Article

Machine Learning Techniques for Chronic Kidney Disease Risk Prediction

Elias Dritsas [ID] **and Maria Trigka** *[ID]

Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece
* Correspondence: trigka@ceid.upatras.gr

Abstract: Chronic kidney disease (CKD) is a condition characterized by progressive loss of kidney function over time. It describes a clinical entity that causes kidney damage and affects the general health of the human body. Improper diagnosis and treatment of the disease can eventually lead to end-stage renal disease and ultimately lead to the patient's death. Machine Learning (ML) techniques have acquired an important role in disease prediction and are a useful tool in the field of medical science. In the present research work, we aim to build efficient tools for predicting CKD occurrence, following an approach which exploits ML techniques. More specifically, first, we apply class balancing in order to tackle the non-uniform distribution of the instances in the two classes, then features ranking and analysis are performed, and finally, several ML models are trained and evaluated based on various performance metrics. The derived results highlighted the Rotation Forest (RotF), which prevailed in relation to compared models with an Area Under the Curve (AUC) of 100%, Precision, Recall, F-Measure and Accuracy equal to 99.2%.

Keywords: healthcare; chronic kidney disease; machine learning; prediction; data analysis

1. Introduction

The human body has two kidneys located at the back of the peritoneal cavity, which are vital organs necessary for its proper functioning. The main function of the kidneys is to regulate the balance of salt, water and other ions and trace elements in the human body, such as calcium, phosphorus, magnesium, potassium, chlorine and acids. At the same time, the kidneys secrete hormones such as erythropoietin, vitamin D and renin. More specifically, erythropoietin stimulates the production and maturation of red blood cells in the bone marrow, while vitamin D regulates calcium and phosphorus in the body, bone structure and many other actions. The kidneys are also the site of the action of hormones that are responsible for regulating blood pressure, fluid balance or bone metabolism and vascular calcifications. Finally, the kidneys eliminate all the useless products of metabolism, as well as drugs and other toxins that enter the body [1].

Diabetes and high blood pressure are the two main causes of chronic kidney disease. Diabetes is characterized by high blood sugar levels, causing damage to the kidneys and heart, blood vessels and eyes. Moreover, poor control of high blood pressure can be a major cause of heart attack, stroke and chronic kidney disease. Other conditions that affect the kidneys are glomerulonephritis, hereditary diseases, dysplasia, kidney stones, tumours, recurrent urinary tract infections, metabolic diseases, obesity and age [2,3].

CKD is a silent disease, as most sufferers have no symptoms until kidney function drops to 15–20% of normal [4]. The main symptoms in the advanced stage of CKD are the feeling of fatigue and lack of energy, concentration problems, decreased appetite, sleep problems, muscle cramps at night, swelling in the legs and ankles, swelling around the eyes, dry skin with intense itching and frequent urination, especially at night [5].

The most important and effective parameter for the evaluation of renal function is the glomerular filtration rate (GFR), which practically evaluates the ability of the kidney

to filter blood. The glomerular filtration rate is the best measure of renal function and is usually assessed (eGFR) by the results of a creatinine blood test. The eGFR value refers to milliliters per minute per 1.73 m^2 (mL/min/1.73 m^2). Renal function can be classified into 5 stages according to eGFR, as shown in Table 1 [6].

Table 1. Five stages of chronic kidney disease.

Stage of CKD	Description	GFR (mL/min/1.73 m^2)
Stage 1	Normal	$\geq$90
Stage 2	Mild CKD	60–89
Stage 3	Moderate CKD	30–59
Stage 4	Severe CKD	15–29
Stage 5	End Stage CKD	<15

Early diagnosis and treatment of CKD is a serious challenge for the medical community. The treating physician (nephrologist) is called on the one hand to slow down the progression of the disease to more advanced stages, and if possible, to suspend it, and on the other hand, to treat the above-mentioned systemic manifestations [7].

The advances in sensor networks, communication technologies, data science and statistical processing have rendered ML techniques as important tools in various health-oriented applications, such as in the early diagnosis of several chronic conditions, the Internet of Things (IoT)-based pervasive (assisted) living environments (smart homes) for elderly fall detection [8], etc. Concerning the diseases, some characteristic ones are the following: Diabetes [9–11], Hypertension [12], Cholesterol [13,14], COVID-19 [15], Chronic Obstructive Pulmonary Disease (COPD) [16], Stroke [17], Cardiovascular Diseases (CVDs) [18], Acute Liver Failure (ALF) [19], Acute Lymphoblastic Leukemia [20], Sleep Disorders [21], Hepatitis [22], Cancer [23], Metabolic Syndrome [24], etc.

In the current research work, a Machine Learning-based approach will be presented for CKD disease. The main contributions of the adopted methodology are the following:

- A data preprocessing step that exploits the Synthetic Minority Oversampling Technique (SMOTE), which is essential to ensure that the dataset instances are distributed in a balanced way and, thus, designs effective classification models to predict the risk for CKD occurrence.
- A features analysis, which includes three specific sub-steps: (i) numerical attributes statistical description, (ii) order of importance measurement by employing three different methods, and (iii) capturing nominal features frequency of occurrence in tabular form.
- A comparative evaluation of various models' performance is presented considering the most common metrics, such as Precision, Recall, F-Measure, Accuracy and AUC.
- A performance evaluation is demonstrated, where all models demonstrated exceptionally high outcomes, with Rotation Forest achieving the highest results in all metrics, thus constituting the main suggestion of this analysis.

The rest of the sections of the work are structured as follows. In Section 2, we present related works that exploit ML under the CKD health condition. Besides, in Section 3, we describe the dataset and analyze the adopted methodology. Furthermore, in Section 4, we present and discuss the research outcomes. Finally, in Section 5, we conclude the paper and set future directions.

2. Related Work

Nowadays, the development of tools and methods for monitoring and predicting various diseases has gained researchers' and clinicians' interest, focusing on those which commonly occur in human life. In this section, we will discuss recent studies that use ML techniques for CKD risk prediction and methods for processing small datasets.

Firstly, in [25], the authors' research was based on clinical and blood biochemical measurements from 551 patients who suffered from proteinuria. For their purpose, several predictive models were compared, including random forest (RF), extreme gradient boosting (XGBoost), logistic regression (LR), elastic net (ElasNet), lasso and ridge regression, k-nearest neighbour (k-NN), support vector machine (SVM) and artificial neural network (ANN). The superior predictive performance was achieved by the models ElasNet, lasso, ridge and logistic regression, reaching a mean value of AUC and precision above 0.87 and 0.8, respectively. Moreover, LR was the first in rank, reaching an AUC of 0.873, with a recall and specificity of 0.83 and 0.82, respectively. The highest recall was attained by ElasNet (0.85), while the highest specificity (0.83) was performed by XGBoost.

In [26], the authors exploited SVM, AdaBoost, linear discriminant analysis (LDA), and gradient boosting (GBoost) algorithms in order to implement highly accurate models for CKD prediction. These models' performance was evaluated considering a dataset derived from the UCI machine learning repository. The gradient boosting classifier achieved the highest accuracy of 99.80%.

The authors in [27] focused on the [28] dataset. LR, Decision Tree (DT), and k-NN algorithms were used in order to train three different models for CKD prediction. The LR achieved better accuracy (97%) in comparison with DT (96.25%) and k-NN (71.25%). Similarly, the [28] dataset is used in [29] research work. The authors examined the performance of the Naïve Bayes (NB), RF and LR models for the risk prediction of CKD that achieved an accuracy of 93.9%, 98.88% and 94.76%, respectively.

Moreover, in [30], the authors used 455 patients' data from the UCI Machine Learning Repository and the real-time dataset from Khulna City Medical College to propose a system for CKD risk prediction. The data were used to train and test RF and ANN using 10-fold cross-validation. The accuracy achieved by the RF and ANN is 97.12% and 94.5%, respectively.

Besides, the research study in [31] was based on a CKD dataset taken from the UCI repository to train and test several classifiers such as ANN, C5.0, chi-square automatic interaction detector, LR, linear SVM with penalty L1 and L2, and random tree (RT). The linear SVM with penalty L2 reached the highest accuracy of 98.86% under SMOTE and all features as input to the ML models. Combining SMOTE with the lasso method for feature selection, the linear SVM achieved a similarly high accuracy of 98.46%. Finally, a deep neural network was applied in the same dataset, attaining the highest accuracy of 99.6%.

The experiments in [32] were conducted on the CKD dataset consisting of 25 attributes and acquired by the UCI Machine Learning Repository. Three ML models, RF, DT and SVM, were selected for the diagnosis of CKD, reaching a prediction accuracy of 99.16%, 94.16% and 98.3%, respectively.

Moreover, in [33], the authors considered a dataset of 26 attributes relevant to CKD. They combined the ANN classifier with four feature-based algorithms: Extra Tree, Pearson correlation, lasso model and chi-square. The highest accuracy (99.98%) was performed by the ANN ensemble with the lasso model.

Furthermore, the research study in [34] used the extra-trees (ExTrees) classifier, AdaBoost, k-NN, GBoost and XGBoost, DT, gaussian Naïve Bayes (NB) and RF. According to the results, k-NN and ExTrees classifiers achieved the best performance with an accuracy of 99% and 98%, respectively.

In addition, in [35], the authors considered a crucial problem in ML that concerns the handling of small medical datasets. They enhanced the regression analysis based on ANNs by introducing additional elements into the formula for calculating the output signal of the existing radial basis function-based (RBF) input-doubling method. Similarly, in [36], the authors designed a new input doubling method based on the classical iterative RBF neural network. The Mean Absolute Error and Root Mean Squared Error were used to validate the highest accuracy of the proposed method by experimenting with a small medical dataset.

A novel approach based on generative adversarial network (GAN) for data augmentation with improved disease classification is applied in [37]. The authors performed their experiments on the NIH chest X-ray image dataset, and the test accuracy of the convolutional neural network (CNN) model is 60.3% compared to the 65.3% test accuracy of the online GAN-augmented CNN model. Finally, in [38], a development of the non-iterative supervised learning predictor is presented based on the Ito decomposition and neural-like structure successive geometric transformations model (SGTM) for managing medical insurance data.

3. Materials and Methods

3.1. Dataset Description

In this research study, we exploited the dataset on [28]. The raw dataset consists of 400 instances represented by 13 input features and 1 for the target class. The features' description is the following:

- **Diastolic Blood Pressure** (Bp - mmHg) [39]: This feature shows the participator's diastolic blood pressure.
- **Specific Gravity** (Sg) [40]: This feature captures the participator's specific gravity value.
- **Albumin** (Al) [41]: This attribute captures the participator's albumin level. It has three categories (72.25% normal, 21.5% above normal and 6.25% well above normal).
- **Glucose** (Su) [42]: This attribute denotes the participator's glucose level. It has three categories (88% normal, 8% above normal and 4% well above normal).
- **Red Blood Cell** (Rbc) [43]: This attribute captures whether the participator's Red Blood Cell is normal or not. It has two categories (88.25% normal and 11.75% abnormal).
- **Blood Urea** (Bu - mmol/L) [44]: This feature captures the amount of urea found in the participant's blood. Blood Urea is measured in millimoles per liter (mmol/L).
- **Serum Creatinine** (Sc - mg/dL) [45]: This feature measures the amount of serum creatinine found in the participant's blood. Serum creatinine is reported as milligrams of creatinine to a deciliter of blood (mg/dL).
- **Sodium** (Sod - mEq/L) [46]: This feature measures the amount of sodium found in the participant's blood. Sodium is a type of electrolyte and is reported as milliequivalents per liter (mEq/L).
- **Potassium** (Pot - mmol/L) [47]: This feature measures the amount of potassium found in the participant's blood and is reported as millimoles per liter (mmol/L).
- **Hemoglobin** (Hemo - gm/dL) [48]: This feature measures the amount of hemoglobin found in the participant's blood and is reported as grams per deciliter (gm/dL).
- **White Blood Cell Count** (Wbcc) [49]: This feature measures the number of white cells in the participant's blood and is reported as Wbc per microliter.
- **Red Blood Cell Count** (Rbcc) [43]: This feature measures the number of red blood cells in the participant's blood and is reported as a million red blood cells per microliter (mcL) of blood.
- **Hypertension** (Htn) [50]: This attribute refers to whether the participant has hypertension or not. A total of 36.75% of participants have hypertension.
- **Chronic Kidney Disease** (CKD): This feature denotes whether the participant suffers from CKD or not. A total of 62.5% of participants have been diagnosed with CKD.

All features are numeric except Al, Su, Rbc, Htn and CKD, which are nominal.

3.2. Chronic Kidney Disease Risk Prediction

In this section, we will focus on class balancing and features importance evaluation in the balanced data. We will also make a brief analysis of the nominal features concerning the CKD class. Moreover, we will describe the models and performance metrics, which will be considered in the experiments.

3.2.1. Data Preprocessing

As for the current dataset, we employed SMOTE [51] to create synthetic data on minority class, i.e., Non-CKD, using a k-NN classifier with k = 5. The instances in the Non-CKD class are oversampled such that the number of instances in both classes is balanced (i.e., 50–50%). After data balancing, in Table 2, we present a statistical description of the numeric features, namely, minimum (Min), maximum (Max), mean and standard deviation.

Table 2. Statistical description of the balanced data.

Feature	Min	Max	Mean $\pm$ std
Hemo	3.1	17.8	13.04 ± 2.68
Sg	1.005	1.025	1.019 ± 0.005
Rbcc	2.1	8	4.84 ± 0.82
Bu	1.5	391	52.59 ± 45.34
Sod	4.5	163	138.44 ± 8.64
Sc	0.4	76	2.65 ± 5.09
Bp	50	180	75.4 ± 12.6
Wbcc	2200	26,400	8310.7 ± 2394.2
Pot	2.5	47	4.56 ± 2.53

3.2.2. Features Analysis

Each record in the dataset is captured by a features vector $\mathbf{x} = (x_1, x_2, x_3, \ldots, x_M)^T$, where $M = 13$ is the features' size. In order to measure the contribution of a feature in the desired class, three ranking methods were selected, i.e., Pearson correlation coefficient (CC), Gain Ratio (GR) and Random Forest. Initially, we evaluate the strength of a feature in predicting the CKD class via Pearson correlation coefficient [52]. Next, we measured GR of the feature x_j [53] based on the formula $GR(x_j) = \frac{H(c) - H(c|x_j))}{H(x_j)}$, where $H(c)$, $H(c|x_j)$ and $H(x_j)$ are the entropy of the class, the conditional entropy of the class given the feature j, x_j and the entropy of the feature x_j, respectively. Random Forest measures, by employing Gini impurity, the ability of a candidate feature in the forest of trees to create the optimal split of the two classes instances [54].

In Table 3, we demonstrate the ranking outcomes of the selected methods in the balanced dataset. Focusing on the Pearson correlation, the highest but moderate association of 0.763 is captured with Hemoglobin, which is a biochemical measure that relates to anemia and CKD progression [55]. Moreover, moderate associations of rank 0.699, 0.645 and 0.621 are noted with the Blood Glucose, Hypertension, and Red blood cells count. In addition, the low association is demonstrated with the rest features, such as Sodium, Red blood cells level and White blood cells count. Finally, the target class records no association of 0.092 with Potassium. The Hemoglobin feature is also first in ranking by Random Forest, while this risk factor is third in order by Gain Ratio. Moreover, a variety in the order of importance is captured among the methods. Since all features are important indicators for kidney operation and thus CKD control by physicians, the models' training and assessment will exploit all of them.

Table 3. Features' importance evaluation (balanced dataset).

Pearson CC		Gain Ratio		Random Forest	
Feature	**Ranking**	**Feature**	**Ranking**	**Feature**	**Ranking**
Hemo	0.763	Sc	0.532	Hemo	0.449
Sg	0.699	Htn	0.441	Rbcc	0.439
Htn	0.645	Hemo	0.381	Sc	0.429
Rbcc	0.621	Sg	0.338	Sg	0.401
Al	0.506	Rbcc	0.337	Sod	0.388
Bu	0.419	Bp	0.295	Pot	0.374
Sod	0.387	Al	0.287	Bp	0.309
Sc	0.334	Bu	0.270	Bu	0.292
Rbc	0.322	Rbc	0.225	Htn	0.277
Bp	0.321	Su	0.190	Wbcc	0.232
Su	0.317	Sod	0.170	Al	0.211
Wbcc	0.207	Wbcc	0.141	Su	0.088
Pot	0.092	Pot	0.136	Rbc	0.086

In Table 4, we isolate nominal features and present the distribution of the instances in both classes in terms of the values of the features. Healthy participants (those who belong to Non-CKD class) have normal levels of albumin, glucose and red blood cells, and they are not hypertensive. Of the total participants, 27.8% have been diagnosed with CKD and have normal albumin levels, and 22.2% have above and well above normal albumin values. Moreover, 40.4% of participants have CKD with normal glucose levels, and 29.4% of them are CKD patients and hypertensive. Finally, the Red blood cells level is normal in 40.6% of them.

Table 4. Nominal features' values in terms of the CKD class (balanced dataset).

Albumin	**CKD = No**	**CKD = Yes**
Above normal	0.00%	17.20%
Well above normal	0.00%	5.00%
Normal	50.00%	27.80%
Glucose	**CKD = No**	**CKD = Yes**
Above normal	0.00%	6.40%
Normal	50.00%	40.40%
Well above normal	0.00%	3.20%
Hypertension	**CKD = No**	**CKD = Yes**
No	50.00%	20.60%
Yes	0.00%	29.40%
Red blood cell	**CKD = No**	**CKD = Yes**
Abnormal	0.00%	9.40%
Normal	50.00%	40.60%

3.3. Machine Learning Models

In this subsection, we will make a brief presentation of the models that will be considered in the risk prediction framework for CKD occurrence. To this end, a variety of

classifiers are utilized in order to evaluate their prediction performance. More specifically, Bayesian networks (BayesNet), Naive Bayes, SVM, LR, ANN, k-NN, J48, Logistic Model Tree (LMT), Reduced Error Pruning Tree (RepTree) Rotation Forest, Decision Tree, Random Forest, Random Tree, AdaBoostM1, Stochastic Gradient Descent (SGD), Stacking and Soft Voting classification methods will be outlined.

3.3.1. Naive Bayes

Naive Bayes classifier [56], following a Bayesian probabilistic model, assigns a subject i with attributes vector $\mathbf{x}_i$ to that class c for which the posterior probability $P(c|x_{i1}, \ldots, x_{iM})$ is maximized.

3.3.2. Bayesian Network

A Bayesian network [57] is a probabilistic graphical model that follows the structure of a directed acyclic graph (DAG). Its nodes are captured as random variables, and the edges demonstrate the conditional (in)dependencies among them.

3.3.3. Support Vector Machine

Support Vector Machine [58] finds the proper boundary that can optimally split subjects into two classes. An instance with an unknown class can be optimally classified using one of the following kernel functions, i.e., linear, polynomial, radial basis or quadratic.

3.3.4. Logistic Regression

Logistic regression [59] is a well-established supervised learning algorithm in the medical community. Logistic regression predicts the probability of the class output (a target categorical variable with values of Yes, No or 0, 1) using a set of independent features. Assuming that p is the probability of a subject being a member of the CKD class, then $1 - p$ is the probability of a subject and is a member of the Non-CKD.

3.3.5. Artificial Neural Network

Multilayer Perceptron (MLP) is a fully connected Neural Network (NN) [60], consisting of an input, an output and a hidden layer. The nodes in the input layer take $\mathbf{x}_i$ and forward it for further processing in the hidden layer that processes the data and passes it to the output layer. Apart from the input layer nodes, every other node in the MLP uses a nonlinear (such as sigmoid) activation function that takes real values as the input and converts them to numbers between 0 and 1. The MLP networks update the weights via backpropagation learning.

3.3.6. k-Nearest Neighbors

The k-Nearest Neighbors classifier measures the distance between an unlabeled instance and every other training instance [61] and designates it into the class where most of its k proximal neighbors originate.

3.3.7. J48

J48 [62] follows a top-down recursive strategy known as divide-and-conquer, and uses information gain measure to choose the attribute at each stage.

3.3.8. Logistic Model Tree

A logistic model tree [63] follows the structure of a standard decision tree with LR functions at the leaves. It builds a single tree consisting of binary splits on numeric attributes, multiple-way splits on nominal ones and LR models at the leaves.

3.3.9. Random Forest

Random Forest is an ensemble of decision trees. It considers the Information Gain or Gini index to find the best subset of features. It classifies an instance by applying majority voting on the outputs of several decision trees [64].

3.3.10. Random Tree

Random Tree [65] ensembles several decision trees. It partitions recursively the training data into segments with similar output features' values and finds the best partition assessing the impurity index.

3.3.11. Reduced Error Pruning Tree

Reduced Error Pruning Tree [66] is a quick learner that uses information variance as the splitting criterion to build a tree, and prunes it using reduced-error pruning.

3.3.12. Rotation Forest

The Rotation Forest [67] uses as a base classifier the decision trees. Prior to training, it applies a rotation transformation matrix to the training data. The feature set is randomly split into subsets, and it is applied principal component analysis (PCA) to create a new feature set for every tree in the ensemble. In this study, RotF uses the J48 decision tree.

3.3.13. AdaBoostM1

The AdaBoostM1 is an adaptive method that combines via weighted majority voting on the predictions of a sequence of L weak classifiers denoted as $G_l(\mathbf{x}_i)$, where $l = 1, 2, \ldots, L$. Assuming a training set consisting of N samples, at each boosting step r, each sample $\mathbf{x}_i$ is weighted, assuming that the initial weights at $r = 1$ are uniform, i.e., $w_{11} = w_{21} = \ldots, w_{N1} = 1/N$. The weights are determined using the error value. The weight of an instance is increased when the previous classification is incorrect, otherwise it is decreased [68]. The higher the error, the higher weight is assigned to the sample. The process is repeated until the error remains constant. The final prediction is derived by

$$G(\mathbf{x}_i) = sign(\sum_{l=1}^{L} \alpha_l G_l(\mathbf{x}_i))(\{-1, +1\}). \tag{1}$$

In (1), the coefficients α_l are estimated based on the classification error and weigh the corresponding $G_l(\mathbf{x}_i)$ giving a higher contribution to the classifiers that are more accurate.

3.3.14. Stochastic Gradient Descent

Stochastic gradient descent [69] is a method to efficiently fit linear classification models such as linear SVM and LR for optimizing an objective function.

3.3.15. Ensemble Learning

Ensemble learning is utilized in machine learning to obtain more accurate predictions than individual models by combining the outputs of several single classification models. Voting and Stacking are the two methods which will be used in this study. In the case of Voting, we focus on the soft method, which averages the probabilities of the single models in each class and designates a test instance to the class with the highest probability [70]. Stacking feeds the outputs of the base models, namely the predicted class labels, as input features to train a meta-classifier, which takes on to predict the final class label [71]. In Figure 1, we demonstrate the above-mentioned ensemble methods, which will be considered in the evaluation part of the study.
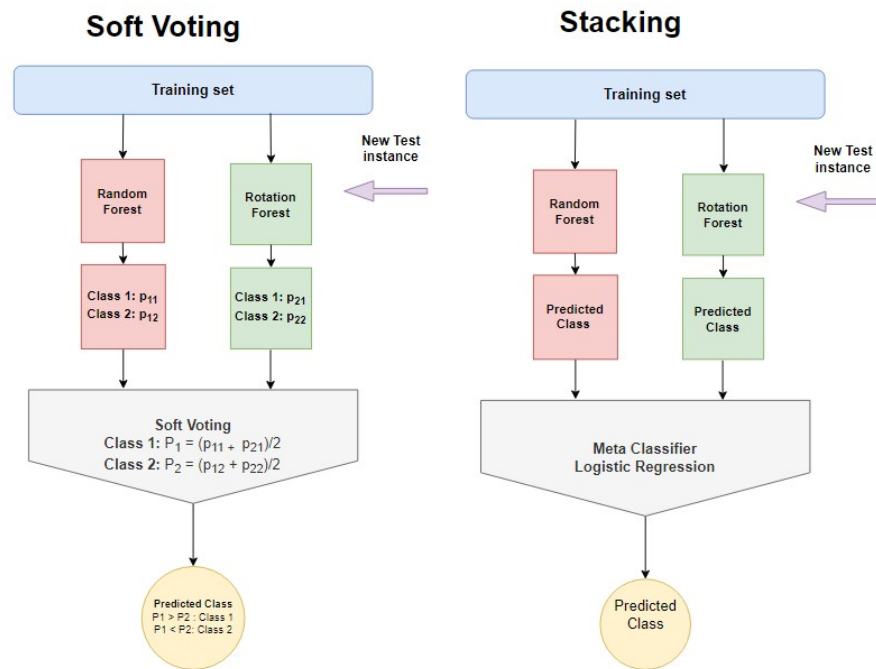
Figure 1. Ensemble learners: Soft voting and stacking.

3.4. Evaluation Metrics

In order to assess the ML models' performance, we consider the most common metrics in the relevant literature, such as precision, recall, accuracy, F-Measure and AUC. Each metric will help us to evaluate the models [72].

Specifically, accuracy summarizes the performance of the classification task and measures the number of correctly predicted instances out of all the data instances. Recall captures the proportion of instances who suffered from CKD and were correctly categorized as CKD, concerning all CKD instances. Precision indicates how many of those who were diagnosed with CKD belong to this class. F-measure is the harmonic mean of the precision and recall and summarizes the predictive accuracy of a model. The aforementioned metrics are defined as follows

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad\qquad \text{Recall} = \frac{TP}{TP + FN} \qquad (2)$$

$$\text{F} - \text{Measure} = 2\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \qquad \text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \qquad (3)$$

where TP, TN, FP and FN stand for the true positive, true negative, false positive (FP) and false-negative, respectively.

Finally, in order to assess the ability of a model to correctly separate the distribution of CKD from Non-CKD subjects, the AUC is utilized. The upper optimal limit of the AUC metric is 1 while the lowest value is 0.

4. Results and Discussion

4.1. Experiments Setup

We based the evaluation of our ML models on the Weka tool [73], and the experiments were conducted in a computing machine, which has the following specifications: 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz 2.70 GHz, 16 GB, Windows 11 Home, 64-bit Operating System and x64-based processor. The experimental results were derived by applying 10-fold cross-validation to measure the models' efficiency in the balanced dataset of 500 instances after SMOTE. Finally, in Table 5, we illustrate the optimal settings of the ML models parameters with which we experimented.

Table 5. Machine learning models' settings.

Models	Parameters
BayesNet	estimator: simpleEstimator searchAlgorithm: K2 useADTree: False
NB	useKernelEstimator: False useSupervisedDiscretization: True
SVM	eps = 0.001 gamma = 0.0 kernel type: linear loss = 0.1
LR	ridge = 10^{-8} useConjugateGradientDescent: False
ANN	hidden layers: 'a' learning rate: 0.3 momentum: 0.2 training time: 500
k-NN	k = 1 Search Algorithm: LinearNNSearch with Euclidean
J48	reducedErrorPruning: False saveInstanceData: False subtreeRaising: True
LMT	errorOnProbabilities: False fastRegression: True numInstances = 15 useAIC: False
RF	maxDepth = 0 numIterations = 100 numFeatures = 0
RT	maxDepth = 0 minNum = 1.0 minVarianceProp = 0.001
DT (RepTree)	maxDepth = −1 minNum = 2.0 minVarianceProp = 0.001
RotF	classifier: J48 numberOfGroups: False projectionFilter: PrincipalComponents
AdaBoostM1	classifier: DecisionStump resume: False useResampling: False
SGD	epochs = 500 epsilon = 0.001 lamda = 10^{-4} learningRate = 0.01 lossFunction: Hinge loss (SVM)
Stacking	classifiers: RF and RotF metaClassifier: LR numFolds = 10
Soft Voting	classifiers: RF and RotF combinationRule: average of probabilities

4.2. Evaluation

In this subsection, we will emphasize the performance evaluation of the classifiers we relied on. Specifically, a variety of ML models are tested in terms of Accuracy, Precision, Recall, F-Measure and AUC. From probabilistic models, we considered BayesNet and NB. From tree-based models, we exploited J48, LMT, RF, RT, RepTree, RotF and AdaBoostM1 (which is based on RotF). The previous models were also compared to SVM, LR, SGD, ANN and k-NN. In addition, we applied ensemble learning, especially Stacking and Soft Voting. In Stacking, we considered as base classifiers, the RotF and RF, and as a meta classifier, the LR. Concerning Soft Voting, the same base classifiers were assumed, and the final prediction was derived from the average of the probabilities.

Moreover, Table 6 illustrates the classifiers' performance after applying SMOTE with 10-fold cross-validation. The RotF model outperforms in comparison to the other models with an accuracy of 99.2%. In addition, we can see that our proposed models demonstrate excellent performance in terms of Precision, Recall, F-Measure and AUC, with percentages over 94%. Moreover, the Stacking and Soft Voting methods and RotF model achieved an AUC of 100%. Finally, the model that presented the lowest similar performance in all metrics was the SVM (linear) with a percentage equal to 94%.

Table 6. ML models' performance with SMOTE and 10-Fold Cross-Validation.

	Accuracy	Precision	Recall	F-Measure	AUC
NB	0.984	0.984	0.984	0.984	0.999
BayesNet	0.984	0.984	0.984	0.984	0.999
SVM (linear)	0.940	0.940	0.940	0.940	0.940
LR	0.974	0.974	0.974	0.974	0.982
ANN	0.968	0.968	0.968	0.968	0.990
k-NN	0.984	0.984	0.984	0.984	0.984
AdaBoostM1	0.978	0.978	0.978	0.978	0.998
SGD	0.974	0.975	0.974	0.974	0.974
RoF	0.992	0.992	0.992	0.992	1
J48	0.974	0.974	0.974	0.974	0.992
LMT	0.982	0.982	0.982	0.982	0.996
RF	0.989	0.989	0.989	0.989	0.999
RT	0.972	0.972	0.972	0.972	0.972
DT	0.974	0.974	0.974	0.974	0.980
Stacking	0.984	0.984	0.984	0.984	1
Soft Voting	0.990	0.990	0.990	0.990	1

The distinct accuracy of the Rotation Forest method relates to the PCA feature transformation that produces rotational matrices with minimal correlations, characterized by a reduced cumulative proportion of matrix diversity. This facilitates the formation of diverse, mutually independent DTs within a Rotation Forest ensemble and thus improves its accuracy [74]. As the results witness, the Rotation Forest forms more accurate individual classifiers than AdaBoostM1 and Random Forest [75].

Also, Table 7 captures the accuracy outcomes of published studies based on the dataset [28] utilizing the same risk factors (namely, features). Specifically, the authors in [29] applied NB, LR and RF, achieving an accuracy of 93.90%, 94.76% and 98.88%, respectively, after 10-fold cross-validation. Our proposed models attained better outcomes in terms of accuracy after SMOTE and 10-fold cross-validation (98.4% NB, 97.4% LR and 98.9% RF). Similarly, the authors in [27] applied LR, k-NN and DT, achieving an accuracy

of 97%, 71.25% and 96.25%, respectively. Our proposed models performed an accuracy of 97.4%, 98.4% and 97.4% for the LR, k-NN and DT models, respectively. We can observe that our proposed models demonstrate slightly better accuracy rates than the comparable research works, except for our k-NN model, which outperforms with a performance gap of 26.15% concerning the respective model of the research work [27].

Table 7. ML models' comparison in terms of accuracy.

	Accuracy		
	Proposed Models	**[29]**	**[27]**
NB	98.4%	93.90%	-
LR	97.4%	94.76%	97%
RF	98.9%	98.88%	-
k-NN	98.4%	-	71.25%
DT	97.4%	-	96.25%

Finally, we have to note the limitations of the current research. The present work considered a public dataset [28] of particular features. Moreover, we relied on data which did not come from a medical unit that could give us diverse features for describing participants' health status. Besides, the acquisition of such data may take considerable time and be difficult from a privacy perspective.

In addition, the features of the dataset do not contain data related to the age and gender of the participants, which would allow us to make the corresponding statistical analysis and processing from a demographic viewpoint. Nevertheless, the dataset is rich in biochemical measurements that can lead us to reliable conclusions.

5. Conclusions

Chronic kidney disease is a condition characterized by progressive loss of kidney function over time. It is a silent disease, as most sufferers have no symptoms. Early diagnosis and treatment of CKD is a serious task for the medical community that resorts to ML theory to design an efficient solution to this challenge.

In the present work, a methodology based on supervised learning is described, which aims to create efficient models for predicting the risk of CKD occurrence by mainly focusing on probabilistic, tree-based and ensemble learning-based models. Moreover, we evaluated SVM, LR, SGD, ANN and k-NN. The derived results highlighted the Rotation Forest, which achieved better performance compared to the other models with an AUC of 100%, Precision, Recall, F-Measure and Accuracy equal to 99.2%. Finally, our proposed models outperformed the published studies based on the same dataset in terms of accuracy.

In future work, we aim to direct our research on Deep Learning methods by applying the Long-Short-term-Memory (LSTM) and CNN and investigate the performance boost that these models may provide. To exploit the capabilities of these models, we aim to follow two directions. The former will apply a data augmentation method to enhance the limited-size dataset before feeding it to the ML models, such as an SVR-based additive input doubling method. In the latter, we will experiment from the beginning with a large-scale non-synthetic dataset.

References

1. Mahadevan, V. Anatomy of the kidney and ureter. *Surgery* **2019**, *37*, 359–364. [CrossRef]
2. Levey, A.S.; Coresh, J. Chronic kidney disease. *Lancet* **2012**, *379*, 165–180. [CrossRef]
3. Koye, D.N.; Magliano, D.J.; Nelson, R.G.; Pavkov, M.E. The global epidemiology of diabetes and kidney disease. *Adv. Chronic Kidney Dis.* **2018**, *25*, 121–132. [CrossRef] [PubMed]
4. CKD. Available online: https://www.urologyhealth.org/urology-a-z/k/kidney-(renal)-failure (accessed on 27 June 2022).
5. Abdel-Kader, K. Symptoms with or because of Kidney Failure? *Clin. J. Am. Soc. Nephrol.* **2022**, *17*, 475–477 [CrossRef]
6. Webster, A.C.; Nagler, E.V.; Morton, R.L.; Masson, P. Chronic kidney disease. *Lancet* **2017**, *389*, 1238–1252. [CrossRef]
7. Wang, Y.N.; Ma, S.X.; Chen, Y.Y.; Chen, L.; Liu, B.L.; Liu, Q.Q.; Zhao, Y.Y. Chronic kidney disease: Biomarker diagnosis to therapeutic targets. *Clin. Chim. Acta* **2019**, *499*, 54–63. [CrossRef]
8. Thakur, N.; Han, C.Y. A study of fall detection in assisted living: Identifying and improving the optimal machine learning method. *J. Sens. Actuator Netw.* **2021**, *10*, 39. [CrossRef]
9. Alexiou, S.; Dritsas, E.; Kocsis, O.; Moustakas, K.; Fakotakis, N. An approach for Personalized Continuous Glucose Prediction with Regression Trees. In Proceedings of the 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Preveza, Greece, 24–26 September 2021; pp. 1–6.
10. Dritsas, E.; Alexiou, S.; Konstantoulas, I.; Moustakas, K. Short-term Glucose Prediction based on Oral Glucose Tolerance Test Values. In Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies-HEALTHINF, Online, 9–11 February 2022; Volume 5, pp. 249–255.
11. Dritsas, E.; Trigka, M. Data-Driven Machine-Learning Methods for Diabetes Risk Prediction. *Sensors* **2022**, *22*, 5304. [CrossRef]
12. Dritsas, E.; Fazakis, N.; Kocsis, O.; Fakotakis, N.; Moustakas, K. Long-Term Hypertension Risk Prediction with ML Techniques in ELSA Database. In Proceedings of the International Conference on Learning and Intelligent Optimization, Athens, Greece, 20–25 June 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 113–120.
13. Fazakis, N.; Dritsas, E.; Kocsis, O.; Fakotakis, N.; Moustakas, K. Long-Term Cholesterol Risk Prediction with Machine Learning Techniques in ELSA Database. In Proceedings of the 13th International Joint Conference on Computational Intelligence (IJCCI), SCIPTRESS, Valletta, Malta, 25–27 October 2021; pp. 445–450.
14. Dritsas, E.; Trigka, M. Machine Learning Methods for Hypercholesterolemia Long-Term Risk Prediction. *Sensors* **2022**, *22*, 5365. [CrossRef]
15. Alballa, N.; Al-Turaiki, I. Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. *Inform. Med. Unlocked* **2021**, *24*, 100564. [CrossRef]
16. Dritsas, E.; Alexiou, S.; Moustakas, K. COPD Severity Prediction in Elderly with ML Techniques. In Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments, Corfu, Greece, 29 June–1 July 2022; pp. 185–189.
17. Dritsas, E.; Trigka, M. Stroke Risk Prediction with Machine Learning Techniques. *Sensors* **2022**, *22*, 4670. [CrossRef] [PubMed]
18. Dritsas, E.; Alexiou, S.; Moustakas, K. Cardiovascular Disease Risk Prediction with Supervised Machine Learning Techniques. In Proceedings of the ICT4AWE, Prague, Czech Republic, 23–25 April 2022; pp. 315–321.
19. Zhang, D.; Gong, Y. The comparison of LightGBM and XGBoost coupling factor analysis and prediagnosis of acute liver failure. *IEEE Access* **2020**, *8*, 220990–221003. [CrossRef]
20. Das, P.K.; Pradhan, A.; Meher, S. Detection of acute lymphoblastic leukemia using machine learning techniques. In *Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 425–437.
21. Konstantoulas, I.; Kocsis, O.; Dritsas, E.; Fakotakis, N.; Moustakas, K. Sleep Quality Monitoring with Human Assisted Corrections. In Proceedings of the International Joint Conference on Computational Intelligence (IJCCI). SCIPTRESS, Virtual, 19–26 August 2021; pp. 435–444.
22. Yarasuri, V.K.; Indukuri, G.K.; Nair, A.K. Prediction of hepatitis disease using machine learning technique. In Proceedings of the 2019 Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), Palladam, India, 12–14 December 2019; pp. 265–269.
23. Saba, T. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *J. Infect. Public Health* **2020**, *13*, 1274–1289. [CrossRef]
24. Yu, C.S.; Lin, Y.J.; Lin, C.H.; Wang, S.T.; Lin, S.Y.; Lin, S.H.; Wu, J.L.; Chang, S.S. Predicting metabolic syndrome with machine learning models using a decision tree algorithm: Retrospective cohort study. *JMIR Med. Inform.* **2020**, *8*, e17110. [CrossRef] [PubMed]
25. Xiao, J.; Ding, R.; Xu, X.; Guan, H.; Feng, X.; Sun, T.; Zhu, S.; Ye, Z. Comparison and development of machine learning tools in the prediction of chronic kidney disease progression. *J. Transl. Med.* **2019**, *17*, 119. [CrossRef]
26. Ghosh, P.; Shamrat, F.J.M.; Shultana, S.; Afrin, S.; Anjum, A.A.; Khan, A.A. Optimization of prediction method of chronic kidney disease using machine learning algorithm. In Proceedings of the 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Bangkok, Thailand, 18–20 November 2020; pp. 1–6.
27. Ifraz, G.M.; Rashid, M.H.; Tazin, T.; Bourouis, S.; Khan, M.M. Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods. *Comput. Math. Methods Med.* **2021**, *2021*, 6141470 . [CrossRef]

28. CKD Prediction Dataset. Available online: https://www.kaggle.com/datasets/abhia1999/chronic-kidney-disease (accessed on 27 June 2022).

29. Islam, M.A.; Akter, S.; Hossen, M.S.; Keya, S.A.; Tisha, S.A.; Hossain, S. Risk factor prediction of chronic kidney disease based on machine learning algorithms. In Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 3–5 December 2020; pp. 952–957.

30. Yashfi, S.Y.; Islam, M.A.; Sakib, N.; Islam, T.; Shahbaaz, M.; Pantho, S.S. Risk prediction of chronic kidney disease using machine learning algorithms. In Proceedings of the 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 1–3 July 2020; pp. 1–5.

31. Chittora, P.; Chaurasia, S.; Chakrabarti, P.; Kumawat, G.; Chakrabarti, T.; Leonowicz, Z.; Jasiński, M.; Jasiński, Ł.; Gono, R.; Jasińska, E.; et al. Prediction of chronic kidney disease-a machine learning perspective. *IEEE Access* **2021**, *9*, 17312–17334. [CrossRef]

32. Revathy, S.; Bharathi, B.; Jeyanthi, P.; Ramesh, M. Chronic kidney disease prediction using machine learning models. *Int. J. Eng. Adv. Technol. (IJEAT)* **2019**, *9*, 6364–6367. [CrossRef]

33. Yadav, D.C.; Pal, S. Performance based Evaluation of Algorithmson Chronic Kidney Disease using Hybrid Ensemble Model in Machine Learning. *Biomed. Pharmacol. J.* **2021**, *14*, 1633–1646. [CrossRef]

34. Baidya, D.; Umaima, U.; Islam, M.N.; Shamrat, F.J.M.; Pramanik, A.; Rahman, M.S. A Deep Prediction of Chronic Kidney Disease by Employing Machine Learning Method. In Proceedings of the 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 28–30 April 2022; pp. 1305–1310.

35. Izonin, I.; Tkachenko, R.; Dronyuk, I.; Tkachenko, P.; Gregus, M.; Rashkevych, M. Predictive modeling based on small data in clinical medicine: RBF-based additive input-doubling method. *Math. Biosci. Eng.* **2021**, *18*, 2599–2613. [CrossRef]

36. Izonin, I.; Tkachenko, R.; Fedushko, S.; Koziy, D.; Zub, K.; Vovk, O. RBF-Based Input Doubling Method for Small Medical Data Processing. In Proceedings of the International Conference on Artificial Intelligence and Logistics Engineering, Kyiv, Ukraine, 20–22 February 2022; Springer: Berlin/Heidelberg, Germany, 2021; pp. 23–31.

37. Bhattacharya, D.; Banerjee, S.; Bhattacharya, S.; Uma Shankar, B.; Mitra, S. GAN-based novel approach for data augmentation with improved disease classification. In *Advancement of Machine Intelligence in Interactive Medical Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 229–239.

38. Tkachenko, R.; Izonin, I.; Vitynskyi, P.; Lotoshynska, N.; Pavlyuk, O. Development of the non-iterative supervised learning predictor based on the ito decomposition and SGTM neural-like structure for managing medical insurance costs. *Data* **2018**, *3*, 46. [CrossRef]

39. Plantinga, L.C.; Miller III, E.R.; Stevens, L.A.; Saran, R.; Messer, K.; Flowers, N.; Geiss, L.; Powe, N.R. Blood pressure control among persons without and with chronic kidney disease: US trends and risk factors 1999–2006. *Hypertension* **2009**, *54*, 47–56. [CrossRef] [PubMed]

40. Shaikh, N.; Shope, M.F.; Kurs-Lasky, M. Urine specific gravity and the accuracy of urinalysis. *Pediatrics* **2019**, *144*. [CrossRef] [PubMed]

41. Erstad, B.L. Serum albumin levels: Who needs them? *Ann. Pharmacother.* **2021**, *55*, 798–804. [CrossRef]

42. Zelnick, L.R.; Batacchi, Z.O.; Ahmad, I.; Dighe, A.; Little, R.R.; Trence, D.L.; Hirsch, I.B.; de Boer, I.H. Continuous glucose monitoring and use of alternative markers to assess glycemia in chronic kidney disease. *Diabetes Care* **2020**, *43*, 2379–2387. [CrossRef]

43. Qiang, Y.; Liu, J.; Dao, M.; Suresh, S.; Du, E. Mechanical fatigue of human red blood cells. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19828–19834. [CrossRef]

44. Seki, M.; Nakayama, M.; Sakoh, T.; Yoshitomi, R.; Fukui, A.; Katafuchi, E.; Tsuda, S.; Nakano, T.; Tsuruya, K.; Kitazono, T. Blood urea nitrogen is independently associated with renal outcomes in Japanese patients with stage 3–5 chronic kidney disease: A prospective observational study. *BMC Nephrol.* **2019**, *20*, 1–10. [CrossRef]

45. Lin, Y.L.; Chen, S.Y.; Lai, Y.H.; Wang, C.H.; Kuo, C.H.; Liou, H.H.; Hsu, B.G. Serum creatinine to cystatin C ratio predicts skeletal muscle mass and strength in patients with non-dialysis chronic kidney disease. *Clin. Nutr.* **2020**, *39*, 2435–2441. [CrossRef]

46. Borrelli, S.; Provenzano, M.; Gagliardi, I.; Ashour, M.; Liberti, M.E.; De Nicola, L.; Conte, G.; Garofalo, C.; Andreucci, M. Sodium intake and chronic kidney disease. *Int. J. Mol. Sci.* **2020**, *21*, 4744. [CrossRef]

47. Kovesdy, C.P.; Matsushita, K.; Sang, Y.; Brunskill, N.J.; Carrero, J.J.; Chodick, G.; Hasegawa, T.; Heerspink, H.L.; Hirayama, A.; Landman, G.W.; et al. Serum potassium and adverse outcomes across the range of kidney function: A CKD Prognosis Consortium meta-analysis. *Eur. Heart J.* **2018**, *39*, 1535–1542. [CrossRef]

48. Kim, J.S.; Choi, S.; Lee, G.; Cho, Y.; Park, S.M. Association of hemoglobin level with fracture: A nationwide cohort study. *J. Bone Miner. Metab.* **2021**, *39*, 833–842. [CrossRef] [PubMed]

49. Sun, Y.; Jiang, L.; Shao, X. Predictive value of procalcitonin for diagnosis of infections in patients with chronic kidney disease: A comparison with traditional inflammatory markers C-reactive protein, white blood cell count, and neutrophil percentage. *Int. Urol. Nephrol.* **2017**, *49*, 2205–2216. [CrossRef] [PubMed]

50. Ku, E.; Lee, B.J.; Wei, J.; Weir, M.R. Hypertension in CKD: Core curriculum 2019. *Am. J. Kidney Dis.* **2019**, *74*, 120–131. [CrossRef] [PubMed]

51. Maldonado, S.; López, J.; Vairetti, C. An alternative SMOTE oversampling strategy for high-dimensional datasets. *Appl. Soft Comput.* **2019**, *76*, 380–389. [CrossRef]

52. Obilor, E.I.; Amadi, E.C. Test for significance of Pearson's correlation coefficient. *Int. J. Innov. Math. Stat. Energy Policies* **2018**, *6*, 11–23.

53. Gnanambal, S.; Thangaraj, M.; Meenatchi, V.; Gayathri, V. Classification algorithms with attribute selection: An evaluation study using WEKA. *Int. J. Adv. Netw. Appl.* **2018**, *9*, 3640–3644.

54. Disha, R.A.; Waheed, S. Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique. *Cybersecurity* **2022**, *5*, 1. [CrossRef]

55. Palaka, E.; Grandy, S.; van Haalen, H.; McEwan, P.; Darlington, O. The impact of CKD anaemia on patients: Incidence, risk factors, and clinical outcomes—A systematic literature review. *Int. J. Nephrol.* **2020**, *2020*, 7692376. [CrossRef]

56. Feng, X.; Li, S.; Yuan, C.; Zeng, P.; Sun, Y. Prediction of slope stability using naive Bayes classifier. *KSCE J. Civ. Eng.* **2018**, *22*, 941–950. [CrossRef]

57. Marcot, B.G.; Penman, T.D. Advances in Bayesian network modelling: Integration of modelling technologies. *Environ. Model. Softw.* **2019**, *111*, 386–393. [CrossRef]

58. Pisner, D.A.; Schnyer, D.M. Support vector machine. In *Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 101–121.

59. Nusinovici, S.; Tham, Y.C.; Yan, M.Y.C.; Ting, D.S.W.; Li, J.; Sabanayagam, C.; Wong, T.Y.; Cheng, C.Y. Logistic regression was as good as machine learning for predicting major chronic diseases. *J. Clin. Epidemiol.* **2020**, *122*, 56–69. [CrossRef] [PubMed]

60. Morariu, D.; Crețulescu, R.; Breazu, M. The WEKA multilayer perceptron classifier. *Int. J. Adv. Stat. It&C Econ. Life Sci.* **2017**, *7*, 1.

61. Ali, N.; Neagu, D.; Trundle, P. Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets. *SN Appl. Sci.* **2019**, *1*, 1559. [CrossRef]

62. Ihya, R.; Namir, A.; Filali, S.E.; Daoud, M.A.; Guerss, F.Z. J48 algorithms of machine learning for predicting user's the acceptance of an E-orientation systems. In Proceedings of the 4th International Conference on Smart City Applications, Casablanca, Morocco, 2–4 October 2019; pp. 1–8.

63. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Bui, D.T. A comparative study of support vector machine and logistic model tree classifiers for shallow landslide susceptibility modeling. *Environ. Earth Sci.* **2019**, *78*, 560. [CrossRef]

64. Reis, I.; Baron, D.; Shahaf, S. Probabilistic random forest: A machine learning algorithm for noisy data sets. *Astron. J.* **2018**, *157*, 16. [CrossRef]

65. Alsharif, N. Ensembling PCA-based Feature Selection with Random Tree Classifier for Intrusion Detection on IoT Network. In Proceedings of the 2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Semarang, Indonesia, 20–21 October 2021; pp. 317–321.

66. Mohamed, W.N.H.W.; Salleh, M.N.M.; Omar, A.H. A comparative study of reduced error pruning method in decision tree algorithms. In Proceedings of the 2012 IEEE International Conference on Control System, Computing and Engineering, Penang, Malaysia, 23–25 November 2012; pp. 392–397.

67. Lu, H.; Meng, Y.; Yan, K.; Gao, Z. Kernel principal component analysis combining rotation forest method for linearly inseparable data. *Cogn. Syst. Res.* **2019**, *53*, 111–122. [CrossRef]

68. Polat, K.; Sentürk, U. A novel ML approach to prediction of breast cancer: Combining of mad normalization, KMC based feature weighting and AdaBoostM1 classifier. In Proceedings of the 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 19–21 October 2018; pp. 1–4.

69. Zhang, Y.; Saxe, A.M.; Advani, M.S.; Lee, A.A. Energy–entropy competition and the effectiveness of stochastic gradient descent in machine learning. *Mol. Phys.* **2018**, *116*, 3214–3223. [CrossRef]

70. Burka, D.; Puppe, C.; Szepesváry, L.; Tasnádi, A. Voting: A machine learning approach. *Eur. J. Oper. Res.* **2022**, *299*, 1003–1017. [CrossRef]

71. Pavlyshenko, B. Using stacking approaches for machine learning models. In Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 21–25 August 2018; pp. 255–258.

72. Moccia, S.; De Momi, E.; El Hadji, S.; Mattos, L.S. Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics. *Comput. Methods Programs Biomed.* **2018**, *158*, 71–91. [CrossRef]

73. WEKA Tool. Available online: https://www.weka.io/ (accessed on 27 June 2022).

74. Bustamam, A.; Musti, M.I.; Hartomo, S.; Aprilia, S.; Tampubolon, P.P.; Lestari, D. Performance of rotation forest ensemble classifier and feature extractor in predicting protein interactions using amino acid sequences. *BMC Genom.* **2019**, *20*, 950. [CrossRef]

75. Jukic, S.; Saracevic, M.; Subasi, A.; Kevric, J. Comparison of ensemble machine learning methods for automated classification of focal and non-focal epileptic EEG signals. *Mathematics* **2020**, *8*, 1481. [CrossRef]