive performance. ML models had high sensitivity, specificity, and accuracy. The research gave the ten most important variables for the model, with pre-existing heart failure being the most important of all [9].

Thorsen-Meyer's research aimed to determine whether ML methods using time series data analysis improve mortality prognosis for intensive care unit (ICU) patients. To approach the prognosis, it is important to provide real-time predictions of 90-day mortality. A recurrent neural network with a temporal resolution of one hour was trained. A model was trained with longitudinal data from patients admitted to the ICU. The model was internally validated using the retention method with 20% of the training data set. The results of the investigation show that the prediction of 90-day mortality was improved. The research concluded that the results need to be confirmed so that this model can be used as a headline tool. The results need to be confirmed in

a randomized clinical trial. The statistical model must be built on a causal model to achieve viable models. The author concludes that the notion of the new model has yet to converge with powerful machine-learning methods [10].

The study presented by Liu aimed to implement an electrocardiogram (ECG) model using AI to detect patients with elevated pulmonary arterial pressure (ePAP). The model also aims to identify prognostic implications related to ePAP. Research developed an AI model to detect ePAP. The model evaluated independent data sets from a hospital ECG database in 41 097 patients. Neural network deep learning tests of the model with 10-fold cross-validation resulted in an area under the curve (AUC) of 0.88 for detecting ePAP. The performance metrics of the model was successful across all ages, sexes, among other characteristics. The model was successful in identifying patients with ePAP. It also predicts patients' future risk of cardiovascular mortality. This model can be very useful to identify patients with ePAP. It is possible to initiate pretreatment to improve the treatment of patients with ePAP [11].

Ferguson's proposed research aimed to develop and validate an RF model to predict the progression of CKD using demographic and laboratory data. As part of the methodology applied, 77,196 individuals with a GFR at risk were considered, in addition to 80 laboratory features, evaluating model discrimination using the AUC and calibrating observational and predicted risks. The model achieved an AUC of 0.88 at two years and 0.84 at five years in internal tests. Within the model, 30% of high-risk individuals account for 87% of the two-year CKD progression events and 77% of the five-year progression events. The research mentions that a machine learning model that leverages routinely collected laboratory data can accurately predict a decline in renal failure [12].

In Ghosh and Khandoker's research, machine learning models were performed to predict CKD. This research proposes using a strategy based on explainable artificial intelligence (XAI), which takes advantage of clinical features. For the predictive model, five methods were considered: XgB, Naive Bayes, Random Forest, Decision Tree (DT) and Logistic Regression. The selection of the algorithms was based on accuracy and AUC. The research used two complementary algorithms, one is SHAP that deals with additive Shapley explanations and the other is LIME that deals with interpretable local explanations of the model to demonstrate the influence of the optimal characteristics of the model. The XgB model performed best, revealing essential variables that significantly impact prediction. The most significant variables of the study are creatinine, hemoglobin, and age. SHAP and LIME algorithms helped to interpret the machine learning models and help healthcare professionals understand the logic of the predicted results [13].

In this paper presented by Kanda, Machine Learning models were proposed to establish efficient detection and evaluation strategies. The application was performed in patients with early-stage type 2 diabetes mellitus (DM2). The study purported to develop a new model to predict the risk of developing CKD or heart failure (HF). The models were derived from a retrospective cohort of 217,054 patients with DM2 with no history of cardiovascular or renal disease. The best performing algorithm was XgBoost. In external validation, the 5-year prediction area under the curve of receptor operating features for diagnosis and hospitalization was 0.718 and 0.83, respectively. The constructed model predicted the risk of developing CKD/ICD in patients with type 2 diabetes mellitus with reasonable probability in the external validation. This clinical approach can help to promote better diagnosis and rapid intervention [14].

In the study by Wang and Chakraborty, a regression model was developed to predict the creatinine value based on twenty-three characteristics. Then, the predicted value is combined with the original twenty-three characteristics to assess the risk of CKD. The subsampling method was used, and a new mean square error (MSE) loss function was proposed. The RF, XgB, and a ResNet neural network-based model were used. R-squared value (R2) results were optimized to select the appropriate subsampling strategy. The regression model was used to predict creatinine. An R2 yield of 0.5590 was obtained. Six values directly affecting creatinine were highlighted: Genre, Age, Hemoglobin, Urine Protein Level, Waist Circumference, and Smoking Habits. They were using the predicted value for creatinine, an AUC of 0.76. The proposed system can assess the risk of CKD, identify the high-risk population, and

recommend screening for the disease through a creatinine test. It is possible to reduce the impact of CKD and facilitate early detection through creatinine testing [15].

AI algorithms have been used in various ways to detect diseases. They are a tool for ensuring better treatment for different types of patients.

3. MATERIALS AND METHODS

This section presents the related concepts, the algorithms used, the origin of the dataset used, and its most important characteristics and columns. It also presents the tools used for data analysis. The section also contains the proposal's overall process, the dataset's treatment, and the creation and training of the model developed. A statistical comparison of the developed models was made.

3.1 Theoretical Foundations

The following are the key theoretical concepts to understand the general proposal's development fully. The concepts mainly comprise the tools for the proposal and the reason for their selection.

Artificial intelligence is a broad term that encompasses the analysis of the potential of modern computers to perform tasks that would usually require the involvement of human intelligence. The development of projects based on deep learning, neural networks, machine learning or natural language processing are among the most significant progress being made with Artificial Intelligence [16].

Without the need to use explicit programming, machine learning allows systems to obtain updates automatically from a set of data. Algorithms are used mainly to analyze data to generate patterns that are then used in decision-making and the ability to make predictions [16].

Random Forest (RF) is a popular tree-based ensemble ML tool. It is highly adaptable to data and can account for the correlation and interactions between data features. This makes RFs particularly attractive for analyzing high-dimensional genomic data [17].

XGBoost is a robust machine-learning algorithm used primarily for classification and regression tasks. This method is based on decision trees, and the optimized implementation of the Gradient boosting algorithm makes it outstanding. The algorithm uses a technique that builds predictive models by combining multiple decision trees. This improves accuracy and reduces the risk of over-adjustment, which helps the training process. This algorithm is highly focused on achieving two characteristics: Computational speed and Model performance [18].

AdaBoost (AdaB) is a machine-learning technique that combines multiple weak classifiers to form a solid and accurate classifier. The new robust classifier is achieved by assigning weights to the training data and adjusting these weights according to the errors made by the classifiers in previous iterations. Later classifiers often focus more on the problematic examples. AdaB is common in many applications, such as face detection systems, feature extraction, and intrusion detection systems. The reason for this is its ability to improve the accuracy of the base models [19].

In the Machine Learning field, the use of Decision Trees (DT) are widely used in methods that develop supervised learning. DT employs a tree-like structure, which, depending on the input characteristics, helps you to make decisions and also to predict results. To create a tree structure the algorithm recursively divides the data according to the selected features, these features are represented by the internal nodes and the possible results or final values are represented by the edges. DT selects the best combination of features and values to split the data at each node [20].

The solid Support Vector Machine (SVM) classification algorithm can be used for both linear and nonlinear classification. The goal of using SVM is to find the hyperplane that best differentiates the classes in the attribute space. To control the trade-off between maximizing the margin and minimizing the classification error SVM uses a cost parameter [20].

Gradient Boosting (GB) is a machine-learning technique that builds prediction models sequentially for regression and classification problems. Each new model focuses on correcting the errors of the previous model by using a gradient descent process to adjust the loss

function. Although the method can be prone to overfitting, this is controlled by regularization techniques. The most common regularization technique is to limit the number of trees and adjust their depth. GB effectively handles complex problems and adapts to different types of data and loss functions. In addition, it allows us to assess the importance of the characteristics, which helps us understand which variables have the most significant impact on the predictions [21].

Python is focused on ease and readability and is well known for having a simple learning curve, which is why these characteristics make it a high-level interpreted programming language. In addition to the benefits of the language itself, you can take advantage of the power of other programming languages at the system level if you need to, the Python community has many tools and libraries available that make the language particularly nice and friendly for programmers dedicated to data science, scientific computing, and machine learning workloads. The latest libraries for machine learning and deep learning have their roots in Python. This language has experienced notable popularity that has increased with research and projects in the field of scientific computing in the last ten years [22].

PyCaret 3.0 is a set of Python-based libraries used mostly in low-code and open-source machine learning developments. It is very useful because it automates workflows and speeds up the development of ML models. It is a complete tool for model management and machine learning that increases efficiency and exponentially accelerates the experimental cycle [23].

Colab is a service hosted by Jupyter Notebook that allows programming and running Python in the browser. Colab requires no configuration and access to GPUs at no additional cost. This tool is suitable for machine learning, data science, and education. Successfully combines executable code and rich text, images, html. LaTex and much more in a single document. Notebooks created in Colab are stored in a Google Drive account for sharing content with other users [24].

3.2 DataSet

The dataset was obtained from Kaggle [25]. These data were obtained from patients who were admitted for two years at Hero DMC Heart Institute, Dayanand Medical College and Hospital Unit, Ludhiana, Punjab, India. The cardiology unit had 14,845 admissions during the study period, corresponding to 12,238 patients and 1921 patients with multiple admissions.

The dataset comprises 56 columns with different types of patient medical data and corresponds to 15,000 records in total [26]. The data were linked to the date of admission of the patients; the date of hospital discharge; demographic information, such as rural or urban location, gender, age, and the type of outpatient or emergency admission. In addition, the patient's medical history and consumption, including alcohol and tobacco use, were also taken into account. Diseases such as hypertension (HTN), diabetes mellitus (DM), cardiomyopathy (CMP) and chronic kidney disease (CKD), coronary artery disease (CAD) were taken into consideration; and laboratory parameters corresponding to hemoglobin (HB), ejection fraction (EF), glucose, creatinine, urea, brain natriuretic peptide (BNP), total lymphocyte count (TLC), elevated cardiac enzymes (RCE) and platelets [26].

BNP are proteins produced by the heart and blood vessels. BNP tests measure the amount of these proteins in a blood sample, and they are used to help detect heart failure.

The left ventricle pumps blood with each contraction this is known as the ejection fraction (EF) . A sixty percent ejection fraction refers to the sixty percent of blood that the left ventricle of the heart pumps with each beat out of the total amount of blood.

A total of 28 characteristics were recorded and analyzed, including pulmonary embolism, heart failure and acute segment elevation myocardial infarction (STEMI). It also has features such as shock. It has been considered the origin of shock in patients when it was any cause other than myocardial infarction, this shock has as its main characteristic a systolic blood pressure lower than 90 mmHg. Patients who suffered shock due to cardiac disease were assigned to a category called cardiogenic shock and patients who suffered shock due to multifactorial pathophysiology were included in both categories. Noting whether the patient was hospitalized or died, findings were documented

within the dataset.

The age range of the dataset starts with patients from four years old to seventy years old. The age column has a higher percentage of patients from forty to sixty-five. The dataset also contains fifty percent of patients suffering from smoking and alcohol consumption. The CKD spine has a high rate of patients of about seventy percent. The Hypertension column indicates that eighty percent of patients present this item. The column on urea and creatinine, which are very important for the study, presents complete data for ninety-four percent of the total. Figure 1 shows the percentage of patients with the abovementioned characteristics concerning the total number of records in the dataset.

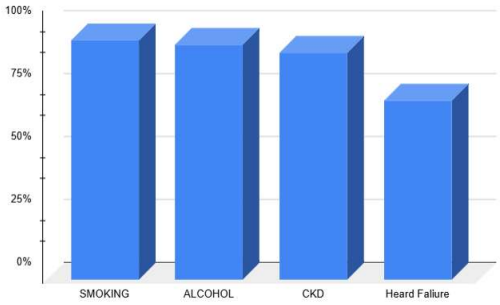

Figure 1:Percentage representation of patients' medical conditions in the dataset

3.3. Proposed model for early detection of heart attacks in patients with CKD

Figure 2 shows the general process of this study. The study consists of the construction of an efficient model with good results in the different metrics that have been selected for application.

Initially, all the variables in the dataset are used for processing, training, and prediction using the various machine learning algorithms. Subsequently, the most influential variables of the dataset are obtained. These variables will be detected and evaluated to determine their level of importance.

Then, the different metrics are applied within the model, giving us the results and qualities of each model built with the different algorithms. A statistical comparison is carried out, which will indicate which model has the best results. It concludes with the analysis of the statistical evaluation results and a response to the initial objectives of this research.

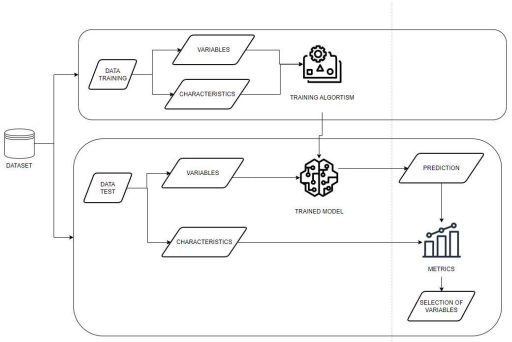

Figure 2: Outline of the general proposal

Dataset processing: The original dataset contains 15,757 records, and several columns were found with missing values. The information was processed using imputation techniques using the mean for the numerical variables and the mode and median for the categorical variables using the PyCaret library. These techniques ensure data consistency for subsequent model training.

In the case of the BNP column, it was decided not to process the missing values because they are considered sensitive data and represent 42.5% of the entire column's data. In addition, the outliers found in the dataset were identified and eliminated. The 'remove_outliers=True' setting was used with a threshold of 2%, and outliers were removed, which could be considered irrelevant information.

Correct data processing helps to optimize the dataset, resulting in cleaner data for further analysis and training. Some fields are filtered to focus only on records directly related to CKD patients.

After this filtering, 1550 records were obtained. These data were used to apply them in the ML models mentioned above. This information ensures that the CKD patient population will be analyzed exclusively.

Model Creation: The programming process was simplified when the model was created with the help of PyCaret and a model for executing the algorithms could be generated. This study is taking 95% of the dataset for training because the model must learn patterns. 5% of the dataset is being taken for tests. In [27], it is mentioned that this division is feasible because of the number of

records we have, obtaining good results.

Algorithm Execution: The selected algorithms were executed, and data on each algorithm was obtained for comparison. The same metrics were obtained for subsequent comparison in all the algorithms executed. These metrics are Accuracy, Area Under the Curve (AUC), True Positive Ratio (Recall), Precision, High Precision and High Sensitivity (F1-score), Model Accuracy Indicator (Kappa), and Matthews Correlation Coefficient (MCC).

Ten folds were obtained for each algorithm from each metric. The most influential parameters for obtaining these folds are detailed below.

In addition to the parameters mentioned above in the table, some important parameters were used in each of the algorithms consistently, which are listed below:

- n_jobs=-1: Number of parallel jobs, in this case, using all cores.
- n_estimators=None: Number of trees in the model, in this case, the maximum.
- max_depth=None: Maximum tree depth until all leaves are pure.
- learning_rate=None: Controls the contribution of each tree to the final model.
- random_state=123: Seed for the random number generator. Setting this value ensures that the results are reproducible.

Table 1 details some of the parameters used in each proposed algorithm, showing the most influential and unique parameters.

Table 1: Parameters by algorithm

Algorithm	Parameters
Random Forest	bootstrap=True, criterion='gini', max_features='sqrt', min_samples_split=2
XgBoost	booster='gbtree', colsample_bytree=None, early_stopping_rounds=None, gamma=None, min_child_weight=None, objective='binary:logistic'
AdaBoost	learning_rate=1.0, n_estimators=50, algorithm='SAMME.R', estimator=None,
Decision Tree	min_impurity_decrease =0.0, max_features =sqrt, class_weight=None, criterion='gini', max_leaf_nodes =None,, ccp_alpha=0.0, splitter='best'.
Support Vector Machine	eta0=0.001, alpha=0.0001, epsilon=0.1, fit_intercept=True, class_weight=None
Gradient Boosting	loss='log_loss', subsample=1.0, max_depth=3

Selection of best variables: Of the fifty-six variables, the ten most influential were selected for the model. The ten most influential variable is Heart failure with reduced ejection fraction (HFREF), which is given to see when the heart muscle does not contract effectively and less oxygen-rich blood is pumped to the body. Heart failure with average ejection fraction (HFNEF) indicates how efficiently the heart pumps blood. B-type natriuretic peptide (BNP) signals the blood vessels to open wider and the kidneys to eliminate water and salt through urine. This process helps reduce strain on the heart by lowering blood pressure and reducing the blood the heart has to pump. Ejection fraction (EF) is the amount of blood pumped out of a full ventricle, with each heartbeat expressed as a percentage. Creatinine (CREATININE) is filtered out of the blood by the kidneys and eliminated from the body in the urine; high creatinine levels indicate a kidney problem. Urea (UREA) is an essential indicator of kidney function and protein balance in the body. Glucose (GLUCOSE) or blood sugar is the key to keeping the body's mechanisms functioning optimally. Platelets (PLATELETS), age (AGE), and total leukocyte count (TLC).

The results obtained by the RF algorithm were considered for choosing the most influential variables in predicting heart attacks. Figure 3 shows the variables that are of direct importance to CKD disease.

Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

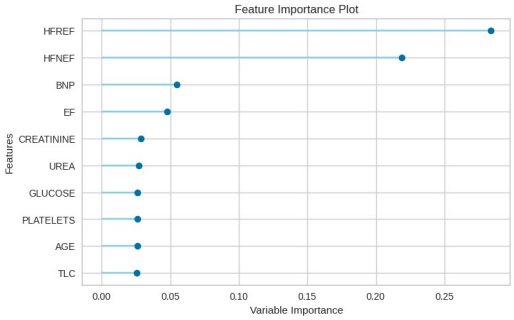

Figure 3: Most Influential Variables in the model according to Random Forest

4. RESULTS AND DISCUSSION

The results obtained from the procedure detailed above are presented below. The results of the seven metrics explained compared to each of the developed algorithms are shown. Finally, a statistical valuation is placed to obtain the best algorithm developed.

Accuracy: Figure 4 shows a comparative table on the Accuracy metric. Accuracy is the proportion of correct predictions a model makes concerning the total number of predictions. As shown in Figure 4, the AdaBoost algorithm performs better in this metric's results. The SVM algorithm shows the lowest performance of all the algorithms.

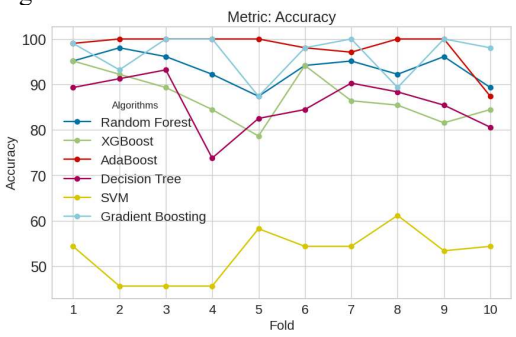

Figure 4: Comparison of the accuracy metric for the six algorithms.

AUC: A comparison table on the AUC metric is shown below in Figure 5. AUC represents the probability that given a randomly chosen positive and negative example, the model will classify the positive as more significant than the negative. Generally, the higher the AUC score, the better the performance of a binary classifier for a given classification task. As can be seen in Figure 4, the algorithm with the best performance for this metric is AdaBoost. The algorithm with the lowest performance is SVM.

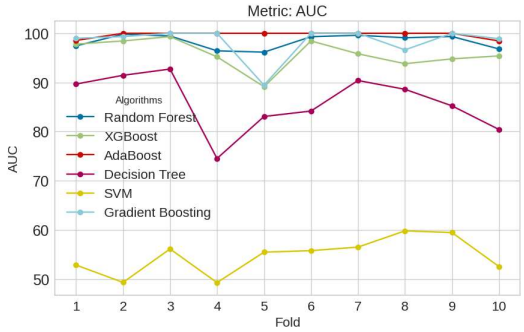

Figure 5: Comparison of the AUC metric for the six algorithms.

Recall**:** A comparative table on the Recall metric is shown below in Figure 6. Recall, or exhaustiveness, is a metric that measures the proportion of positive cases that a classification model correctly identifies about the total number of real positive cases. As can be seen in Figure 5, the algorithm with the best performance for this metric is Gradient Boosting. The algorithm with the lowest performance is SVM. It can also be noted that SVM maintained high results in the first half of the folds, but in the second half, it declined.

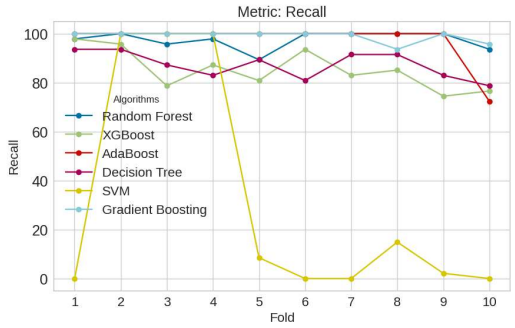

Figure 6: Comparison of the Recall metric for the six algorithms.

Precision: A comparison table on the precision metric is shown below in Figure 7. Accuracy is a simple and intuitive metric that indicates the percentage of correct predictions from the total number of predictions. Precision measures the reliability of the model when predicting a positive outcome. As can be seen in Figure 6, the best algorithm for this metric is AdaBoost. The algorithm with the lowest performance is SVM.

Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

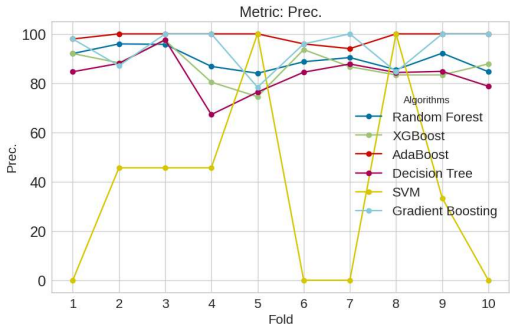

Figure 7: Comparison of the Precision metric for the six algorithms.

F1-score: A comparative table on the F1-score metric is shown below in Figure 8. The F1 score is a metric that for a single value merges the recall and precision of a classifier. It is determined as the balance of the harmonic mean of recovery and precision. The F1 Score considers both precision and recall to calculate a score that tends to be lower than either of the individual metrics when one is significantly lower than the other. This reflects that the F1 Score penalizes extreme imbalances between precision and recall, ensuring that both are reasonably high to obtain a high F1 Score. As can be seen in Figure 7, the algorithm that performs best for this metric is Gradient Boosting. The algorithm with the lowest performance is SVM.

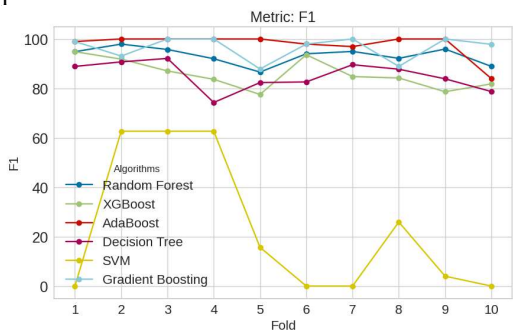

Figure 8: Comparison of the F1 metric for the six algorithms.

Kappa: The graph of the Kappa metric in Figure 9 shows the performance of the models in this metric, with the Gradient boosting model and AdaBoost reaching the maximum in folds 3, 4, and 9. On the other hand, the SVM model obtained the minimum in fold 9. It helps evaluate the precision and consistency of classification models in machine learning tasks.

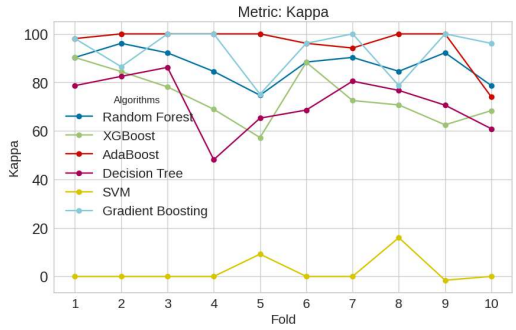

Figure 9: Comparison of the Kappa metric for the six algorithms.

MCC: The MCC metric plot in Figure 10 shows the performance of the models in this metric, with the Gradient boosting and AdaBoost model reaching the maximum at folds 3, 4, and 9. The SVM model obtained the minimum at fold 9. MCC is used to measure quality when talking about binary classification. This metric considers true and false positives and negatives, providing a balanced assessment of model performance, even with unbalanced data. It is regarded as one of the best metrics when there is a significant difference between data categories.

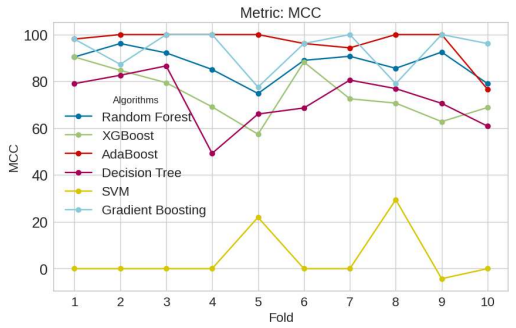

Figure 10: Comparison of the MCC metric for the six algorithms.

Shapiro-Wilk Normality Tests and Levene Homogeneity of Variance Tests: The Shapiro-Wilk statistical test was performed to evaluate the results of each metric and confirm the normal distribution of the data. This technique is efficient in small data sets; these records are, at most, 2000 observations, and its performance is more optimal in small sample sizes. Next, Levene's test was applied to ensure that the data variances remained constant. In most cases, the data passed both evaluations. Finally, the ANOVA test for repeated samples was performed, and at least two averages showed significant differences in the results of the seven metrics.

It is essential to identify where these significant differences are found. For this reason, the Student's t-test for related samples was applied, with a 99% confidence interval. Peer comparisons were performed, and those that did not show relevant differences were classified into groups. Finally, groups ranging from the most outstanding results (Group 1) to the worst were established, as shown in Table 3.

As Table 2 shows, the algorithm that has remained in group 1 (G1) in all metrics is AdaBoost. AdaBoost has the best results throughout our study, followed by Gradient Boosting and Random Forest.

Adaboost is the best algorithm, and it matches the explanations of [28] in their study on selecting improved features and ensemble learning for cardiovascular disease prediction. It sets Adaboost as one of the most promising algorithms for improving predictive modeling. Their results show that AdaBoost removed impurities from the model and boosted its efficiency. It also proves to be an effective algorithm for predicting heart disease. The use of AdaBoost in various industries is described in [29]. In the medical field, this algorithm has evidenced significant potential for improving classification performance; it is adequate for multiple medical uses. In addition, it indicates that the application of Adaboost performs a crucial role in the early detection of diseases, improves patients' medical examination results, and decreases the costs associated with medical care.

Table 2: Groupings and means of each metric by the corresponding algorithm.

Metric / Algorithm		AdaB	GB	RF	XgB	DT	SVM
Accuracy	Group	G1	G1	G1	G2	G2	G3
	Mean	98.156	96.505	93.59	87.185	85.922	52.719
AUC	Group	G1	G2	G2	G3	G4	G5
	Mean	99.698	98.327	98.34	95.816	86.028	54.713
Recall	Group	G1	G1	G1	G2	G2	G3
	Mean	97.234	98.936	97.44	85.319	87.234	32.553
Precision	Group	G1	G1	G1	G1	G1	G2
	Mean	98.784	94.383	89.57	86.726	83.385	37.022
F1-score	Group	G1	G1	G1	G2	G2	G3
	Mean	97.773	96.451	93.30	85.805	85.092	23.363
Kappa	Group	G1	G1	G1	G2	G2	G3
	Mean	96.229	93.047	87.19	74.136	71.796	2.36
MCC	Group	G1	G1	G1	G2	G2	G3
	Mean	96.518	93.425	87.55	74.428	72.116	4.714

5. CONCLUSIONS

It is concluded that the study's objective was met since a model for predicting heart attacks in patients with CKD was developed. This model has been developed using six different algorithms and is functional for new datasets that can be predicted. An algorithm with better results was obtained when building the model. The algorithm with the most impact on the study was Adaboost, with a confidence of 99.69%. This algorithm builds a model that exceeds all the metrics with respect to other algorithms. In addition, it is concluded that the Support Vector Machine algorithm builds a model with the lowest results, so its application in models similar to the one used in this study is not recommended.

This study is important because it is used in the health field, such as the prognosis of heart attacks in people with CKD. The proposed model allows classifications or predictions of this prognosis to be made with good results.

REFERENCES

[1] M. Wainstein et al., "International Society of Nephrology Global Kidney Health Atlas: structures, organization, and services for the management of kidney failure in Latin America," Kidney Int. Suppl., vol. 11, no. 2, pp. e35–e46, May 2021, doi: 10.1016/j.kisu.2021.01.005.

[2] E. R. Castillo-Velarde and A. H. Caro, "Satisfacción del tipo de acceso vascular y calidad de vida en pacientes de una unidad de hemodiálisis en Perú," Rev. Colomb. Nefrol., vol. 11, no. 1, Art. no. 1, Apr. 2024, doi: 10.22265/acnef.11.1.703.

[3] J.-J. Boffa and C. Cartery, "Insuficiencia renal crónica o enfermedad renal crónica," EMC - Tratado Med., vol. 19, no. 3, pp. 1–8, Sep. 2015, doi: 10.1016/S1636-5410(15)72803-5.

[4] Á. L. Martín De Francisco, L. Aguilera García, and V. Fuster Carulla, "Enfermedad cardiovascular, enfermedad renal y otras enfermedades crónicas. Es necesaria una intervención más temprana en la enfermedad renal crónica," Aten. Primaria, vol. 41, no. 9, pp. 511–514, Sep. 2009, doi: 10.1016/j.aprim.2008.12.007.

[5] H. Juan Carlos Flores, "Enfermedad renal crónica: epidemiología y factores de riesgo," Rev. Médica Clínica Las Condes, vol. 21, no. 4, pp. 502–507, Jul. 2010, doi: 10.1016/S0716-8640(10)70565-4.

[6] L. J. S. Madriz, D. C. S. Benavides, J. F. S. Pérez, L. D. P. González, and N. P. C. Arias, "Inteligencia Artificial aplicada al diagnóstico médico: una revisión actual," Rev. Científica Salud Desarro. Hum., vol. 5, no. 2, Art. no. 2, Jul. 2024, doi: 10.61368/r.s.d.h.v5i2.183.

[7] I. Rojek, P. Kotlarz, M. Kozielski, M. Jagodzinski, and Z. Krolikowski, "Development of AI-Based Prediction of Heart Attack Risk as an Element of Preventive Medicine," ELECTRONICS, vol. 13, no. 2, p. 272, Jan. 2024, doi: 10.3390/electronics13020272.

[8] R. K. Halder et al., "ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application," J. Pathol. Inform., vol. 15, p. 100371, Dec. 2024, doi: 10.1016/j.jpi.2024.100371.

[9] Y. Hamatani et al., "Machine Learning Risk Prediction for Incident Heart Failure in Patients With Atrial Fibrillation," JACC Asia, vol. 2, no. 6, pp. 706–716, Nov. 2022, doi: 10.1016/j.jacasi.2022.07.007.

[10] H.-C. Thorsen-Meyer et al., "Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records," Lancet Digit. Health, vol. 2, no. 4, pp. e179–e191, Apr. 2020, doi: 10.1016/S2589-7500(20)30018-2.

[11] C.-M. Liu et al., "Artificial Intelligence-Enabled Electrocardiogram Improves the Diagnosis and Prediction of Mortality in Patients With Pulmonary Hypertension," JACC Asia, vol. 2, no. 3, pp. 258–270, Jun. 2022, doi: 10.1016/j.jacasi.2022.02.008.

[12] T. Ferguson et al., "Development and External Validation of a Machine Learning Model for Progression of CKD," KIDNEY INTERNATIONAL REPORTS, vol. 7, no. 8. ELSEVIER SCIENCE INC, STE 800, 230 PARK AVE, NEW YORK, NY 10169 USA, pp. 1772–1781, Aug. 2022. doi: 10.1016/j.ekir.2022.05.004.

[13] S. K. Ghosh and A. H. Khandoker, "Investigation on explainable machine learning models to predict chronic kidney diseases," SCIENTIFIC REPORTS, vol. 14, no. 1. NATURE PORTFOLIO, HEIDELBERGER PLATZ 3, BERLIN, 14197, GERMANY, Feb. 14, 2024. doi: 10.1038/s41598-024-54375-4.

[14] E. Kanda et al., "Machine learning models for prediction of HF and CKD development in early-stage type 2 diabetes patients," SCIENTIFIC REPORTS, vol. 12, no. 1. NATURE PORTFOLIO, HEIDELBERGER PLATZ 3, BERLIN, 14197, GERMANY, Nov. 21, 2022. doi: 10.1038/s41598-022-24562-2.

[15] W. Wang, G. Chakraborty, and B. Chakraborty, "Predicting the Risk of Chronic Kidney Disease (CKD) Using Machine Learning Algorithm," APPLIED SCIENCES-BASEL, vol. 11, no. 1. MDPI, ST ALBAN-ANLAGE 66, CH-4052 BASEL, SWITZERLAND, Jan. 2021. doi: 10.3390/app11010202.

[16] Y. González-Pérez, A. Montero Delgado, and J. M. Martinez Sesmero, "Acercando la inteligencia artificial a los servicios de farmacia hospitalaria," Farm. Hosp., vol. 48, pp. S35–S44, Jul. 2024, doi: 10.1016/j.farma.2024.02.007.

[17] X. Chen and H. Ishwaran, "Random forests for genomic data analysis," Genomics, vol. 99, no. 6, pp. 323–329, Jun. 2012, doi: 10.1016/j.ygeno.2012.04.003.

[18] S. Frutos Serrano, "Comparación entre XGBoost y Regresión Lineal Múltiple para la predicción de la evolución del precio de las acciones," Jun. 2022, Accessed: Aug. 31, 2024. [Online]. Available: https://hdl.handle.net/20.500.14352/3266

[19] Y. Ding, H. Zhu, R. Chen, and R. Li, "An Efficient AdaBoost Algorithm with the Multiple Thresholds Classification," Appl. Sci., vol. 12, no. 12, Art. no. 12, Jan. 2022, doi: 10.3390/app12125872.

[20] A. Arifuddin, G. S. Buana, R. A. Vinarti, and A. Djunaidy, "Performance Comparison of Decision Tree and Support Vector Machine

Algorithms for Heart Failure Prediction," Procedia Comput. Sci., vol. 234, pp. 628–636, 2024, doi: 10.1016/j.procs.2024.03.048.
[21] B. Fernandes, A. González-Briones, P. Novais, M. Calafate, C. Analide, and J. Neves, "An Adjective Selection Personality Assessment Method Using Gradient Boosting Machine Learning," Processes, vol. 8, no. 5, Art. no. 5, May 2020, doi: 10.3390/pr8050618.
[22] S. Raschka, J. Patterson, and C. Nolet, "Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence," Information, vol. 11, no. 4, Art. no. 4, Apr. 2020, doi: 10.3390/info11040193.
[23] "PyCaret 3.0 | Docs." Accessed: Aug. 28, 2024. [Online]. Available: https://pycaret.gitbook.io/docs/
[24] "Google Colab." Accessed: Sep. 03, 2024. [Online]. Available: https://colab.research.google.com/
[25] A. Sahani and G. Wander, "Hospital Admissions Data." [Data Set]. Available: https://www.kaggle.com/datasets/ashishsahani/hospital-admissions-data/data?select=HDHI+Admission+data.csv
[26] S. C. Bollepalli et al., "An Optimized Machine Learning Model Accurately Predicts In-Hospital Outcomes at Admission to a Cardiac Unit," Diagnostics, vol. 12, no. 2, p. 241, Jan. 2022, doi: 10.3390/diagnostics12020241.
[27] R. Boiarsky, N. Singh, A. Buendia, G. Getz, and D. Sontag, "A Deep Dive into Single-Cell RNA Sequencing Foundation Models," Oct. 23, 2023, bioRxiv. doi: 10.1101/2023.10.19.563100.
[28] S. P. Praveen et al., "Enhanced feature selection and ensemble learning for cardiovascular disease prediction: hybrid GOL2-2 T and adaptive boosted decision fusion with babysitting refinement," Front. Med., vol. 11, p. 1407376, Jul. 2024, doi: 10.3389/fmed.2024.1407376.
[29] T. Miller et al., "Boosting modern society: advancements and applications of the adaboost algorithm in diverse domains," Sci. Collect. «InterConf», no. 152, Art. no. 152, Apr. 2023.