IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 June 2022, accepted 4 July 2022, date of publication 7 July 2022, date of current version 13 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3189018

RESEARCH ARTICLE

Machine-Learning Model to Predict the Intradialytic Hypotension Based on Clinical-Analytical Data

LUIS MENDOZA-PITTÍ[1,2], JOSÉ MANUEL GÓMEZ-PULIDO[1], MIGUEL VARGAS-LOMBARDO[2], (Senior Member, IEEE), JUAN A. GÓMEZ-PULIDO[3], (Senior Member, IEEE), MARÍA-LUZ POLO-LUQUE[4], AND DIEGO RODRÍGUEZ-PUYOL[5]

[1]Health Computing and Intelligent Systems Research Group (HCIS), Department of Computer Science, Ramón y Cajal Institute for Health Research (IRYCIS), University of Alcalá, 28805 Alcalá de Henares, Spain
[2]SNI, Senacyt Panama, E-Health and Supercomputing Research Group (GISES), Technological University of Panama, Panama City 0819-07289, Panama
[3]Department of Technologies of Computers and Communications, University of Extremadura, 10003 Cáceres, Spain
[4]Department of Nursing and Physiotherapy, University of Alcalá, 28805 Alcalá de Henares, Spain
[5]Department of Medicine and Medical Specialties, Research Foundation of the University Hospital Príncipe de Asturias, IRYCIS, University of Alcalá, Alcalá de Henares, 28805 Madrid, Spain

Corresponding authors: Miguel Vargas-Lombardo (miguel.vargas@utp.ac.pa) and José Manuel Gómez-Pulido (jose.gomez@uah.es)

This work was supported in part by the clinical team of the Nephrology Section of the Hospital Universitario Príncipe de Asturias, in part by the National Secretariat for Science, Technology and Innovation (SENACYT, Panama) and the Institute for the Training and Use of Human Resources (IFARHU), for the doctoral study scholarship, in part by the SENACYT through the National Research System (SNI-SENACYT) National Research System of which one of the authors is a member, and in part by the Fresenius Medical Care AG company.

ABSTRACT Predicting whether patients will experience intradialytic hypotension (IDH) during hemodialysis (HD) is not an easy task. IDH is associated with multiple risk factors, meaning that traditional statistical models are unable to find the relationships that affect it. In this context, the use of models based on machine learning (ML) can allow the discovery of complex relationships, since they can solve problems without being explicitly programmed. In this work we developed, evaluated and identified an ML-based model that is capable of predicting at the beginning of the HD session whether a patient will suffer from IDH during its prolonged development. To develop the ML models, we used the hold-out and cross-validation methods; while, to evaluate the performance of the models we used the metrics F1-score, Matthews Correlation Coefficient, areas under the receiver operating characteristic (AUROC) and precision-recall curve (AUPRC). In this sense, we selected and used a reduced combination of variables from clinical records and blood analytics, which have proven to be decisive for the occurrence of IDH. The predictive results obtained through our work confirmed that the best ML model was based on the XGBoost model, achieving values of 0.969 and 0.945 for AUROC and AUPRC respectively. Therefore, our study suggests that the XGBoost model has a very high predictive capacity for the appearance of an IDH in HD patients and presents great versatility and flexibility in terms of supporting informed decision-making by medical staff.

INDEX TERMS Clinical-analytical data, hemodialysis, intradialytic hypotension, machine learning, predicting model, XGBoost.

I. INTRODUCTION

Chronic kidney disease (CKD) is the sixth fastest growing cause of death, affecting approximately 11–13% of the population [1]. It is estimated that 850 million people worldwide suffer from kidney disease, and it is responsible for at least 2.4 million deaths per year. Furthermore, the incidence of CKD is expected to increase due to aging of the population, since although it can develop at any age, its prevalence is higher among older adults. One in five men and one in four women aged 65–74 suffer from CKD, and half of people aged 75 or older have CKD to some degree. The prevalence

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

of CKD in Spain is approximately 15%, whereas in the United States it is approximately 13%, and it can therefore be observed that these rates are quite similar [2]. Furthermore, Spain has one of the highest rates of CKD compared to other European countries [3].

Dialysis or renal transplantation are treatments for advanced CKD, which occurs in approximately 1% of people with CKD. They are also one of the most expensive chronic disease treatments [2], [4]. In this case, patients may be treated by hemodialysis (HD).

HD is a treatment that replaces two of the kidney's main functions: periodic filtering of the blood and elimination of excess fluids from the body. To perform HD, artificial filters called dialyzers are used in which a large amount of blood is circulated through the filter via vascular access. However, fluid removal is one of the main causes of hypotension in patients [5].

Hypotension due to dialysis, known as intradialytic hypotension (IDH), is one of the most frequent complications in clinical practice. IDH has a prevalence ranging from 5–40% [6], [7]. This large range of values is because there is no clear consensus and uniformity in many investigations on the definition of the onset of IDH [8]–[10]. Nevertheless, several studies have used one or more of the following criteria to define IDH; (i) a threshold/nadir in systolic blood pressure (SBP); (ii) an absolute reduction in SBP; and (iii) a requirement for intervention [9], [10]. Furthermore, identifying the factors that cause the occurrence of IDH is a complex task, since multiple factors are involved, such as the modality of dialysis, patient type or medical criteria [7], [11]. Therefore, identification of the factors that most influence its occurrence would improve decision-making by clinical staff (doctors, nurses, and assistants, among others).

HD is usually carried out in sessions that last approximately four hours and are performed on alternate days. Large amounts of data are generated during these HD sessions, often consisting of different records (e.g., demographic data, laboratory records or medical prescriptions). In this sense, the use of specialized techniques such as big data and artificial intelligence (AI) are of great use in the analysis of variabilities and intrinsic relationships in patient registries with and without IDH.

Recent developments in the field of AI have stimulated its application in daily clinical practice [12], [13], because it is useful for processing massive and complex sets of big data and can facilitate diagnostic and therapeutic decisions, promote medical innovation, and reduce costs, among its other advantages. In view of this, AI-based models such as machine learning (ML) have received increasing interest in recent years, since they can significantly contribute to disease detection, early diagnosis, and prediction and/or automatic classification of diseases [1], [14]. This means that medical staff have more information for decision-making, which consequently improves medical care and patient outcomes.

Although the application of ML models is relatively mature in other specialties, this has not been the case in the field of nephrology, as the lack of evidence and the limited scope of research in kidney disease have not allowed this specialty to benefit from these technologies [15], [16]. Therefore, the development of intelligent ML-based models may support nephrology medical staff in the context of identifying the occurrence of IDH in patients receiving HD.

Some of the ML models that are most used in the different specialties of medicine are: Logistic Regression, Random Forest, Multi-layer Perceptron and Extreme Gradient Boosting (XGBoost),

For this reason, the present research focuses on developing, evaluating and identifying an intelligent ML model that allows predictions regarding whether a patient will have IDH during an HD session. In addition, this work aims to determine the most relevant parameters associated with the occurrence of IDH based on the massive analysis of data relating to an HD session, including the clinical parameters measured at the beginning of the session and the most recent parameters of the blood analytics available at the time session. Data were provided by nephrology staff of the Hospital Príncipe de Asturias in Madrid, Spain. Different techniques were applied to this dataset (explained Section III), in order to obtain a useful dataset for the development of ML models that are able to determine whether or not the patient will have IDH at the beginning of HD treatment.

In this research study, different ML models (Logistic Regression, Random Forest, Multi-layer Perceptron and eXtreme Gradient Boosting [XGBoost]) were developed and studied. These ML models are widely used in fields such as medicine, energy, agriculture, among others [17]–[19] and, in addition, they are good at performing classification tasks; therefore, we decided to evaluate their performance in our case study. The models were developed by applying hold-out and cross-validation techniques (explained Section III). The evaluation criteria used to compare these models were the F1-score, Matthews Correlation Coefficient (MCC), areas under the receiver operating characteristic (AUROC) and precision-recall curve (AUPRC) metrics, as these are very useful when evaluating the performance of dichotomous classification models [20], [21].

After evaluating the different models, the results showed that the model based on the XGBoost demonstrated the best performance, obtaining 0.86 and 0.81 for F1-score and MCC respectively. In addition, obtaining an area under the curve (AUC) for the receiver operating characteristic (ROC) of 0.969 and an AUC for precision-recall (PR) of 0.945. Consequently, the results of this study suggest that the XGBoost model is feasible for its application as a predictor of IDH in HD sessions, due to its versatility and flexibility in terms of supporting informed decision-making by medical staff. It also provides medical staff with insight into the identification of the most decisive clinical and analytical parameters that affect the occurrence of IDH during an HD session.

The article is organized as follows. In Section II, we review and describe some prior work related to the context of our study. In Section III, we briefly detail the methodological

procedure employed in the development of the ML models. In Section IV, we evaluate and present our results for several different metrics for each of the models developed. In Section V, we analyze and discuss the most important findings. Finally, in Section VI, we highlight our main conclusions and suggestions from the study and describe future research work

II. LITERATURE REVIEW

According to several studies [19], [22]–[25] ML-based models are widely used for prediction/classification in different specialties of medicine, such as neurology, cardiology, and pulmonology; in contrast, the study and development of ML is in its early stages in the nephrology field [15], [16]. However, the technological boom and the enormous volume of stored data can offer great benefits to this medical specialty [16], [26].

IDH is one of the most common complications in patients with CKD [7], [27]–[29] It is characterized by hemodynamic changes during HD treatment, which are due to multiple factors [7]. The implementation of intelligent models that are able to predict/classify the occurrence of IDH would allow health care personnel to provide early and effective treatments, and to make more informed decisions. Some of the most relevant work in this context is described below.

In [29], the authors predicted IDH using photoplethysmography signal fingerprinting (PPG), and subsequently employed a genetic algorithm for the extraction of the most important features and compared it with other algorithms for automatic IDH classification. They found that using their proposed method with the Adaboost algorithm showed good performance in classifying IDH and pre-IDH patients, where they obtained 90.6%, 86% and 93% accuracy, sensitivity and specificity respectively. Although these researchers obtained good results in terms of IDH classification, their work did not incorporate the analysis with other parameters (such as biochemical or analytical factors) that influence the occurrence of IDH.

The authors of [30], [31], studied the influence of heart rate variation and/or decreased oxygen saturation in patients with a view to predicting IDH. In the first work they reported satisfactory results, where they achieved an AUC of 0.63. Additionally, in the second study they indicated good results when predicting IDH with 80% accuracy. Therefore, they suggested that these parameters were useful for predicting IDH during HD; however, these researchers require justification of the effectiveness and consistency of these models.

Other researchers conducted a prospective cohort study [32] in which they analyzed heart rate variability together with some clinical parameters and patient laboratory results to predict IDH one month in advance. Their multivariate negative binomial model showed a significant ability to predict IDH using these parameters, where the model obtained an AUC of 0.804.

In [33] the authors proposed an intelligent model that was capable of alerting about Blood Pressure irregularities during HD. They posed the need to predict the Systolic Blood Pressure (SBP) of the following HD sessions, as well as to improve the quality of life of patients receiving HD. To do this, they used a dataset with 9,245 records, which presented 248 registered patients with their clinical variables (i.e., age, dialysis duration, blood flow, etc.). They compared different ML models (i.e., linear model, random forest, support vector regression, XGBoost and LASSO regression) and ensemble method. They noted that the random forest model (R2=0.95, RMSE=6.6) and XGBoost (R2=1.0, RMSE=1.83) obtained comparable performance in the training phase; however, in the test phase the random forest model (R2=0.49, RMSE=16.2) performed better. The ensemble method (R2=0.50, RMSE=16.01) performed best for the next SBP prediction.

Similarly, the study in [34] presented an intelligent early warning system that was able to predict IDH by checking the next Blood Pressure (BP) reading. They used a database of 653 patients, of whom underwent 55,516 HD sessions, leaving 285,705 valid BP records. The authors applied time-dependent logistic regression analysis to build predictive models, using 13 different patient clinical variables (i.e., age, genre, dry weight, dialysis duration, etc.). Their models obtained favorable results of 86% and 81% in terms of sensitivity and specificity for both nadir systolic BP (SBP) of $<$90 mmHg and $<$100 mmHg. In addition, their model obtained results of 64% and 64% in terms of sensitivity and specificity for SBP of $\geq$20 mmHg.

In [35], a deep neural network (DNN) model was proposed with the potential to determine the clinical factors that are related to the occurrence of IDH during an HD session. The researchers collected demographic data, HD clinical variables and laboratory data to identify factors associated with IDH. They tested different ML models (i.e., support vector machine, artificial neural network, random forest, decision tree, k-nearest neighbor, naïve Bayes) and showed that the proposed DNN-model was superior to the alternative models. They found that the 4-factor locus achieved great performance in the evaluation metrics (accuracy of 64.97% and sensitivity of 87.97%). However, the results obtained were not sufficiently satisfactory for this method to be employed in a medical setting during an HD session.

Likewise, in [36], a model based on a recurrent neural network (RNN) was developed to give real-time predictions of the risk of a patient presenting with IDH. The authors used different datasets where defined intradialytic hypotension when nadir systolic blood pressure (SBP) was $<$90 mmHg (IDH-1) or when a decrease in SBP $\geq$20 mmHg and/or a decrease in mean arterial pressure $\geq$10 mmHg on the basis of the initial SBPs (IDH-2) or prediction time SBPs (IDH-3) occurred within 1 hour. They compared the evaluation metrics resulting from the RNN model and other ML models (i.e., logistic regression, multi-layer perceptron and LightGBM), and they found that the RNN model had the best values in the different datasets. The RNN model achieved AUROCs of 0.94, 0.87 and 0.79 for IDH-1, IDH-2 and IDH-3 respectively.

In [37], ML algorithms were applied to develop models predicting hypotension after initiating continuous renal replacement therapy (CRRT). The authors defined hypotension as a reduction in mean arterial pressure of 20 mmHg from baseline within six hours. They tested different ML models including a support vector machine (SVM), a deep neural network (DNN), a light gradient boosting machine (LightGBM) and an extreme gradient boosting machine (XGBoost). Of these, the XGBoost model presented the best performance with an AUROC of 0.828. The authors noted that the use of ML algorithms could improve the predictability of hypotension after initiating CRRT.

The study in [38], used time-series differencing to extract the characteristics that most influenced the occurrence of IDH. This was called the time-relevant difference and was calculated based on the current time and the previous three IDH occurrence values. The researchers tested the generated features with non-time-series algorithms and showed that there was an improvement in the performance of the different algorithms when using this type of approach. For example, the LightGBM model achieved a performance of 86%, 66%, 88.9%, 85.1% and 0.946 for accuracy, precision, sensitivity, specificity and AUROC respectively. Therefore, they considered that this approach would be useful in small hospitals or settings where data were collected in a shorter time, as it gave acceptable values.

The aforementioned studies presented several models to predict or classify the occurrence of IDH in patients receiving HD. Each of these studies used a different approach or methods to treat the data arising from HD sessions. However, most of the related studies have sought to predict IDH while the patient is receiving HD, meaning that medical staff need to constantly monitor or receive alerts on the likelihood of IDH occurring in the patient. This can cause stress to nephrology staff, which may cause them to provide erroneous treatments to the patient.

Moreover, it is well known that the datasets in the nephrological domain are massive and heterogeneous, since they include a large number of variables (such as demographic data, laboratory results, analytical variables, etc.) and are collected during or after each patient's HD session, which means that the datasets have a high degree of dimensionality. In this sense, it would not be feasible to develop ML-based models with this large number of variables, as it would affect the performance and efficiency of each of these models [39].

For these reasons, our aim in this study is to provide new insight in predicting the occurrence of IDH in patients receiving HD. We developed and evaluated an intelligent ML-based model that will work at the start of the HD session, and it will allow medical staff to know whether the patient will have or will not have IDH during the HD session. To achieve this, we consider a combination of known clinical and analytical variables at the beginning of the HD session that can be associated with the occurrence of IDH through massive data analysis, which is described in the next section.

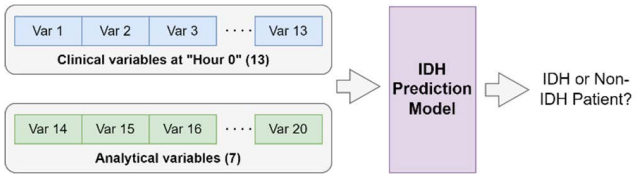

FIGURE 1. Model to predict IDH at the beginning of the HD session.

In the future, our proposed model will be implemented and deployed as a service in a real second-generation microservices ecosystem.

III. METHODOLOGY

In this section, the methodological process performed in the study is described and the data source is discussed.

A. IDH PREDICTIVE MODEL CONCEPT

The objective of the model (classifier) is to detect at the beginning of an HD session whether a patient is susceptible to IDH. For this purpose, a model built from data corresponding to multiple HD patients' sessions has been developed, including a set of analytical variables, from laboratory tests, close to the date of the HD session and the values of another set of clinical variables recorded at the beginning of the HD session in the so-called "hour 0" and of the detection of a drop in SBP throughout the dialysis session as indicated in the following Subsection. In this way, the predictive model will be fed at the beginning of the HD session by the data of the analytical variables available for the patient and the clinical values recorded at the beginning of said session at the so-called hour 0 (see Fig. 1).

B. DATA SOURCE

The dataset was obtained from the Hospital Príncipe de Asturias in Madrid, Spain, and contained records relating to clinical variables of patients who had received HD treatment between January 2016 and October 2019. These records had ethical approval from the hospital and the database was completely anonymized. We used as a starting point, the dataset of clinical variables that was analyzed and transformed in our previous study [40].

Unlike the previous dataset, in this work we defined IDH as a decrease in systolic blood pressure (SBP) of 20 mmHg or more if any of the SBPs measured at "Hour 1", "Hour 2", "Hour 3", and "Hour 4" was less than the SBP measured at "Hour 0" [41] that is illustrated in Fig. 2. Additionally, as a novel contribution, this work has incorporated new data on analytical variables extracted from the hospital laboratory.

In view of the above, the dataset used for this study was obtained by merging two different databases, as shown in Fig. 3. The first was drawn from the records of dialysis sessions and the variables were automatically recorded by the dialyzer itself in the HD sessions, while the second was based on the variables arising from a blood analysis performed on the patient. The two databases were merged in such a way that each HD session was linked to both the data from that HD

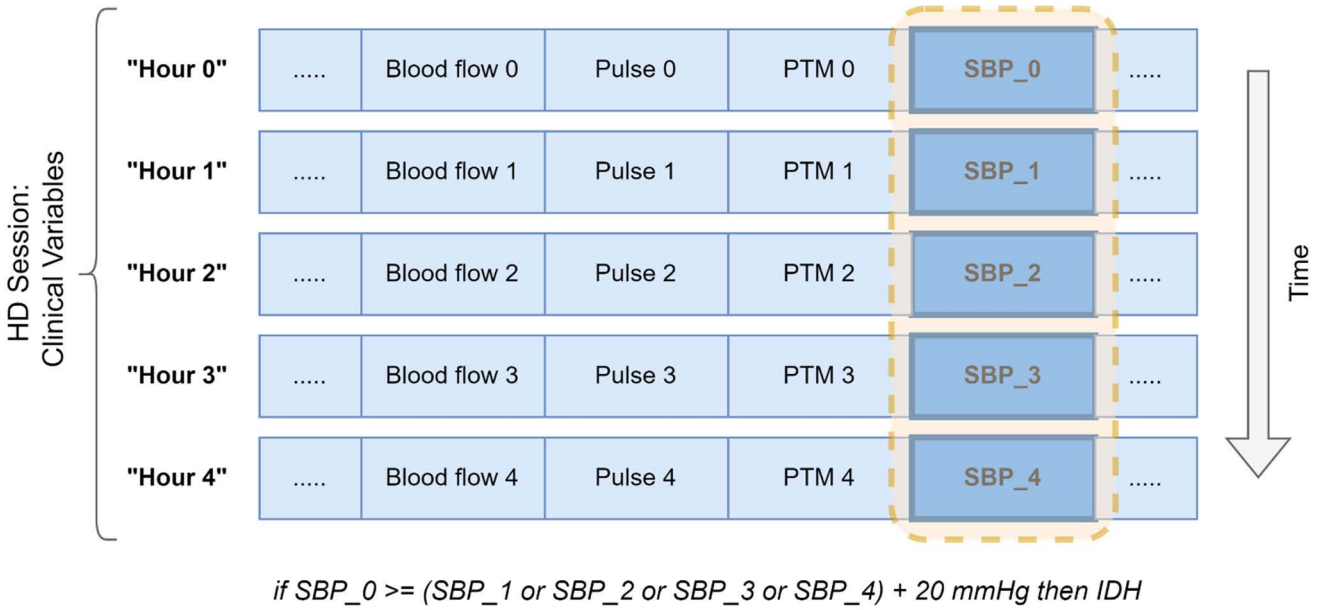

FIGURE 2. Identification and coding of an IDH during an HD session.

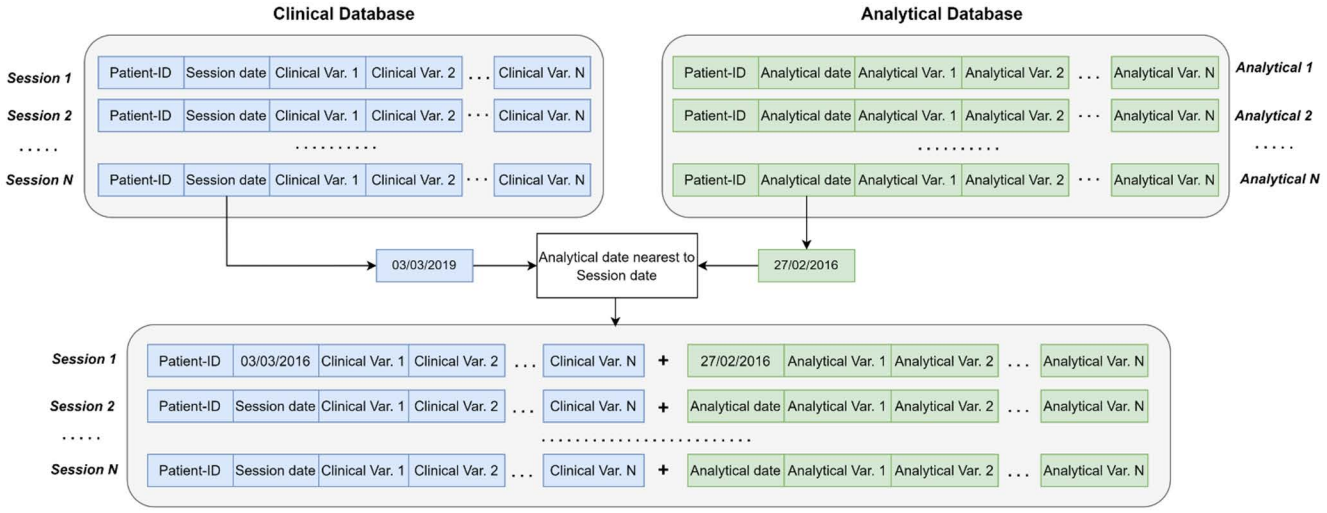

FIGURE 3. Procedure performed to merge the clinical and analytical databases.

session and the blood test carried out closest to that session. For example, if the patient underwent a quarterly blood test in an HD session, the quarterly blood test closest in time to that session was linked.

1) VARIABLE SELECTION

Once merged, the dataset contained more than 200 variables reflecting the clinical and analytical data related to the HD session. However, it is necessary to reduce this number of variables (parameter optimization), since the development of clinical models with a large number of variables is not advisable or practical. In fact, the selection of variables is an important and costly process in the development of clinical prediction models, so, if these variables are excessive or insignificant and do not improve the model, the medical staff, under their experience and clinical knowledge, determine which are the most important and influential variables for the development of the model [42], [43].

In this sense, we perform a second optimization process by reducing the dimensionality of the parameters. For this, we use ensemble trees [44] as a data mining technique and we consider the criteria and domain knowledge of the expert medical staff to identify the relevance of clinical and analytical variables. This left 20 variables (13 clinical and seven analytical), which are shown in Table 1 in the following subsection.

2) STUDY POPULATION

The generated dataset contained information from 22,234 sessions with 299 patients, in which 80% (17,793) of the

TABLE 1. Clinical and analytical variables from patients' hemodialysis sessions.

Type of variable	Variables	Counts (%)	Mean (± SD)
Clinical Variables Recorder at "Hour 0"	**(1) Age (years)**		66.4 (15.7)
	(2) Sex: n (%)		
	Male	16629 (74.8 %)	
	Female	5605 (25.2 %)	
	(3) Weight gain/Weight (before) (kg/kg)		0.024 (0.015)
	(4) Blood flow (mL/min)		379 (61.1)
	(5) Bath conductivity (mS/cm)		14 (0.9)
	(6) PTM (mmHg)		144.1 (55.9)
	(7) SBP (mmHg)		131.7 (28.1)
	(8) Pulse (bpm)		70.3 (13.7)
	(9) Ultrafiltration (mL)		147.1 (81.1)
	(10) Body temperature (°C)		36.5 (0.4)
	(11) Dialyzers: n (%)		
	Evodial	390 (1.8 %)	
	Sureflux 2.1	452 (2.0 %)	
	ELISIO 210	5181 (23.3 %)	
	FILTRYZER NF-2.1H	582 (2.6 %)	
	FX 100	4212 (18.9 %)	
	FX 80	929 (4.2 %)	
	FX CorDiax 800	8450 (38.0 %)	
	Solacea 21H	2038 (9.2 %)	
	(12) Baths: n (%)		
	ACF 3A1	1558 (7.0 %)	
	ACF 3A12	2517 (11.3 %)	
	ACF 3A2	1583 (7.1 %)	
	ACF 3A5	15103 (67.9 %)	
	ACF 3A6	1473 (6.6 %)	
	(13) Dialyzer techniques: n (%)		
	HD – High flow	426 (1.9 %)	
	HD – Low flow	5 (0.0 %)	
	HDF - On line	21803 (98.1 %)	
Analytical Variables prior to the session	**(14) Leukocyte (x10e3/µL)**		6.7 (2.6)
	(15) Urea (mg/dL)		108.4 (37.5)
	(16) Potassium (mmol/L)		4.7 (0.9)
	(17) Phosphate (mg/dL)		4.6 (1.4)
	(18) Ionized calcium (mmol/L)		1.2 (0.1)
	(19) PH venous blood (ND)[1]		7.4 (0.1)
	(20) PCR (mg/L)		18.4 (31.9)

[1]Not Dimensional.

patient sessions did not involve IDH and 20% (4,441) did report the occurrence of IDH during the HD session.

All of the clinical and analytical variables used in this study are detailed in Table 1. It should be noted that categorical variables are presented as a quantity (percentage), and numerical variables are presented as a mean (±standard deviation (SD)).

C. DATA PRE-PROCESSING

Data pre-processing was performed in three phases, to determine how well the model performed in terms of predicting the occurrence or non-occurrence of IDH in patients' HD sessions. These three pre-processing phases are described below:

1) DATA CLEANING

Some variables with null values and several outliers were found in the records of the patient sessions, which were eliminated based on the domain knowledge provided by the expert medical staff involved in this study.

TABLE 2. Encoding categorial variables of the dataset.

Categorical Variables	Encoding
Sex	$\{s \dots S\}_{s=1}^{S}$, S = 2
Dializers	$\{d \dots D\}_{d=1}^{D}$, D = 8
Baths	$\{b \dots B\}_{b=1}^{B}$, B = 5
Dializer techniques	$\{t \dots T\}_{t=1}^{T}$, T = 3

2) ENCODING AND STANDARDIZATION OF DATA

Once the dataset was cleaned, we proceeded to encode the categorical variables in the dataset. Since our goal in this process was to avoid unnecessarily increasing the dimensionality of the dataset, we limited the number of dummy variables created from the categorical values [45]. Therefore, we applied the Label Encoding approach, since it allows us to identify the existing values of the categorical variables and replace them with a numerical value. An example of these encoded categorical variables is shown in Table 2.

We then developed a neural network-based model, specifically a multi-layer perceptron (MLP), for the prediction of

IDH and non-IDH patients. It is well-known that it is not recommended to train this type of model on raw data, so we applied standard normalization to the numerical variables of the dataset for the MLP.

3) DATA BALANCING

The dataset was unbalanced with respect to the IDH and non-IDH classes. To remedy this situation and improve the classification performance of the ML models, we balanced the data by applying the synthetic minority oversampling technique (SMOTE) and edited nearest neighbors (ENN) [46]–[49]. Using this method, we performed multiple tests to greatly reduce the non-IDH class while maintaining sufficiently high consistency between the proportions of data in both classes.

After applying SMOTE-ENN, the dataset was reduced to 14,803 sessions, of which 10,294 (70%) belonged to the non-IDH class and approximately 4,509 (30%) to the IHD class, meaning that the dataset was better distributed for classification. Although, the generated dataset does not show an optimal balance in the classes, this will not affect the performance of the ML models, since both classes are well represented and do not overlap, therefore, the ML models will be able to correctly classify the studied classes [50]. All ML models developed in this study were trained, evaluated and validated on this dataset.

D. MACHINE LEARNING MODELS

In this work, we developed and evaluated different ML models (classifiers) to predict the occurrence of IDH or non-IDH at the beginning of the HD session. In the following, we briefly describe each of these classifiers.

1) LOGISTIC REGRESSION

Logistic Regression (LR) is one of the most common ML models studied in the literature. This model employs a logistic function to predict the probability of a variable being classified in a two-class or target problem. This resulting variable is considered dichotomous [51]. Although, LR is easy to implement, interpret and efficient to train, it has difficulties in dealing with nonlinear problems and, moreover, the major limitation of LR is the assumption of linearity between the dependent variable and the independent variables [52].

2) RANDOM FOREST

Random forest (RF) is an ML algorithm, which is based on the ensemble method and is usually trained with the bagging (bootstrapping + aggregation) method. The idea behind the bagging method is that it combines multiple decision trees and each of them are individually trained with a different subset of the dataset features, thus generating a more accurate and stable prediction or classification [53], [54]. Some of the goodness of RF is that it performs well for nonlinear problems, is robust to outliers, has good accuracy in classification tasks. However, it is slow to train with large datasets, it is not easy to interpret and has biases when dealing with categorical variables.

3) MULTI-LAYER PERCEPTRON

The Multi-Layer Perceptron (MLP) is a type of artificial neural network, which consists of a series of neurons (nodes) that are fully connected. MLP networks are generally composed of three layers; an input layer, one or more hidden layers and an output layer. In addition, they employ nonlinear functions, called activation function, at the nodes of the hidden and output layers [54], [55]. This type of network allows distinguishing data that are not linearly separable, the prediction is very fast when the model is trained, it has the ability to learn in real time. On the contrary, MLP networks require the adjustment of several hyperparameters, it is not easy to know the influence of the independent variables on the dependent one, because they are a black-box and, in addition, they are sensitive to the scaling of the variables.

4) EXTREME GRADIENT BOOSTING

The eXtreme Gradient Boosting (XGBoost) model is an ensemble method based on decision trees, that uses a gradient boosting framework [33], [56]. The XGBoost incorporates techniques (i.e., regularization, sampling, pruning) as a measure to avoid the overfitting problem present in the gradient boosting (GB) algorithm. It can be run in parallel and using multiple cores and can therefore make use of the high computational power available in graphics cards and can even be run on server clusters. Another of the advantages of XGBoost is that it does not require scaling or normalization the data and it also has the ability to efficiently handle missing values [55].

5) HYPERPARAMETERS OF MODELS

Hyperparameter optimization or hyperparameter tuning consists of objectively search different values for model hyperparameters and choose a subset that results in a model that achieves the best performance on a given dataset. In this sense, the hyperparameters were tuned to each model using random search cross-validate algorithm, which allows testing different ranges of hyperparameter values and thus generating models with reliable and stable performance. It should be noted that each model has its own set of hyperparameters, therefore, we selected those hyperparameters that most influence the performance and accuracy of the model [54]. The best values of the hyperparameters of the different models are shown in Table 3.

E. MODEL DEVELOPMENT

In this work, each of the models studied were developed using the Python programming language (v3.6). The focus of the classifier models is to predict whether a patient will have IDH or non-IDH at the beginning of the HD session. We considered combining hold-out and cross-validation methods, as they allow to evaluate the predictive ability of the model, prevent overfitting and generate robust and reliable performances [57]–[60].

TABLE 3. Hyperparameter values obtained in the models.

Model	Hyperparameter values
Logistic Regression	Penalty = L2; max iterations = 50.
Random Forest	Max depth = 7, max features = auto, number of estimators = 30.
Multi-layer Perceptron	Size hidden layers=[64,32,1], activation= [relu and sigmoid], batch size= 64, epochs= 50, optimizer= Adam, dropout= 0.2.
XGBoost	Max depth = 6, number of estimators = 46, learning rate = 0.2, subsample = 0.75, grow policy = lossguide.

Within this framework, we first applied the hold-out method [61], where we divided the data into an 80% training-set and a 20% testing-set. With the training-set, the different ML models were evaluated by applying the Stratified K-fold cross validation method [62]–[64], which allows us to evaluate the performance and robustness of the models handling this data set. Subsequently, testing-set was used to evaluate the actual ability of the models to generalize from this unseen data, as well as to avoid introducing biases in the model training phase [65], [66]. The performance of the models was calculated using different metrics described in the following subsection. The entire development process is detailed in the pseudocode shown in Algorithm 1.

Algorithm 1 Pseudo-Code for the Development and Evaluation of Models

1. Begin
2. Load data source (Clinical-Analytical Data)
3. Apply pre-processing methods in the Clinical-Analytical Data.
4. Data splitting in training and testing
5. Apply Stratified K-fold for each ML model using training data
6. Validate each ML model using testing data
7. Calculate performance metrics
8. End

F. MODEL EVALUATION AND VALIDATION

We used Stratified K-fold cross-validation to evaluate models [62]–[64]. This consisted of dividing the training-set into k folds while maintaining the ratio between the classes (IDH and non-IDH). We applied five folds to each model due to the number of samples in the training-set [57]; at each iteration, a different fold was selected to train the model, and it was tested on the rest of the folds.

Although, the Stratified K-fold cross-validation allowed us to have a first evaluation of the performance of the models, these were validated with the testing-set, since this is a data set not used by the models and, therefore, allows us to validate the robustness and reliability of each one.

1) PERFORMANCE METRICS

Because there are a wide variety of metrics to evaluate models, we consider combining several metrics and interpreting their results in a holistic manner. [66]. For this reason, for each of the models, we evaluated several performance metrics, which include accuracy, precision, recall, F1-score and Matthews Correlation Coefficient (MCC) [20], [67], [68]. Equations for each of these metrics are (1)–(5), as shown at the bottom of the next page, where TP, TN, FP and FN represent true positives, true negatives, false positives and false negatives, respectively.

We also generated other evaluation and validation measures that are important to our study and in the medical context are of great importance [69]. These are Receiver Operating Characteristic curve (ROC) and Precision-Recall curve (PRC), since they allow determining the predictive performance of the model, provide readers with a general understanding of the utility of the model, and also allow evaluating the performance of the model when the data present some imbalance [21], [60], [70]. Likewise, for ROC and PRC, the areas under the curve (AUC) of both metrics were calculated and denoted as AUROC and AUPRC respectively.

To interpret and calculate the AUROC and AUPRC curves, we used the DeLong test [71]. This involved comparing the values obtained by the ML models, where the p-values were bilateral and those less than 0.05 were considered to be statistically significant. All comparisons were performed using the testing-set.

2) RECEIVER OPERATING CHARACTERISTIC CURVE (ROC)

This is one of the most widely used evaluation criteria and information on the ability of a model to distinguish between classes (in this case, IDH and non-IDH patients) [54], [67]. It does this by comparing the true positive rate (TPR) with the false positive rate (FPR). As TPR (recall or sensitivity) has already been defined above, the FPR can be defined as shown below:

$$FPR = \frac{FP}{FP + TN} \tag{6}$$

3) PRECISION-RECALL CURVE (PRC)

The PRC represents the performance of each of the different ML models against different thresholds, rather than a single metric (such as accuracy or F1-score). A classifier will show good performance if it has high accuracy and high recall. The equations for this curve were previously defined in (2) and (3).

It is important to note that, for our study we consider the F1-score, AUROC, AUPRC and MCC metrics as the main measures of comparison of the studied models, since they are useful to evaluate the performance of dichotomous classification models and to validate the robustness of the models with respect to a dataset with some imbalance as the one studied here [21].

IV. RESULTS

ML models have been developed in the Python programming language (v3.6) and implemented on a machine with an NVIDIA Jetson Nano Developer Kit [72]. To avoid biases in the comparison of the results obtained, all models studied

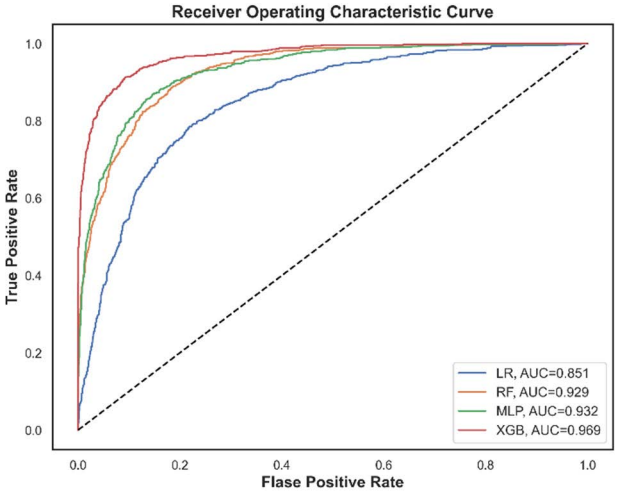

FIGURE 4. Receiver operation characteristic curve (ROC) and the area under the curve (AUC) for the different classification models developed here: LR, logistic regression; RF, random forest; MLP, multi-layer perceptron; XGB, XGBoost.

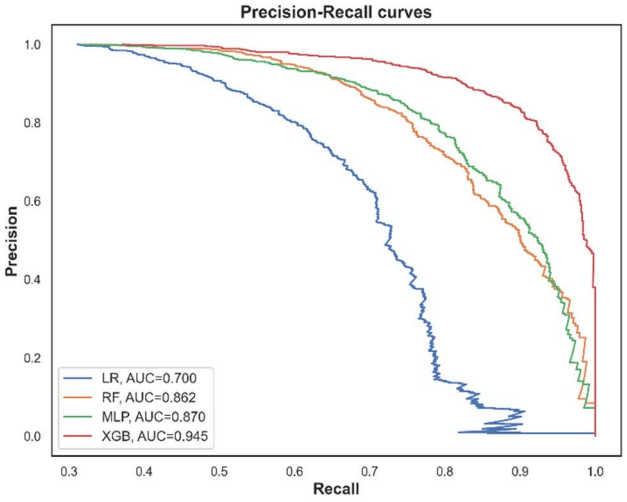

FIGURE 5. Precision-recall curve and area under the curve for the different classification models developed here. LR, logistic regression; RF, random forest; MLP, multi-layer perceptron; XGB, XGBoost.

in this work were also implemented using the same hardware and language resources as the model finally proposed.

To evaluate the performance of the models, we first used the training-set with the Stratified K-fold cross-validation method (Section III-E). During the model training process, the different metrics described in Section III-F were calculated. The results obtained by each model in predicting whether the patient will have a class (IDH or non-IDH) are shown in Table 4. Importantly, the values were calculated as the mean (±SD) for the metrics of accuracy, precision, recall, F1-score and MCC for both classes (in this case, IDH and non-IDH patients).

As a follow-up of this activity, we applied on the trained models the data remaining from applying the hold-out method (testing-set), as these were never seen or used by the models. Therefore, the results obtained by the models on the metrics applied in this work are shown in Table 5.

Once the models were trained and evaluated, we proceeded to perform a more exhaustive analysis of the models using the test-sets. In this sense, we plotted the ROC and PRC curves shown in Figs. 4 and 5. These plots visualize the ability of the models to predict whether the patient will have IDH or non-IDH at the beginning of the HD session.

As a complement, we calculated and compared the AUROC and AUPRC metrics using the DeLong test of the pRoc library [73]. The results of comparing these metrics in the models allowed us to interpret the ability, robustness and reliability of each of the models to discriminate between IDH and non-IDH classes. The results for the ROC and PRC curves for each model are shown in Table 6, with the AUC (AUROC and AUPRC) values for each model and their 95% confidence intervals (CI). Similarly, the p-value was obtained when comparing the ROC curve of the XGBoost model using the DeLong test, since this model showed the best performance.

Comparing the data in Tables 4, 5 and 6, it can be seen that the XGBoost model has the best predictive capacity when determining whether the patient will have IDH or non-IDH at the start of the HD session. However, it is important to verify that this resulting model has the capacity to generalize with records of future HD patients and, in addition, to verify that the model does not present overfitting. Therefore, Fig. 6 shows the learning curve of the XGBoost model, where

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2 * (Precision * Recall)}{Precision + Recall} \tag{4}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{5}$$

TABLE 4. Classification results of models using Stratified k-fold cross-validation method (training-set).

Models	IDH Patient			Non-IDH Patient			Accuracy	MCC
	Precision	Recall	F1-score	Precision	Recall	F1-score		
LR	0.72 (0.002)	0.56 (0.015)	0.63 (0.017)	0.82 (0.003)	0.90 (0.007)	0.86 (0.004)	0.80 (0.005)	0.49 (0.013)
RF	0.87 (0.008)	0.61 (0.019)	0.71 (0.015)	0.85 (0.005)	0.96 (0.005)	0.90 (0.003)	0.85 (0.006)	0.63 (0.009)
MLP	0.82 (0.012)	0.74 (0.012)	0.78 (0.006)	0.88 (0.007)	0.938 (0.01)	0.91 (0.002)	0.87 (0.009)	0.70 (0.016)
XGB	**0.90 (0.013)**	**0.82 (0.014)**	**0.86 (0.013)**	**0.92 (0.006)**	**0.96 (0.005)**	**0.94 (0.004)**	**0.92 (0.007)**	**0.81 (0.017)**

LR, logistic regression; RF, random forest; MLP, multi-layer perceptron; XGB, XGBoost.

TABLE 5. Classification results of models using testing-set.

Models	IDH Patient			Non-IDH Patient			Accuracy	MCC
	Precision	Recall	F1-score	Precision	Recall	F1-score		
LR	0.71	0.59	0.64	0.83	0.89	0.86	0.80	0.51
RF	0.85	0.60	0.71	0.84	0.95	0.90	0.85	0.62
MLP	0.84	0.72	0.77	0.88	0.94	0.91	0.87	0.68
XGB	**0.90**	**0.83**	**0.86**	**0.93**	**0.96**	**0.94**	**0.92**	**0.81**

LR, logistic regression; RF, random forest; MLP, multi-layer perceptron; XGB, XGBoost.

TABLE 6. Results of areas under the curve (AUCs) to classify intradialytic hypotension of ML models.

Models	AUROC (95% CI)	p-value[1]	AUPRC (95% CI)
XGB	0.969 (0.963 to 0.975)		0.945 (0.929 to 0.958)
MLP	0.932 (0.922 to 0.941)	< 0.0001	0.87 (0.847 to 0.890)
RF	0.929 (0.920 to 0.938)	< 0.0001	0.862 (0.838 to 0.890)
LR	0.851 (0.837 to 0.866)	< 0.0001	0.70 (0.669 to 0.728)

[1]Comparison with the ROC curve of the XGBoost model using DeLong test.

TABLE 7. Comparison of our obtained model (XGBoost) with other studies.

Studies	Variables		Metrics				
	Clinical	Analytical	Acc.	F1	MCC	AUROC	AUPRC
S. Park et al., [32]	12	25	-	-	-	0.80	-
J. Chen et al., [35]	12	13	0.65	-	0.19	0.65	-
H. Lee et al., [36]	15	11	-	0.69	-	0.87	0.78
M. W. Kang et al., [37]	9	> 60	0.76	0.66	0.49	0.828	-
J.-Y. Yang et al., [38]	28	-	0.86	-	-	0.94	-
Our model	**13**	**7**	**0.92**	**0.86**	**0.81**	**0.97**	**0.94**

the graph shows the learning curve of the LogLoss error metric. This metric is based on probability and is used to measure the performance of the model.

Additionally, we compared the results of the XGBoost model obtained with the most recent works related to our research (Section II and III-B). In this sense, Table 7 shows the different studies and their respective results for the Accuracy (Acc.), F1-score (F1), MCC, AUROC and AUPRC.

Finally, we plotted the importance of each of the features according to the XGBoost model obtained in this study. In this way, the feature importance plot allowed us to determine the utility or value of each of the features (variables) involved in the construction of the model [74]. The calculation of the score for each feature was determined by its impact (gain) on the performance measure: the higher the value, the greater the impact of the feature (variable) on the performance of the model. Fig. 7 shows the importance of each feature according to the XGBoost model.

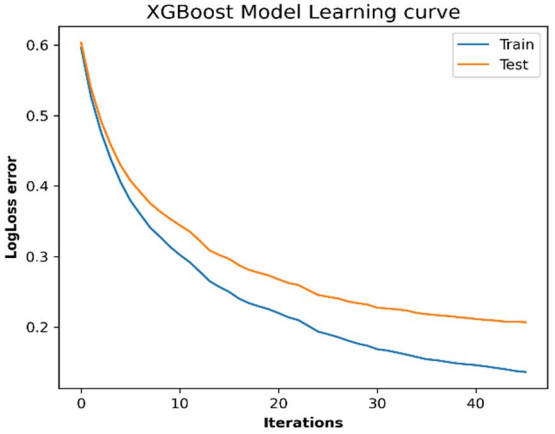

FIGURE 6. XGBoost model learning curve.

V. DISCUSSION

The purpose of this research was to develop and evaluate several ML models (classifiers) that would allow us to predict whether a patient will suffer from IDH or be non-IDH during an HD session based on a combination of clinical and analytical variables. Another objective of this research was to identify the model that has the best ability to discriminate between these classes.

Based on the performance analysis carried out here, we can observe from Table 4 that the results were acceptable in terms of the metrics evaluated (Section III-F). The XGBoost model achieved higher values by classifying both classes (IDH and non-IDH patients) compared to the MLP, RF and LR models. To use this predictive model in clinical practice, the results for predicting whether the patient will suffer from IDH were the most important. The XGBoost model obtained values of 0.92, 0.90, 0.82, 0.86 and 0.81 for the accuracy, precision, recall, F1-score and MCC, respectively, followed by MLP with 0.87, 0.82, 0.74, 0.78 and 0.70 and RF with 0.85, 0.87, 0.61, 0.71 and 0.63, respectively. The worst performer was LR, with values of 0.80, 0.72, 0.56, 0.63 and 0.49.

In the same way, we performed a validation of the different models by applying the testing-set, with the objective of validating the performance of the models with data that

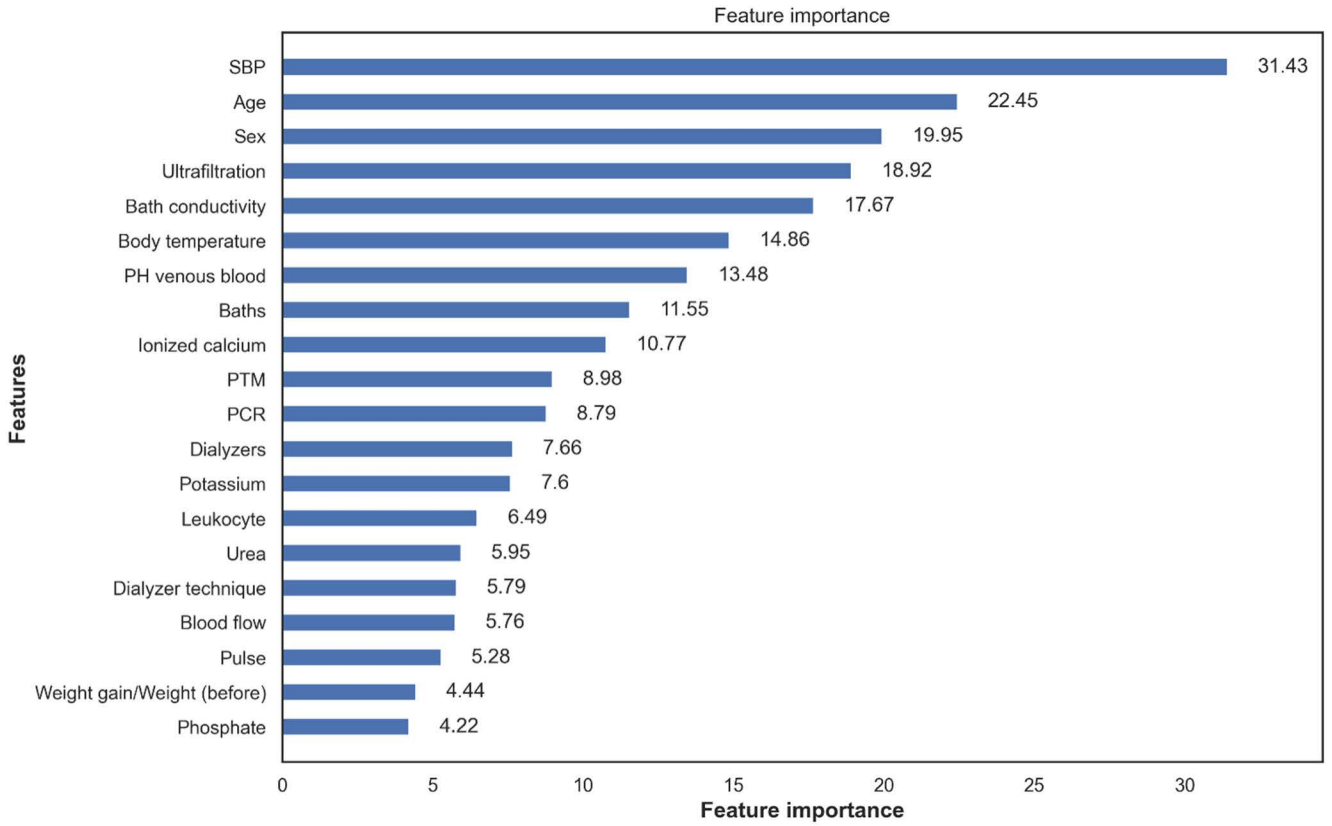

FIGURE 7. **Importance of each feature as determined using the XGBoost model.**

were not used for training. In this sense, it is evident in Table 5 that the XGBoost model also maintains a better performance over the other models when using these records and, in addition, the MCC metric is superior to the rest of the models; therefore, the model has a robustness when classifying patients (IDH or non-IDH). These results present an encouraging view of the XGBoost model in terms of identifying the occurrence of IDH or non-IDH in patients and may contribute to providing additional information in clinical practice. Furthermore, the degree of significance obtained from the different clinical and analytical variables opens possibilities for nephrology clinicians to study the complex task of determining the multiple factors associated with the occurrence of IDH [7], [11].

However, we also considered evaluating the different models by analyzing their ROC and PRC curves, as well as their areas under the AUROC and AUPRC curves. With these considerations in mind, it is evident from the graphs in Figs. 4 and 5, that the XGBoost model presents a good predictive ability when determining whether the patient will suffer IDH or non-IDH at the beginning of the HD session compared to the rest of the models studied. Furthermore, the DeLong test was used to compare and interpret these graphs. The results for the areas under the curves (AUROC and AUPRC) observed in Table 6 show a very good performance of the XGBoost model compared to other developed models. Additionally, there was a statistically significant difference ($p < 0.0001$) in the AUROC, meaning that this model was able to correctly identify and distinguish between the two classes (IDH and non-IDH) [75]–[77].

In addition, it is worth noting that the XGBoost model has excelled in each of the main metrics of this study (Section III-F). Within this context, we observe in Fig. 6 that the model has shown favorable performance for both training-set and testing-set and, furthermore, it is apparent from the LogLoss error curves that the model does not overfit, as the XGBoost model itself employs mechanisms to avoid this [55], [56], [78]. Therefore, this model would be able to generalize over records of future patients receiving HD.

We can highlight some important findings from these results. First, the XGBoost model achieved more accurate results than the rest of the ML models. This model also achieved better results for all metrics compared to previous work [40], where only the clinical variables of the patient HD were used, and presented better results than other studies [33], [35]–[37]. A possible explanation for these good results may lie in the novel incorporation of the clinical and analytical variables that were most influential, according to the applied ensemble-trees and the domain knowledge of nephrology experts in the creation of the ML predictor model. Reducing the dimensionality of the variables (clinical and analytical) provided a significant improvement and increased robustness

in the model's ability to classify the data into classes (IDH and non-IDH). The developed predictor will be able to help medical staff to treat a patient at the beginning of the HD session to avoid the probable occurrence of IDH.

The second finding was that the XGBoost model demonstrated superior and favorable performance compared to other works using more complex models [33], [35], [36]. Although DNN-based models have a complex and abstract structure, which gives them the ability to learn/capture the different hidden patterns in the information provided. However, they have certain disadvantages compared to models such as the one based on the XGBoost, as they are sensitive to the scale of the data, dependent on the structure and size of the data, require a large computational load to adjust the large number of parameters of the model, and in the case of binary classification, require well balanced data [79]. Furthermore, DNN-based models are currently considered to be black-box models, meaning that their internal behavior is unknown and it is not possible to identify which variables or features contribute most to a given decision [35], [36].

In the third finding, Table 7 shows that our obtained model (XGBoost) achieves a formidable performance compared to the methods applied in other recent studies. It is important to note that the different works used a greater number of variables (clinical and/or analytical) compared to ours; therefore, not necessarily dealing with more variables generates better performance. In addition, using a smaller set of variables allows to obtain a more useful model, the medical staff to better identify the influence of the variables and, in addition, the collection of patient records is more agile.

Another significant finding can be observed from Fig. 7, which shows the variables or features, according to the XGBoost model, that most influenced the classification between patients who did or did not suffer from IDH during an HD session. As discussed above, the ability of this model to determine the patient features that contribute most to a potential case of IDH is important for decision-making by medical staff [80], [81].

Although the results obtained here are encouraging, there are some limitations to this study. This was a retrospective cohort study, and the ability of the model to classify IDH from new patient records would need to be validated, meaning that prospective validation is required. It would also be desirable to test the model' s capability with patient records from other hospital centers and therefore significantly expand the number of HD sessions. Finally, it is important to emphasize that the model obtained may be more robust when applying other optimization techniques.

VI. CONCLUSION AND FUTURE WORKS

In this study, we developed and evaluated ML models with the ability to predict whether a patient will suffer from IDH during an HD session from the available analytical and clinical values recorded at the beginning of the HD session. In particular, we have focused on data provided by the nephrology staff at the Hospital Príncipe de Asturias, Madrid, Spain. Our aim was to develop a flexible, robust model to support decision-making by hospital medical staff, to anticipate and treat the occurrence of IDH in patients receiving HD in a timely manner. We based the approach on a combination of variables (clinical and analytical) that were strongly determinant for the occurrence of IDH during an HD session.

From this predictive point of view, the XGBoost model achieved the best performance compared to the rest of the ML models, obtaining very reliable results of 0.92, 0.90, 0.82, 0.86 and 0.81 for accuracy, precision, recall, F1-score and MCC respectively. It also showed a favorable performance in the area under the curve (AUC) of receiver operating characteristic of 0.969 (95% confidence intervals [CI], 0.963 to 0.975), which was the highest compared to the other ML models ($p < 0.0001$). Similarly, the AUC of precision-recall was 0.945 (95% CI, 0.929 to 0.958). This suggests that models based on the XGBoost may be useful for application in medical specialties such as nephrology due to their ability to be implemented and deployed within different computational paradigms (such as the cloud, fog or edge) [82] and their versatility in terms of utilizing the high computational power offered by graphics cards. In addition, this model provides information on the way in which the importance of features can influence the occurrence of IDH during the HD session; therefore, this is a key advantage over other black-box models such as DNNs [79].

Finally, in future work, we will improve the performance of the model obtained after applying other optimization techniques. Also, we will implement and deploy this model within a second-generation microservices ecosystem that focuses on the telemonitoring of patients receiving HD, to take advantage of all the benefits and opportunities offered by this novel software architecture [83]. We will also build different models from another point of view, with the aim of classifying the occurrence of IDH and the range of SBP decrease that is encountered (e.g., mild, moderate or severe). This could provide improvements in the actions taken by nephrology staff in reference to the treatment that needs to be applied to HD patients. As a consequence, personalized methods or mechanisms of patient care could be developed.

CONFLICTS OF INTEREST

No conflict of interest exists. The authors wish to confirm that, there are no known conflicts of interest associated with this publication.

ACKNOWLEDGMENT

The authors thank the Clinical Team of the Nephrology Section of the Hospital Universitario Príncipe de Asturias for their advice and transfer of the clinical and analytical data. They also would like to thank the National Secretariat for Science, Technology and Innovation (SENACYT, Panama) and the Institute for the Training and Use of Human Resources (IFARHU), for the Doctoral Study Scholarship awarded within the IFARHU-SENACYT Doctoral Scholarship Program, in 2018. They also thank the Fresenius Medical

Care AG company for their collaboration in carrying out this work.

REFERENCES

[1] E. Dovgan, A. Gradišek, M. Luštrek, M. Uddin, A. A. Nursetyo, S. K. Annavarajula, Y.-C. Li, and S. Syed-Abdul, "Using machine learning models to predict the initiation of renal replacement therapy among chronic kidney disease patients," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0233976, doi: 10.1371/journal.pone.0233976.

[2] M. Gorostidi, M. Sánchez-Martínez, L. M. Ruilope, A. Graciani, J. J. de la Cruz, R. Santamaría, M. D. del Pino, P. Guallar-Castillón, F. de Álvaro, F. Rodríguez-Artalejo, and J. R. Banegas, "Prevalencia de enfermedad renal crónica en españa: Impacto de la acumulación de factores de riesgo cardiovascular," *Nefrología*, vol. 38, no. 6, pp. 606–615, Nov. 2018, doi: 10.1016/j.nefro.2018.04.004.

[3] K. Brück *et al.*, "CKD prevalence varies across the European general population," *J. Amer. Soc. Nephrol.*, vol. 27, no. 7, pp. 2135–2147, Jul. 2016, doi: 10.1681/ASN.2015050542.

[4] J. Arrieta, A. Rodríguez-Carmona, C. Remón, M. Pérez-Fontán, F. Ortega, J. A. Sánchez-Tomero, and R. Selgas, "Cost comparison between haemodialysis and peritoneal dialysis outsourcing agreements," *Nefrología*, vol. 32, no. 2, pp. 247–248, 2012, doi: 10.3265/Nefrologia.pre2011.Dec.11311.

[5] M. Kitamura and A. Saito, "Dialysis hypotension: A review of recent studies of causative factors," *Nephrology*, vol. 6, no. 3, pp. 109–112, Jun. 2001, doi: 10.1046/J.1440-1797.2001.00053.X.

[6] J. A. Chou, K. Kalantar-Zadeh, and A. T. Mathew, "A brief review of intradialytic hypotension with a focus on survival," *Seminars Dialysis*, vol. 30, no. 6, pp. 473–480, Nov. 2017, doi: 10.1111/SDI.12627.

[7] M. Kanbay, L. A. Ertuglu, B. Afsar, E. Ozdogan, D. Siriopol, A. Covic, C. Basile, and A. Ortiz, "An update review of intradialytic hypotension: Concept, risk factors, clinical implications and management," *Clin. Kidney J.*, vol. 13, no. 6, pp. 981–993, Dec. 2020, doi: 10.1093/CKJ/SFAA078.

[8] P. B. Reeves and F. R. M. Causland, "Mechanisms, clinical implications, and treatment of intradialytic hypotension," *Clin. J. Amer. Soc. Nephrol.*, vol. 13, no. 8, pp. 1297–1303, Aug. 2018, doi: 10.2215/CJN.12141017.

[9] D. F. Keane, J. G. Raimann, H. Zhang, J. Willetts, S. Thijssen, and P. Kotanko, "The time of onset of intradialytic hypotension during a hemodialysis session associates with clinical parameters and mortality," *Kidney Int.*, vol. 99, no. 6, pp. 1408–1417, Jun. 2021, doi: 10.1016/j.kint.2021.01.018.

[10] M. M. Assimon and J. E. Flythe, "Definitions of intradialytic hypotension," *Seminars Dialysis*, vol. 30, no. 6, pp. 464–472, Nov. 2017, doi: 10.1111/SDI.12626.

[11] C. Barbieri, I. Cattinelli, L. Neri, F. Mari, R. Ramos, D. Brancaccio, B. Canaud, and S. Stuard, "Development of an artificial intelligence model to guide the management of blood pressure, fluid volume, and dialysis dose in end-stage kidney disease patients: Proof of concept and first clinical assessment," *Kidney Diseases*, vol. 5, no. 1, pp. 28–33, 2019, doi: 10.1159/000493479.

[12] J. Saez-Rodriguez, M. M. Rinschen, J. Floege, and R. Kramann, "Big science and big data in nephrology," *Kidney Int.*, vol. 95, no. 6, pp. 1326–1337, Jun. 2019, doi: 10.1016/j.kint.2018.11.048.

[13] J. U. Becker, D. Mayerich, M. Padmanabhan, J. Barratt, A. Ernst, P. Boor, P. A. Cicalese, C. Mohan, H. V. Nguyen, and B. Roysam, "Artificial intelligence and machine learning in nephropathology," *Kidney Int.*, vol. 98, no. 1, pp. 65–75, Jul. 2020, doi: 10.1016/j.kint.2020.02.027.

[14] K. Shailaja, B. Seetharamulu, and M. A. Jabbar, "Machine learning in healthcare: A review," in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Mar. 2018, pp. 910–914, doi: 10.1109/ICECA.2018.8474918.

[15] A. Verma, V. C. Chitalia, S. S. Waikar, and V. B. Kolachalama, "Machine learning applications in nephrology: A bibliometric analysis comparing kidney studies to other medicine subspecialities," *Kidney Med.*, vol. 3, no. 5, pp. 762–767, Sep. 2021, doi: 10.1016/j.xkme.2021.04.012.

[16] Q. Li, Q.-L. Fan, Q.-X. Han, W.-J. Geng, H.-H. Zhao, X.-N. Ding, J.-Y. Yan, and H.-Y. Zhu, "Machine learning in nephrology: Scratching the surface," *Chin. Med. J.*, no. 6, pp. 687–698, Mar. 2020, doi: 10.1097/CM9.0000000000000694.

[17] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," *Sensors*, vol. 21, no. 11, p. 3758, May 2021, doi: 10.3390/S21113758.

[18] E. García-Martín, C. F. Rodrigues, G. Riley, and H. Grahn, "Estimation of energy consumption in machine learning," *J. Parallel Distrib. Comput.*, vol. 134, pp. 75–88, Dec. 2019, doi: 10.1016/j.jpdc.2019.07.007

[19] S. Panicker, "Use of machine learning techniques in healthcare: A brief review of cardiovascular disease classification," in *Proc. 2nd Int. Conf. Commun. Inf. Process. (ICCIP)*, 2020, pp. 1–12, doi: 10.2139/SSRN.3681833.

[20] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Dec. 2020, doi: 10.1186/S12864-019-6413-7.

[21] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, Mar. 2015, Art. no. e0118432, doi: 10.1371/journal.pone.0118432.

[22] A. Mujumdar and V. Vaidehi, "Diabetes prediction using machine learning algorithms," *Proc. Comput. Sci.*, vol. 165, pp. 292–299, Jan. 2019, doi: 10.1016/j.procs.2020.01.047.

[23] P. Khan, M. F. Kader, S. M. R. Islam, A. B. Rahman, M. S. Kamal, M. U. Toha, and K.-S. Kwak, "Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances," *IEEE Access*, vol. 9, pp. 37622–37655, 2021, doi: 10.1109/ACCESS.2021.3062484.

[24] J. Lee, J. Woo, A. R. Kang, Y.-S. Jeong, W. Jung, M. Lee, and S. H. Kim, "Comparative analysis on machine learning and deep learning to predict post-induction hypotension," *Sensors*, vol. 20, no. 16, p. 4575, Aug. 2020, doi: 10.3390/S20164575.

[25] N. Caballé-Cervigón, J. L. Castillo-Sequera, J. A. Gómez-Pulido, J. M. Gómez-Pulido, and M. L. Polo-Luque, "Machine learning applied to diagnosis of human diseases: A systematic review," *Appl. Sci.*, vol. 10, no. 15, p. 5135, Jul. 2020, doi: 10.3390/APP10155135.

[26] P. Panwong and N. Iam-On, "Predicting transitional interval of kidney disease stages 3 to 5 using data mining method," in *Proc. 2nd Asian Conf. Defence Technol. (ACDT)*, Jan. 2016, pp. 145–150, doi: 10.1109/ACDT.2016.7437659.

[27] P. Reach, M. Touzot, Y. Lombardi, C. Maheas, E. Sacco, A. Fels, H. Beaussier, P. Ureña-Torres, G. Chatellier, C. Ridel, and M. Zuber, "Electrochemical skin conductance by Sudoscan®: A new tool to predict intradialytic hypotension," *Nephrol. Dialysis Transplantation*, vol. 36, no. 8, pp. 1511–1518, Jul. 2021, doi: 10.1093/ndt/gfab183.

[28] B. Sars, F. M. van der Sande, and J. P. Kooman, "Intradialytic hypotension: Mechanisms and outcome," *Blood Purification*, vol. 49, nos. 1–2, pp. 158–167, 2020, doi: 10.1159/000503776.

[29] M. Shahabi, V. R. Nafisi, and F. Pak, "Prediction of intradialytic hypotension using PPG signal features," in *Proc. 22nd Iranian Conf. Biomed. Eng. (ICBME)*, Nov. 2015, pp. 399–404, doi: 10.1109/ICBME.2015.7404178.

[30] C. Perazzini, P. G. Bolasco, L. Corazza, M. Tramonti, E. Mancini, A. Santoro, and S. Severi, "Prediction of intradialytic hypotension based on oxygen saturation variations," in *Proc. Comput. Cardiol.*, 2013, pp. 1223–1226.

[31] V. K. Y. Chan, L. W. C. Chan, and D. H. K. Chow, "Oxygen saturation and heart rate variations as predictors of intradialytic hypotension," *Nephrol. Nursing J.*, vol. 45, no. 1, pp. 53–61, 2018.

[32] S. Park, W.-J. Kim, N.-J. Cho, C.-Y. Choi, N. H. Heo, H.-W. Gil, and E. Y. Lee, "Predicting intradialytic hypotension using heart rate variability," *Sci. Rep.*, vol. 9, no. 1, p. 2574, Dec. 2019, doi: 10.1038/S41598-019-39295-Y.

[33] J.-C. Huang, Y.-C. Tsai, P.-Y. Wu, Y.-H. Lien, C.-Y. Chien, C.-F. Kuo, J.-F. Hung, S.-C. Chen, and C.-H. Kuo, "Predictive modeling of blood pressure during hemodialysis: A comparison of linear model, random forest, support vector regression, XGBoost, LASSO regression and ensemble method," *Comput. Methods Programs Biomed.*, vol. 195, Oct. 2020, Art. no. 105536, doi: 10.1016/j.cmpb.2020.105536.

[34] C.-J. Lin, C.-Y. Chen, P.-C. Wu, C.-F. Pan, H.-M. Shih, M.-Y. Huang, L.-H. Chou, J.-S. Tang, and C.-J. Wu, "Intelligent system to predict intradialytic hypotension in chronic hemodialysis," *J. Formosan Med. Assoc.*, vol. 117, no. 10, pp. 888–893, Oct. 2018, doi: 10.1016/j.jfma.2018.05.023.

[35] J.-B. Chen, K.-C. Wu, S.-H. Moi, L.-Y. Chuang, and C.-H. Yang, "Deep learning for intradialytic hypotension prediction in hemodialysis patients," *IEEE Access*, vol. 8, pp. 82382–82390, 2020, doi: 10.1109/ACCESS.2020.2988993.

[36] H. Lee, D. Yun, J. Yoo, K. Yoo, Y. C. Kim, D. K. Kim, K.-H. Oh, K. W. Joo, Y. S. Kim, N. Kwak, and S. S. Han, "Deep learning model for real-time prediction of intradialytic hypotension," *Clin. J. Amer. Soc. Nephrol.*, vol. 16, no. 3, pp. 396–406, Mar. 2021, doi: 10.2215/CJN.09280620.

[37] M. W. Kang, S. Kim, Y. C. Kim, D. K. Kim, K.-H. Oh, K. W. Joo, Y. S. Kim, and S. S. Han, "Machine learning model to predict hypotension after starting continuous renal replacement therapy," *Sci. Rep.*, vol. 11, no. 1, p. 17169, Dec. 2021, doi: 10.1038/S41598-021-96727-4.

[38] J.-Y. Yang, H. W. Hu, C.-H. Liu, K.-Y. Chen, C.-H. Un, C.-C. Huang, C.-C. Chen, C.-C.-K. Lin, H. Chang, and H.-M. Lin, "Differencing time series as an important feature extraction for intradialytic hypotension prediction using machine learning," in *Proc. IEEE 3rd Eurasia Conf. Biomed. Eng., Healthcare Sustainability (ECBIOS)*, May 2021, pp. 19–20, doi: 10.1109/ECBIOS51820.2021.9510749.

[39] S. S. Thakur, S. S. Abdul, H.-Y. Chiu, R. B. Roy, P.-Y. Huang, S. Malwade, A. A. Nursetyo, and Y.-C. Li, "Artificial-intelligence-based prediction of clinical events among hemodialysis patients using non-contact sensor data," *Sensors*, vol. 18, no. 9, p. 2833, Aug. 2018, doi: 10.3390/S18092833.

[40] J. A. Gómez-Pulido, J. M. Gómez-Pulido, D. Rodríguez-Puyol, M.-L. Polo-Luque, and M. Vargas-Lombardo, "Predicting the appearance of hypotension during hemodialysis sessions using machine learning classifiers," *Int. J. Environ. Res. Public Health*, vol. 18, no. 5, p. 2364, Feb. 2021, doi: 10.3390/IJERPH18052364.

[41] K. Workgroup, "K/DOQI clinical practice guidelines for cardiovascular disease in dialysis patients," *Amer. J. Kidney Diseases*, vol. 45, no. 3, pp. S1–S153, Apr. 2005.

[42] M. Z. I. Chowdhury and T. C. Turin, "Variable selection strategies and its importance in clinical prediction modelling," *Family Med. Community Health*, vol. 8, no. 1, Feb. 2020, Art. no. e000262, doi: 10.1136/FMCH-2019-000262.

[43] E. W. Steyerberg and Y. Vergouwe, "Towards better clinical prediction models: Seven steps for development and an ABCD for validation," *Eur. Heart J.*, vol. 35, no. 29, pp. 1925–1931, Aug. 2014, doi: 10.1093/eurheartj/ehu207.

[44] G. Mehrpoor, M. M. Azimzadeh, and A. Monfared, "Data mining: A novel outlook to explore knowledge in health and medical sciences," *Int. J. Travel Med. Global Health*, vol. 2, no. 2, pp. 87–90, 2014. [Online]. Available: http://www.ijtmgh.com/article_33289.html

[45] C. Seger, "An investigation of categorical variable encoding techniques in machine learning: Binary versus one-hot and feature hashing," School Elect. Eng. Comput. Sci. (EECS), KTH, Stockholm, Sweden, Tech. Rep. 2018:596, 2018, p. 34.

[46] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Dec. 2002.

[47] M. Beckmann, N. F. F. Ebecken, and B. S. L. P. de Lima, "A KNN undersampling approach for data balancing," *J. Intell. Learn. Syst. Appl.*, vol. 7, no. 4, p. 104, 2015.

[48] B. R. Manju and A. R. Nair, "Classification of cardiac arrhythmia of 12 lead ECG using combination of SMOTEENN, XGBoost and machine learning algorithms," in *Proc. 9th Int. Symp. Embedded Comput. Syst. Design (ISED)*, Dec. 2019, pp. 1–7, doi: 10.1109/ISED48680.2019.9096244.

[49] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics," *Inf. Sci.*, vol. 250, pp. 113–141, Nov. 2013, doi: 10.1016/j.ins.2013.07.007.

[50] B. Krawczyk, "Learning from imbalanced data: Open challenges and future directions," *Prog. Artif. Intell.*, vol. 5, no. 4, pp. 221–232, 2016, doi: 10.1007/S13748-016-0094-0.

[51] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A comparative analysis of logistic regression, random forest and KNN models for the text classification," *Augmented Hum. Res.*, vol. 5, no. 1, p. 12, Dec. 2020, doi: 10.1007/S41133-020-00032-0.

[52] G. M. Fitzmaurice and N. M. Laird, "Multivariate analysis: Discrete variables (logistic regression)," in *International Encyclopedia of the Social & Behavioral Sciences*, N. J. Smelser and B. S. Baltes, Eds. Oxford, U.K.: Pergamon, 2001, pp. 10221–10228.

[53] K. Kirasich, T. Smith, and B. Sadler, "Random forest vs logistic regression: Binary classification for heterogeneous datasets," *SMU Data Sci. Rev.*, vol. 1, no. 3, p. 9, 2018.

[54] A. Geron, *Hands-on Machine Learning With Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. Newton, MA, USA: O'Reilly Media, Inc., 2019.

[55] J. Wu, Y. Li, and Y. Ma, "Comparison of XGBoost and the neural network model on the class-balanced datasets," in *Proc. IEEE 3rd Int. Conf. Frontiers Technol. Inf. Comput. (ICFTIC)*, Nov. 2021, pp. 457–461, doi: 10.1109/ICFTIC54370.2021.9647373.

[56] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[57] S. Yadav and S. Shukla, "Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification," in *Proc. IEEE 6th Int. Conf. Adv. Comput. (IACC)*, Feb. 2016, pp. 78–83, doi: 10.1109/IACC.2016.25.

[58] A. Vabalas, E. Gowen, E. Poliakoff, and A. J. Casson, "Machine learning algorithm validation with a limited sample size," *PLoS ONE*, vol. 14, no. 11, Nov. 2019, Art. no. e0224365, doi: 10.1371/journal.pone.0224365.

[59] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology*, S. Ranganathan, M. Gribskov, K. Nakai, and C. B. Schönbach, Eds. Oxford, U.K.: Elsevier, 2019, pp. 542–545.

[60] G. S. Handelman, H. K. Kok, R. V. Chandra, A. H. Razavi, S. Huang, M. Brooks, M. J. Lee, and H. Asadi, "Peering into the black box of artificial intelligence: Evaluation metrics of machine learning methods," *Amer. J. Roentgenol.*, vol. 212, no. 1, pp. 38–43, Jan. 2019, doi: 10.2214/AJR.18.20224.

[61] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. IJCAI*, 1995, vol. 14, no. 2, pp. 1137–1145.

[62] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statist. Surv.*, vol. 4, pp. 40–79, Jan. 2010, doi: 10.1214/09-SS054.

[63] K. R. Lakshmi, Y. Nagesh, and M. V. Krishna, "Performance comparison of three data mining techniques for predicting kidney dialysis survivability," *Int. J. Adv. Eng. Technol.*, vol. 7, no. 1, p. 242, 2014.

[64] J. G. Moreno-Torres, J. A. Saez, and F. Herrera, "Study on the impact of partition-induced dataset shift on k-fold cross-validation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 8, pp. 1304–1312, Aug. 2012, doi: 10.1109/TNNLS.2012.2199516.

[65] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart disease identification method using machine learning classification in E-Healthcare," *IEEE Access*, vol. 8, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[66] S. A. Hicks, I. Strümke, V. Thambawita, M. Hammou, M. A. Riegler, P. Halvorsen, and S. Parasa, "On evaluation metrics for medical applications of artificial intelligence," *Sci. Rep.*, vol. 12, no. 1, p. 5979, Dec. 2022, doi: 10.1038/S41598-022-09954-8.

[67] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Int. J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.

[68] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mob. Inf. Syst.*, vol. 2018, Dec. 2018, Art. no. 3860146, doi: 10.1155/2018/3860146.

[69] R. Kumar and A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers," *Indian Pediatrics*, vol. 48, no. 4, pp. 277–287, 2011, doi: 10.1007/S13312-011-0055-4.

[70] C. J. Chiew, N. Liu, T. H. Wong, Y. E. Sim, and H. R. Abdullah, "Utilizing machine learning methods for preoperative prediction of postsurgical mortality and intensive care unit admission," *Ann. Surg.*, vol. 272, no. 6, pp. 1133–1139, Dec. 2020, doi: 10.1097/SLA.0000000000003297.

[71] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, pp. 837–845, Sep. 1988, doi: 10.2307/2531595.

[72] F. Dustin. (2019). *Jetson Nano Brings AI Computing to Everyone*. Nvidia. Accessed: Sep. 23, 2021. [Online]. Available: https://developer.nvidia.com/blog/jetson-nano-ai-computing/

[73] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Müller, "pROC: An open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinf.*, vol. 12, no. 1, p. 77, Mar. 2011, doi: 10.1186/1471-2105-12-77.

[74] M. Li, X. Fu, and D. Li, "Diabetes prediction based on XGBoost algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, p. 72093, 2020, doi: 10.1088/1757-899x/768/7/072093.

[75] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, Apr. 1982, doi: 10.1148/RADIOLOGY.143.1.7063747.
[76] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine," *Clin. Chem.*, vol. 39, no. 4, pp. 561–577, 1993.
[77] B. Snider, E. A. McBean, J. Yawney, S. A. Gadsden, and B. Patel, "Identification of variable importance for predictions of mortality from COVID-19 using AI models for Ontario, Canada," *Frontiers Public Health*, vol. 9, p. 798, Jun. 2021, doi: 10.3389/FPUBH.2021.675766.
[78] D. Zhang, L. Qian, B. Mao, C. Huang, B. Huang, and Y. Si, "A data-driven design for fault detection of wind turbines using random forests and XGboost," *IEEE Access*, vol. 6, pp. 21020–21031, 2018, doi: 10.1109/ACCESS.2018.2818678.
[79] O. Niel and P. Bastard, "Artificial intelligence in nephrology: Core concepts, clinical applications, and perspectives," *Amer. J. Kidney Diseases*, vol. 74, no. 6, pp. 803–810, Dec. 2019, doi: 10.1053/j.ajkd.2019.05.020.
[80] M. Hueso, A. Vellido, N. Montero, C. Barbieri, R. Ramos, M. Angoso, J. M. Cruzado, and A. Jonsson, "Artificial intelligence for the artificial kidney: Pointers to the future of a personalized hemodialysis therapy," *Kidney Diseases*, vol. 4, no. 1, pp. 1–9, 2018, doi: 10.1159/000486394.
[81] S. Ansari, S. Molaei, K. Oldham, M. Heung, K. R. Ward, and K. Najarian, "An extended Kalman filter with inequality constraints for real-time detection of intradialytic hypotension," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2227–2230, doi: 10.1109/EMBC.2017.8037297.
[82] I. Odun-Ayo, R. Goddy-Worlu, V. Geteloma, and E. Grant, "A systematic mapping study of cloud, fog, and edge/mobile devices management, hierarchy models and business models," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 4, no. 2, pp. 91–101, 2019.
[83] H. Calderon-Gomez, L. Mendoza-Pitti, M. Vargas-Lombardo, J. M. Gomez-Pulido, J. L. Castillo-Sequera, J. Sanz-Moreno, and G. Sencion, "Telemonitoring system for infectious disease prediction in elderly people based on a novel microservice architecture," *IEEE Access*, vol. 8, pp. 118340–118354, 2020, doi: 10.1109/ACCESS.2020.3005638.

LUIS MENDOZA-PITTÍ was born in Panama, in 1992. He received the B.E. degree in engineering systems and computing from the Technological University of Panama, Chiriqui, Panama, in 2015. He is currently pursuing the Ph.D. degree in information and knowledge engineering with the University of Alcalá, Alcalá de Henares, Spain. From 2013 to 2016, he was a Student Research Assistant with the Technological University of Panama. His current research interests include smart health, cloud computing, software engineering, and machine learning models.

JOSÉ MANUEL GÓMEZ-PULIDO was born in Cáceres, Spain. He received the B.S. and M.S. degrees in telecommunications engineering from the University Polytechnic of Madrid, Spain, in 1988 and 1994, respectively, and the Ph.D. degree in telecommunication from the University of Alcalá, Madrid, in 2004. In 1992, he was an Assistant Researcher and a Professor with the University Polytechnic of Madrid. From 1993 to 2004, he was with the Signal Theory and Communications Department, University of Alcalá. He is currently an Assistant Professor with the Computational Science Department, University of Alcalá. He has participated in more than 24 research projects several research projects with Spanish and European companies, related with radio propagation in mobile communications, signal processing in radar, ultrasound and laser applications, EMI–EMC design in on-board satellite instrumentation, radio propagation materials proprieties, and numerical technics applied in telecommunication. His research interests include wireless sensors networks, smart grids, and new technologies in the data mining and artificial intelligence world applied to the ICT.

MIGUEL VARGAS-LOMBARDO (Senior Member, IEEE) received the Ph.D. degree from the Polytechnic University of Madrid. He is currently a member of the National Research System (SNI-SENACYT), Panama. He is also a Professor and a Researcher with the Technological University of Panama.

JUAN A. GÓMEZ-PULIDO (Senior Member, IEEE) received the Ph.D. degree in physics (electronics specialty) from Complutense University, Madrid, Spain, in 1993. He is currently a Full Professor of computer organization and design of processors with the Department of Technology of Computers and Communications, Universidad de Extremadura, Spain. He has authored or coauthored 70 ISI journals, almost 100 book chapters, and 300 peer-reviewed conference proceedings. His main research interests include machine learning applied to big-data analysis, reconfigurable and embedded computing based on FPGA devices, bioinformatics, optimization, and evolutionary computing.

MARÍA-LUZ POLO-LUQUE received the B.S. degree in nursing from the Fundación Jiménez Díaz of Madrid, Spain, in 1979, the M.S. degree in diététique superieure from the Henri Poincaré University of Nancy, France, the M.S. degree in nursing sciences from the University of Alicante, Spain, the M.S. degree in gerontology from the University of Santiago de Compostela, Spain, and the Ph.D. degree in health sciences from the University of Alcalá, Spain, in 2011. She has conducted research at the SU.VI.MAX Center, France, regarding methods of measuring food consumption and data analysis. She has had an extensive career as a Nurse at the Transplant Unit, Hospital Fundación Jiménez Díaz, Hospital Necker, Paris, France. She is currently a Professor of health sciences at the University of Alcalá and participates in the UAH Cooperation Group (ACORDES), which aims to promote the resilience of the most vulnerable communities through social and gender equity. Her main research interests include geriatric nursing, nursing cares in highly complex chronic patients, nursing research in health and social problems, and health computing and intelligent systems in nursing cares.

DIEGO RODRÍGUEZ-PUYOL was born in Madrid. He received the Ph.D. degree from Autonoma University, Madrid. He is currently a Full Professor of medicine at the University of Alcalá and the Chief of nephrology at the University Hospital "Príncipe de Asturias." He performed his medicine studies at Complutense University. He is a Nephrologist with a significant clinical, teaching, and translational research expertise. His work has been devoted to the study of renal disease pathogenesis and progression, synthesis of new drugs useful for treating renal diseases, and big data applied to the solution of clinical problems. He has published more than 170 original manuscripts in different international publications.