

# eXtreme Gradient Boosting-based method to classify patients with COVID-19

Antonio Ramón,<sup>1</sup> Ana Maria Torres,<sup>2</sup> Javier Milara,<sup>1,3</sup> Joaquín Cascón,<sup>2</sup> Pilar Blasco,<sup>1</sup> Jorge Mateo <sup>2</sup>

<sup>1</sup>Pharmacy Department, General University Hospital Consortium of Valencia, Valencia, Spain

<sup>2</sup>Institute of Technology, Universidad de Castilla-La Mancha, Cuenca, Spain

<sup>3</sup>Pharmacy Department, University of Valencia, Valencia, Spain

## Correspondence to

Dr Jorge Mateo, Institute of Technology, Universidad de Castilla-La Mancha, Cuenca, Castilla-La Mancha, Spain; [jomateo2010@gmail.com](mailto:jomateo2010@gmail.com)

Accepted 15 June 2022  
Published Online First  
18 July 2022

## ABSTRACT

Different demographic, clinical and laboratory variables have been related to the severity and mortality following SARS-CoV-2 infection. Most studies applied traditional statistical methods and in some cases combined with a machine learning (ML) method. This is the first study to date to comparatively analyze five ML methods to select the one that most closely predicts mortality in patients admitted with COVID-19. The aim of this single-center observational study is to classify, based on different types of variables, adult patients with COVID-19 at increased risk of mortality. SARS-CoV-2 infection was defined by a positive reverse transcriptase PCR. A total of 203 patients were admitted between March 15 and June 15, 2020 to a tertiary hospital. Data were extracted from the electronic medical record. Four supervised ML algorithms (k-nearest neighbors (KNN), decision tree (DT), Gaussian naïve Bayes (GNB) and support vector machine (SVM)) were compared with the eXtreme Gradient Boosting (XGB) method proposed to have excellent scalability and high running speed, among other qualities. The results indicate that the XGB method has the best prediction accuracy (92%), high precision (>0.92) and high recall (>0.92). The KNN, SVM and DT approaches present moderate prediction accuracy (>80%), moderate recall (>0.80) and moderate precision (>0.80). The GNB algorithm shows relatively low classification performance. The variables with the greatest weight in predicting mortality were C reactive protein, procalcitonin, glutamyl oxaloacetic transaminase, glutamyl pyruvic transaminase, neutrophils, D-dimer, creatinine, lactic acid, ferritin, days of non-invasive ventilation, septic shock and age. Based on these results, XGB is a solid candidate for correct classification of patients with COVID-19.

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ COVID-19 causes severe acute respiratory syndrome manifesting clinically from asymptomatic to mild forms with cough, fever and myalgia, to triggering bilateral pneumonia with severe respiratory failure and multiorgan damage, which can lead to death.
- ⇒ A wide variety of clinical, laboratory and demographic variables associated with severity and mortality from COVID-19 have been identified, including but not limited to age, previous healthy status and laboratory parameters.
- ⇒ Most studies do not perform a comprehensive risk assessment to predict COVID-19-related mortality due to the increased number of clinical, laboratory and anthropometric variables which limits conclusions.

## WHAT THIS STUDY ADDS

- ⇒ Machine learning, as part of artificial intelligence, is a useful tool to assign variables to predict COVID-19 mortality.
- ⇒ The eXtreme Gradient Boosting (XGB) model of machine learning was superior to decision tree, Gaussian naïve Bayes, k-nearest neighbor and support vector machines in predicting variables for COVID-19 mortality.
- ⇒ The variables that best predict COVID-19 mortality were levels of C reactive protein, procalcitonin, glutamate oxaloacetate transferase and glutamate pyruvate transferase transaminases, number of neutrophils, D-dimer, creatinine, septic shock and age.

## INTRODUCTION

COVID-19, caused by a coronavirus-2 (SARS-CoV-2) infection and causing severe acute respiratory syndrome, first emerged in Wuhan, Hubei, China in December 2019.<sup>1</sup> The virus is highly transmissible, even more than SARS-CoV-2, manifesting clinically from asymptomatic or mild forms with cough, fever and myalgia, to triggering bilateral pneumonia with severe respiratory failure that requires mechanical ventilation and/or multiorgan damage that

can lead to death.<sup>3</sup> During the first wave, the mortality rate due to COVID-19 was less than 3%, although the fatality rate for severe cases is high, according to the WHO. The current global epidemiological situation is characterized by a high percentage of the population immunized against SARS-CoV-2, as well as an increase in the proportion of mild and asymptomatic cases, with the case fatality rate being less than 1%.<sup>4</sup> In Spain, as of February 11, 2022, 10,555,197 cases of COVID-19 have



© American Federation for Medical Research 2022. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Ramón A, Torres AM, Milara J, et al. *J Invest Med* 2022;**70**:1472–1480.

**HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY**

- ⇒ The present work indicates that machine learning is a useful tool to predict mortality in hospitalized patients.
- ⇒ Between different types of machine learning procedures, XGB is the best tool that predicts mortality and can be used routinely to identify which patients have an increased risk of worsening.
- ⇒ This work identifies laboratory parameters that better predict mortality and which can be potentially used to stratify patients at risk.

been confirmed, including a total of 95,606 deaths.<sup>4</sup> Case fatality rates help to understand the severity of the disease, identify populations at risk and assess the quality of health-care. Predicting the clinical course of this disease based on several variables is of vital importance for proper patient management.

A wide variety of clinical, laboratory and demographic variables associated with severity and mortality from COVID-19 have been identified.<sup>5,6</sup> Most studies did not perform comprehensive risk assessment to predict COVID-19-related mortality.<sup>7,8</sup> To circumvent these drawbacks, machine learning (ML) models have emerged designed to make accurate predictions using data from a multitude of variables, as opposed to classic statistical models created to make inferences about relationships between variables. ML, as part of artificial intelligence (AI), uses statistical and mathematical algorithms that allow the opting of patterns that help in making complex decisions.<sup>9</sup> These algorithms can be used to develop predictive models and reduce the complexity of clinical phenotypes. They are used in biomedicine as elements of clinical decision support and as generators of new clinical knowledge. For example, they have been used in the prediction of hospitalization for heart disease.<sup>10</sup>

On the other hand, progress has been made in the modeling of clinical data in electronic medical records (EMR) and specifically in the ability of ML techniques to predict mortality.<sup>11</sup> ML constitutes an integrative method that allows observation of the combined effect of multiple variables and their interactions, allowing generation of knowledge about the disease from patients' EMR data, and is a very useful tool in conditions where structured numerical data are readily available.

ML algorithms have been explored in different fields of COVID-19, mainly in the detection of outbreaks and spread of SARS-CoV-2,<sup>12</sup> prediction of incidence rates,<sup>13</sup> early diagnosis,<sup>14</sup> prediction of risk of complications and severity,<sup>15</sup> as well as prediction of mortality risk.<sup>16–20</sup> Syeda *et al*<sup>21</sup> recently conducted a systematic review on the role of AI as a comprehensive and critical technology in combating the COVID-19 crisis in the fields of epidemiology, diagnosis and disease progression. Only 14.6% of the studies were related to the latter. Thus, a more precise approach to COVID-19 mortality is needed. To our knowledge, this is the first study to develop, compare and validate five ML models in predicting in-hospital mortality in patients admitted with COVID-19 in a tertiary-level hospital and research reference hospital in Spain during the first wave

of the pandemic. Demographic, clinical and laboratory data easily extractable from the hospital EMR were used for prediction. The study is structured in a brief introduction highlighting the case fatality or mortality rate as a key variable in the study of a newly emerging disease such as COVID-19 and the importance of applying ML to numerous variables associated with hospital mortality. The Materials and methods section describes the types of variables included and the data collected and used for the different ML models applied. The Results section includes, among others, the accuracy values of the five validated algorithms for predicting hospital mortality from COVID-19. The Discussion section compares the results obtained in this study with the results of other studies using ML. Finally, the Conclusion section highlights the eXtreme Gradient Boosting (XGB) method over the other ML methods as a method for predicting mortality, facilitating patient stratification and optimizing medical resources.

**MATERIALS AND METHODS****Data sources**

Patient data were obtained from different internal sources of the hospital, such as the EMR (Hosix. Net. Ink.), which includes a module for registration of results of clinical analysis and a module for electronic prescription of drugs and the prescription program of the intensive care unit (ICU) (IntelliSpace Critical Care and Anesthesia, V.H.02.00, Philips Iberica). With this information, a data collection questionnaire (DCQ) was constructed individually by patient.

**Study design and population**

This is a retrospective observational study carried out in a tertiary-level hospital that attends a monthly average of 12,000 emergencies and 2000 hospital admissions. A total of 203 patients admitted to the hospital with SARS-CoV-2 were included. Inclusion criteria were all patients admitted to the Valencia University General Hospital Consortium with SARS-CoV-2 infection confirmed microbiologically by reverse transcriptase PCR assay of a nasopharyngeal swab between March 15 and June 15, 2020. The patients selected were admitted to the hospital during a period of  $\geq 7$  days. Exclusion criteria were patients  $\leq 18$  years old and patients with missing clinical data of more than one clinical/laboratory variable during this period. Participants gave informed consent before taking part in study.

**Study data**

Data on demographic, clinical and laboratory variables were included in the CRD. The questionnaire was divided into eight sections.

**Patient characteristics**

Demographic variables such as age and sex and the following clinical variables were included: weight, height and presence of comorbidities of interest (hypertension, diabetes mellitus, chronic obstructive pulmonary disease, asthma, other chronic respiratory disease (eg, pulmonary dysplasia, cystic fibrosis), use of oxygen therapy or presence of tracheostomy, heart failure, ischemic heart disease, pulmonary hypertension, recent catheterization, renal failure (RF),

cirrhosis, history of neurological, active haematological or oncological neoplasia (with active treatment, diagnosis or recurrence/metastasis <5 years, excluding diagnosis of squamous cell and basal cell carcinoma), and HIV). In the event that the patient presented another type of serious underlying pathology, it was specified in an open-text section. The following were taken into account as pharmacological treatment prior to admission: ACE inhibitors/angiotensin-2 receptor antagonists, non-steroidal anti-inflammatory drugs, and antihistamines and/or montelukast, as well as whether the patient was a healthcare professional or if the previous stay was in a residence or another healthcare center.

#### Initial data on arrival at the hospital

These included date of admission to the emergency room, date of admission to the hospital, date of onset of symptoms, date of microbiological confirmation, limitation of life support treatment and its date, and whether the patient required admission to the ICU. If the patient was admitted to the ICU, data on the risk of mortality (CURB-65 scale (Confusion, Urea nitrogen, Respiratory rate, Blood pressure, 65 years of age and older)), level of altered consciousness (Glasgow Scale) and other clinical variables were included: fever ( $\geq 38^\circ\text{C}$ ), respiratory rate  $>24$  breaths per minute and systolic blood pressure  $<90$  mm Hg in the first 24 hours, baseline oxygen saturation ( $\text{SpO}_2$ ), and number of quadrants affected on the chest radiograph (1–4).

#### Data on admission to the ICU

These included date of admission, Acute Physiology And Chronic Health Evaluation (APACHE) II scores and Sepsis related Organ Failure Assessment (SOFA) scores.

#### Analytical data

The closest analysis after hospital admission (emergency/admission), the first analysis since admission to the ICU and the last analysis of the hospital stay were included. The laboratory parameters collected were leukocytes, neutrophils, lymphocytes, platelets, C reactive protein (CRP), glutamate oxaloacetate transferase (GOT), glutamate pyruvate transferase (GPT), lactate dehydrogenase (LDH), serum creatinine, hemoglobin, procalcitonin (PCT), lactic acid, creatine phosphokinase (CPK), D-dimer and ferritin.

#### Pharmacological treatment

This was taken into account if the patient participated in a clinical trial. The drugs considered were lopinavir/ritonavir, remdesivir, interferon beta, hydroxychloroquine, chloroquine, darunavir/cobicistat, darunavir/ritonavir, darunavir/cobicistat/tenofovir/emtricitabine, fosamprenavir, tocilizumab, sarilumab, ciclosporin, anakinra, tacrolimus, eculizumab, azithromycin, immunoglobulins, baricitinib and tofacitinib. For all these drugs, the dosage regimen and duration of treatment were included, and in the case of tocilizumab/sarilumab the levels of interleukin 6 (in pg/mL) and D-dimer (in  $\mu\text{g/mL}$ ) were taken into account before and after treatment, as well as where treatment was started (ICU/no ICU). Other treatments included antibiotics, vaso-pressors, prescribed and/or bolus corticosteroids, and use of low molecular weight heparin, distinguishing between

prophylactic or treatment doses. The corticosteroids included were methylprednisolone, hydrocortisone, dexamethasone and prednisone.

#### Microbiological tests

The isolated micro-organism was taken into account in all cases. The tests were tracheal aspirate, blood cultures, presence of influenza and/or coinfection, pneumococcal antigen and *Legionella* antigen in urine.

#### Techniques performed during admission

The following were included: oxygen therapy, non-invasive ventilation (NIV), mechanical ventilation, ventilation in prone position, hemodialysis/hemofiltration and extracorporeal membrane oxygenation system.

#### Final evolution of the patient

First, the severity of the SARS-CoV-2 infection was indicated according to the classification of severity levels of respiratory infections included in the COVID-19 clinical management protocol of the Ministry of Health on June 18, 2020. Complications during admission (acute respiratory distress syndrome (ARDS), sepsis, septic shock, nosocomial pneumonia (not COVID-19), other nosocomial infection (not COVID-19, not pneumonia), and acute renal and liver failure) were included. As a final assessment, improvement in symptoms (fever, cough, etc) together with radiological improvement and/or alveolar pressure / inspired oxygen fraction (PaFi)  $\geq 300$  mm Hg or  $\text{SpO}_2 >93$  without oxygen administration during the first 7, 14, 21 or 28 days of admission, depending on their duration, were recorded. A distinction was made between hospital discharge or exitus. The date and destination of discharge (home, residence or support center, or unknown destination), date of discharge of ICU, date of exitus and days of admission were included. Whether the patient was readmitted within 14 days after discharge was also taken into account.

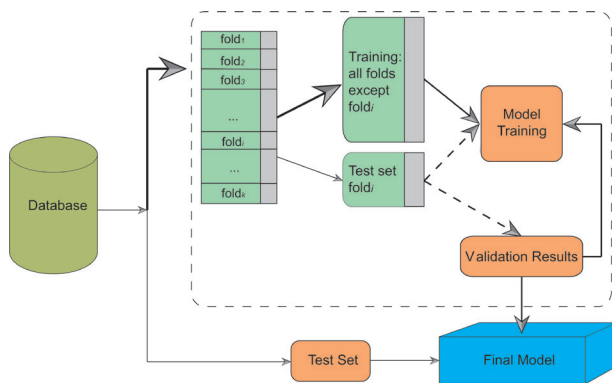
#### Model development

An XGB-based method was implemented in this study because it is a flexible, highly efficient, portable and flexible supervised learning algorithm. The main advantages are that it is fast to run and is scalable and allows parallel computing.<sup>22–25</sup> XGB algorithms are developed under the framework of gradient boosting. XGB features parallel tree boosting (also known as gradient-boosted decision trees), which solves many data science problems accurately and quickly. XGB is adopted to build a COVID-19 patient classification model. Given a data set  $S = x_p, y_p$ , the XGB model was designed using the following:

$$\hat{y}_j = \sum_{p=1}^P t_p(x_i) \quad (1)$$

where  $x_i$  is the input vector with  $m$  time variables,  $\hat{y}_j$  shows the predicted output,  $y_j$  represents the output,  $t_p$  represents a tree with leaf weight  $w_p$  and structure  $u_p$ ,  $j=1; 2; \dots; n$ , and  $P$  corresponds to the number of trees.

The regularized objective function for the proposed method is shown in equation 2. In this case, it is different from that of ensemble methods. In the proposed method, a second-order Taylor expansion is implemented to



**Figure 1** Training and validation scheme for machine learning methods.

approximate the objective function of XGB in order to improve prediction accuracy.<sup>22 23</sup> To control the complexity of the model and avoid overfitting, the regulation term is used, which is represented by the weights of the leaf nodes and the tree depth.

$$R = \sum_j r(y_j; y_i) + \sum_p \psi(t_p) \quad (2)$$

$$\Psi(t_p) = \lambda f_p + \frac{1}{2} \gamma \|w_p\|^2 \quad (3)$$

As can be seen in equation 3,  $f_p$  corresponds to the tree pruning used to control overfitting.  $f_p$  shows the number of leaves on the tree. Pruning is a method to improve generalization in trees. Once the trees are built, the proposed XGBoost performs a ‘pruning’ step that, starting at the bottom (where the leaves are) and moving up to the root node, looks to see if the gain falls below  $\lambda$ . If the first node encountered has a gain value below  $\lambda$ , then the node is pruned and the pruner moves up the tree to the next node. If, on the other hand, the node has a gain greater than  $\lambda$ , the node is left and the pruner does not check the parent nodes.<sup>23 24 26</sup> The  $R()$  function penalizes the complexity of the method. The learning rate is shown by  $\lambda$  and  $w$  is the vector of leaf scores.  $R()$  represents a function that measures the difference between the target output  $y_j$  and the expected output  $\hat{y}_j$ . To control the complexity weight of the system, a parameter  $\gamma$  is employed.<sup>23 24 26</sup> To improve performance, this study seeks to minimize equation 2.

The functions of the functions in equation 2 are incorporated in the tree set model.<sup>23 24 26</sup> Because of this, equation 2 cannot be optimized through traditional Euclidean space optimization systems. Therefore, in this study,  $y_j$  was the  $j$ -th sample estimate at  $s$ -th iteration. With all these, equation 2 would look like the one shown in equation 4.

$$R^s = \sum_{j=1}^r \left( \hat{y}_j^{(s-1)}, y_j + C_s(x_j) \right) + \sum_p \psi(C_s) \quad (4)$$

To reduce the objective function, the tree generated  $C_s$  by the  $j$ -th sample at the  $s$ -th iteration is added. Moreover, in the proposed method, the second-order approximation has been applied to optimize the objective function.<sup>22-24</sup>

$$R^s \approx \sum_{j=1}^r \left[ \hat{y}_j^{(s-1)}, y_j + h_j C_s(x_j) + \frac{1}{2} b_j C_s^2(x_j) \right] + \psi(C_s) \quad (5)$$

where  $h_j = \partial_{\hat{y}_j^{(s-1)}} \left( \hat{y}_j^{(s-1)}, y_j \right)$  represents the first-order gradient statistic for the loss function  $R()$  and  $b_j = \partial_{\hat{y}_j^{(s-1)}}^2 \left( \hat{y}_j^{(s-1)}, y_j \right)$  shows the second. The optimal weight  $w_{r_v}$  of the license  $v$  for a fixed structure  $u(x)$  can be estimated as:

$$w - r_v = - \frac{\sum_{j \in K_v} (b_j)}{\sum_{j \in K_v} (b_j + \gamma)} \quad (6)$$

Finally, the optimal value can be achieved by means of equation 7 for the proposed method.

$$\tilde{R}(s)(u) = -\frac{1}{2} \frac{\left( \sum_{j \in K_v} (b_j) \right)^2}{\sum_{j \in K_v} (b_j + \gamma)} + \lambda F \quad (7)$$

For this study, the proposed method was compared with different ML methods in order to classify patients into two groups: patients without risk and patients with risk of mortality from COVID-19. The methods involved decision tree (DT),<sup>27</sup> Gaussian naïve Bayes (GNB),<sup>28 29</sup> k-nearest neighbors (KNN),<sup>30 31</sup> support vector machines (SVM)<sup>32 33</sup> and the proposed method XGB.<sup>23 24</sup> The MatLab Statistical and Machine Learning Toolbox (MatLab V2021a; The MathWorks, Natick, Massachusetts, USA) was used to implement the models. A fivefold cross-validation was applied to avoid overfitting. The database was divided into two groups, 70% was used for training and 30% for testing, and patients were not shared.

The phases implemented for the whole study are described in figure 1. As can be seen, the subjects to be studied were first chosen. Once the database was created, training and validation of the ML methods were carried out.

## Performance evaluation

In this paper, the different methods were compared with the following metrics: degenerate Youden index (DYI), specificity, precision (also known as positive predictive value), recall (also known as sensitivity), balanced accuracy, receiver operating characteristic (ROC) and area under the curve (AUC). The  $F_1$  score is described as:

$$F1_{score} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

Matthew’s correlation coefficient (MCC) was also used to test the performance of the ML methods, defined as:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (9)$$

where TP represents the number of true positives, FP is the number of false positives, TN shows the number of true negatives and FN corresponds to the number of false negatives. Cohen’s kappa index was used to estimate the overall performance of the system.<sup>34</sup>

## RESULTS

This section describes the results obtained by using patient records for training and validation of COVID-19 mortality classification. The performance of the proposed system was



**Table 1** Mean value and SD of balanced accuracy, recall, precision,  $F_1$  score, AUC, MCC, DYI and kappa of the machine learning models and the proposed method implemented in this study

| Methods | Balanced accuracy | Recall     | Precision  | $F_1$ score |
|---------|-------------------|------------|------------|-------------|
| SVM     | 81.24±0.54        | 81.31±0.61 | 81.26±0.64 | 81.54±0.67  |
| DT      | 83.25±0.78        | 83.17±0.72 | 83.15±0.73 | 83.31±0.74  |
| GNB     | 80.08±0.67        | 80.11±0.75 | 79.91±0.62 | 80.07±0.66  |
| KNN     | 85.42±0.59        | 85.48±0.45 | 86.01±0.35 | 86.08±0.41  |
| XGB     | 92.37±0.31        | 92.43±0.36 | 92.55±0.24 | 92.42±0.27  |

| Methods | AUC       | MCC        | DYI        | Kappa      |
|---------|-----------|------------|------------|------------|
| SVM     | 0.81±0.02 | 72.03±0.65 | 80.96±0.68 | 72.65±0.64 |
| DT      | 0.83±0.02 | 72.94±0.72 | 82.79±0.63 | 73.58±0.69 |
| GNB     | 0.80±0.02 | 71.65±0.75 | 79.87±0.71 | 72.61±0.73 |
| KNN     | 0.85±0.02 | 75.16±0.43 | 85.23±0.47 | 75.86±0.51 |
| XGB     | 0.92±0.02 | 84.23±0.26 | 92.39±0.24 | 85.14±0.25 |

AUC, area under the curve; DT, decision tree; DYI, degenerate Youden index; GNB, Gaussian naïve Bayes; KNN, k-nearest neighbors; MCC, Matthew's correlation coefficient; SVM, support vector machine; XGB, eXtreme Gradient Boosting.

compared with different ML methods that are accepted in the scientific community.

Table 1 presents the results achieved from the classification methods such as SVM, DT, GNB and KNN and the proposed system for mortality classification of patients with COVID-19. As can be seen, the systems based on SVM and GNB obtained lower accuracy value than the rest of the methods; these values are close to 81%. As for the DT and KNN methods, they show improved classification capability by obtaining an accuracy value of 83%. On the other hand, the proposed XGB system achieved an accuracy value of 92%, a significant increase over the previous methods, which translates to better prediction. The algorithms that come closest to XGB in terms of precision and recall values are KNN and DT, which again performed better than SVM and GNB. As can be seen in table 1, the same thing happens with parameter  $F_1$  score, where XGB obtained higher values, which imply an improvement in classification.

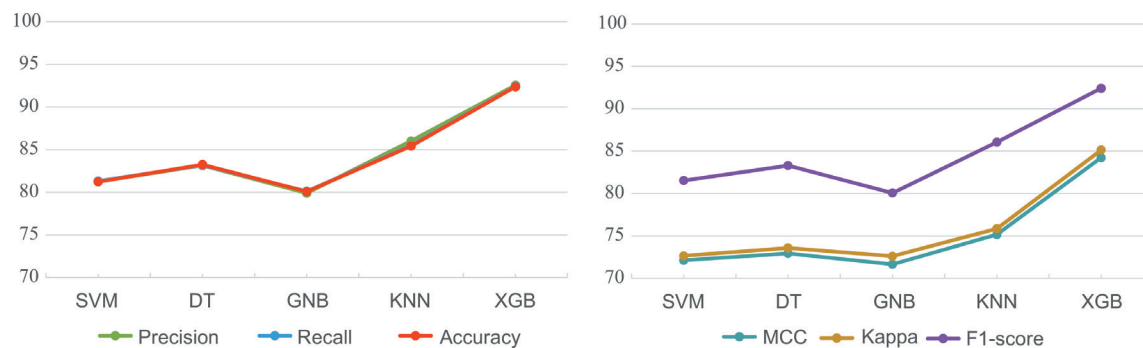
To test the performance of the proposed XGB system in classifying mortality of patients with COVID-19, other

parameters widely used in the literature, such as AUC, MCC, DYI and kappa index, were calculated. For this analysis, one of the most reliable statistical indices available, the MCC, was used. This coefficient produces a high score only if the prediction has been performed well in the four categories of the matrix. The results in the four categories of the confusion matrix (true positives, false negatives, true negatives and false positives) are proportional to the size of the positive elements and the size of the negative elements in the data set. As can be observed in table 1, the proposed method, XGB, achieved a value of 84.23%, increasing the values achieved by KNN and DT, which presented 75.16% and 72.94%. Both SVM and GNB showed worse performance in this parameter. As for the kappa index, XGB obtained a value close to 85%, improving the value of KNN and DT by 9.28% and 11.56%, respectively. The same is true for the AUC and DYI parameters: the XGB method achieved a higher value, which means it can better classify mortality in patients with COVID-19.

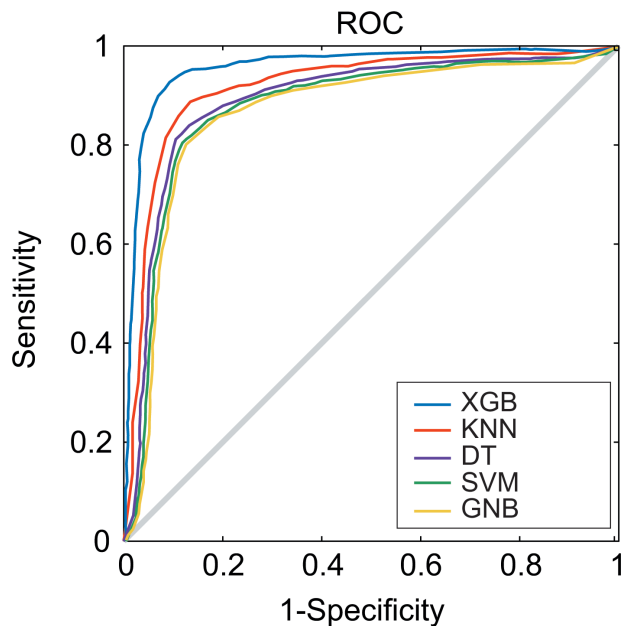
Figure 2 shows a summary of the comparison between the XGB method and the other classifiers with respect to accuracy, recall and precision. XGB achieved values of 0.924, 0.924 and 0.925, respectively, while those of KNN were 0.854, 0.855 and 0.860. Figure 2 also shows the values obtained for MCC, kappa and  $F_1$  score. The proposed method obtained values of 0.842, 0.851 and 0.924, respectively. The next closest system to XGB is KNN, with values of 0.752, 0.758 and 0.860. In all parameters, it can be observed how the proposed method shows better performance in predicting mortality.

On the other hand, ROC was used to compare the classification capability of the proposed system with that of other ML methods. The curve is the result of plotting, for each threshold value, the sensitivity and specificity.<sup>35</sup> In figure 3, the results obtained by the different systems of classification between patients with COVID-19 mortality and those who survive are shown, where a larger area can be appreciated for the XGB method, which implies better classification of the two classes; the values can be seen in table 1.

For clarity, all metrics have been grouped for each data set (training and test) and are presented as a radar plot. A perfect score on all metrics would be represented by a circle the size of the entire grid. In our study, model training sets have higher scores on all training set metrics and generally have lower scores on the test set. The shape of the graphs



**Figure 2** Graphical representation of precision, recall, accuracy, MCC, kappa and  $F_1$  score values in percentages. DT, decision tree; GNB, Gaussian naïve Bayes; KNN, k-nearest neighbors; MCC, Matthew's correlation coefficient; SVM, support vector machine; XGB, eXtreme Gradient Boosting.



**Figure 3** ROC curves for the five assessed machine learning predictors. DT, decision tree; GNB, Gaussian naïve Bayes; KNN, k-nearest neighbors; ROC, receiver operating characteristic; SVM, support vector machine; XGB, eXtreme Gradient Boosting.

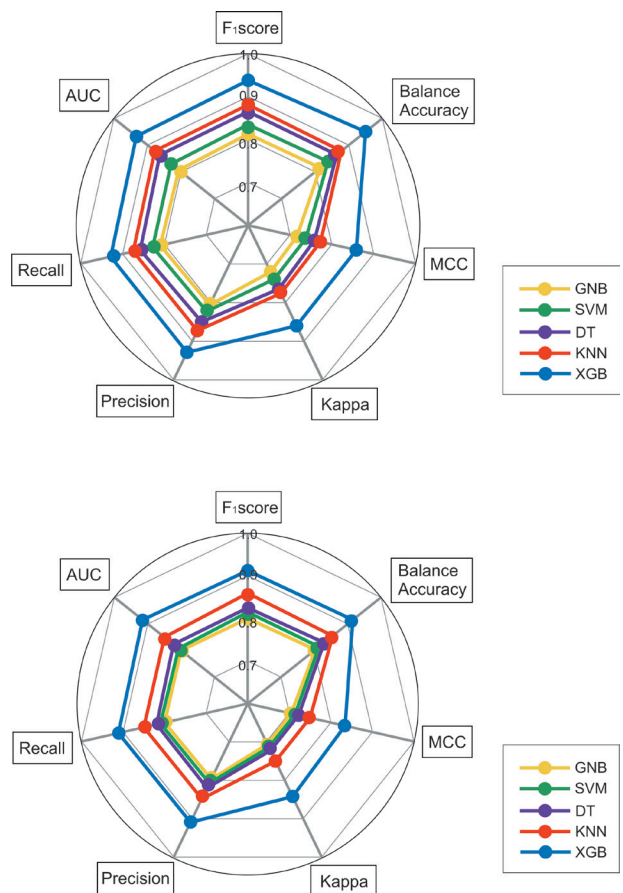
can also be indicative of the quality of the models. The larger the area of the circle of the test set, the better the prediction method will be. The proposed XGB system (figure 4) is a good example of a balanced model. The training and test sets give rise to similar pie charts. These similarities are due to the system obtaining an optimal training point, with no overfitting or underfitting, and therefore the method has high generalizability. That is, given a new input, the system does well to provide a correct output. As can be seen, the GNB method performed the worst on most metrics. In view of the results obtained, we can say that the proposed XGB system manages to classify patients with COVID-19 with high accuracy and in an automatic way, confirming the fact that this tool would be of great help in clinical practice.

## DISCUSSION

The current SARS-CoV-2 pandemic is associated with high morbidity and mortality.<sup>36–37</sup> Most mortality prediction models for COVID-19 that use ML are based partially or totally on subjective clinical data, which may vary depending on the study.<sup>38</sup> As far as we know, this is the first study to develop, compare and evaluate five supervised ML methods in the Spanish population to predict mortality in patients admitted with COVID-19 in a tertiary hospital.

## ML models analyzed and related work

Unlike other studies,<sup>17–20</sup> 273 clinical, demographic and laboratory predictors were included to fit the models. Of all the ML classifiers applied, the XGB method was the pattern recognition method that managed to more precisely discriminate between patients at risk of mortality from COVID-19 and those who are not. This model was analyzed and compared with different supervised ML methods described in the literature, such as GNB, DT,



**Figure 4** Radar plot of the training phase (top) and test (bottom) for prediction of mortality in patients with COVID-19. AUC, area under the curve; DT, decision tree; GNB, Gaussian naïve Bayes; KNN, k-nearest neighbors; MCC, Matthew's correlation coefficient; SVM, support vector machine; XGB, eXtreme Gradient Boosting.

KNN or SVM. Current ML classification methods, used in biomedical applications, have shown that supervised algorithms, whether regression or classification, such as GNB, DT, KNN or SVM, usually have higher average accuracy than their unsupervised counterparts.<sup>39–40</sup> In addition, individually applied methods are limited in their precision, but combination of methods, when applied correctly, can have higher overall classification precision, as is the case with the proposed XGB method.<sup>39–40</sup> In our study, the SVM and GNB methods performed the worst, with KNN the method that most closely approximates the precision values of the proposed method. This is in line with the results of studies describing these supervised ML algorithms in predicting mortality from COVID-19.<sup>41–44</sup> Different studies<sup>19–20</sup> using the XGB method for COVID-19 mortality prediction obtained accuracy values above 90%, as in our case, but in North American and Chinese populations. The number of variables included in these studies was much lower than in our study, and in addition pharmacological treatments, both before and during hospital stay, were not considered as variables of interest. Our study provides a similar radar plot between the training and test phases, indicating that the system does not lose much predictive capability. The results show that the proposed model can handle large

**Table 2** Comparison of XGB method as a classification in biomedical applications

|                                     | Accuracy | AUC  | Recall |
|-------------------------------------|----------|------|--------|
| Ma <i>et al</i> <sup>46</sup>       | 0.83     | 0.87 | 0.95   |
| Torlay <i>et al</i> <sup>47</sup>   | 0.93     | 0.91 | –      |
| Ogunleye <i>et al</i> <sup>50</sup> | 0.98     | 0.99 | 0.98   |
| Shi <i>et al</i> <sup>74</sup>      | 0.92     | –    | 0.92   |
| Yu <i>et al</i> <sup>51</sup>       | 0.97     | –    | 0.96   |
| Zhong <i>et al</i> <sup>52</sup>    | 0.78     | 0.78 | 0.55   |

AUC, area under the curve; XGB, eXtreme Gradient Boosting.

data dimensions, avoiding overtraining, and significantly improves the performance of other classification methods. It achieved higher values for precision, recovery and accuracy than those achieved by the other methods. This guarantees its reliability for the automatic classification of the desired result. XGB is a predictive model that has excellent scalability and high execution speed.<sup>45</sup> It has been applied in biomedicine (table 2) to classify patients with cancer,<sup>46</sup> epilepsy,<sup>47</sup> atrial fibrillation<sup>48</sup> and those at risk of hypertension,<sup>49</sup> and to diagnose chronic kidney disease.<sup>50</sup> Yu *et al*<sup>51</sup> and Zhong *et al*<sup>52</sup> took advantage of the XGB method to predict the location of submitochondrial and essential proteins in their respective work.

### Predictors of mortality and related work

In our study, the predictors of mortality, in order of weighting, were CRP, PCT, GOT, NIV days, neutrophils, GPT, D-dimer, creatinine, septic shock, age, lactic acid and ferritin. White cell counts and platelets were also weighted to a lesser degree. Of the patients, 52.7% were male and 65.5% of the total were  $\geq 65$  years old. Of the patients, 22.7% were deceased and 16.2% were admitted to the ICU, with both percentages higher than in other studies.<sup>18 53 54</sup> Consistent with other studies,<sup>16 17 20</sup> advanced age was the main demographic predictor of hospital mortality in patients with COVID-19. The study by Sánchez-Montañes *et al*<sup>18</sup> applied different ML methods, with age being the most important predictor of mortality. The systematic review by Zheng *et al*<sup>55</sup> included 3027 patients and showed age  $\geq 65$  years (OR 6.06, 95% CI 3.98 to 9.22) as the factor that was most associated with progression of COVID-19. Other authors that used other ML models, such as the artificial neural network<sup>56</sup> or the deep learning model,<sup>36</sup> also highlighted age as a predictor of progression to a severe/critical clinical picture of severity and/or mortality. The clinical predictors included the scores obtained on the APACHE II, SOFA and CURB-65 scales. Although all of these are useful in predicting mortality in patients with COVID-19,<sup>57</sup> in our study they did not have a significant weight. On the other hand, comorbidities such as diabetes and hypertension have been described as risk factors for poor prognosis and progression in patients with COVID-19.<sup>58 59</sup> In our study, we did not find an association between these comorbidities and mortality from COVID-19, as is the case in other studies.<sup>53</sup> History of cardiac comorbidities and CPK measurement were taken into account as a marker of cardiac dysfunction, unlike in other studies

which preferentially used elevated cardiac troponin as an indicator of cardiac injury.<sup>15 36</sup> SARS-CoV-2 interacts with the cardiovascular system on multiple levels and heart problems are associated with higher mortality in patients with COVID-19,<sup>15 36</sup> although in our study there was no association in this regard. Other predictors that were positively associated with mortality were septic shock and NIV days, as described in the systematic review by Adamidi *et al*.<sup>43</sup> Most of the studies in this review showed SpO<sub>2</sub> and respiratory failure as predictors of mortality instead of talking about patients with NIV. Elevated blood urea nitrogen (BUN) and D-dimer and lymphocytopenia were associated with extrapulmonary disorders and possible multiorgan damage caused by COVID-19,<sup>16</sup> all of which were a result of septic shock due to infection. The laboratory parameters were obtained after patients' admission. Those related to altered kidney function, such as BUN and serum creatinine, were associated with a worse prognosis in these patients,<sup>60</sup> similar to our case. Different studies have identified acute kidney injury (AKI) as a sequela in patients with severe COVID-19, many of whom died.<sup>61</sup> A Cox regression analysis showed that proteinuria, hematuria, and elevated BUN and creatinine levels, among other characteristics, were significantly associated with death of patients with COVID-19.<sup>62</sup> This analysis suggested that patients with COVID-19 who developed AKI are at risk of mortality ~5.3 times greater than those without AKI. As in our research, other studies using ML<sup>17 19 20 36 57 63</sup> identified the following laboratory parameters as predictors of severity and mortality: CRP, lactic acid, PCT, ferritin, D-dimer, GOT, GPT and neutrophils. PCT is elevated during bacterial infection, but less so during viral infection, suggesting that bacterial coinfection leads to worse outcomes in patients with COVID-19.<sup>36</sup> Elevated serum ferritin is associated with ARDS.<sup>64</sup> Wu *et al*<sup>65</sup> conducted a retrospective cohort study of 201 patients with COVID-19 and found that elevated serum ferritin was an independent risk factor related to the development of ARDS, but no similar association was observed in terms of mortality, possibly due to insufficient sample size. The meta-analysis of Henry *et al*<sup>66</sup> confirmed serum ferritin as a possible biomarker of progression to critical illness in patients with COVID-19. D-dimer has also been associated with mortality in patients with COVID-19.<sup>19 62</sup> This is a marker of hypercoagulability and thrombosis that has been found to be elevated in patients with COVID-19.<sup>19</sup> Concentrations greater than 1 µg/mL are associated with poor prognosis in the initial stages of the disease.<sup>53</sup> Elevated GOT levels due to liver dysfunction have been seen in severe cases of COVID-19.<sup>67</sup> Jiang *et al*<sup>63</sup> used supervised learning and found that elevation in GPT was predictive of severe ARDS in patients with COVID-19. Elevation of both enzymes and therefore liver disease are considered predictors of severity in these patients.<sup>68</sup> Finally, low levels of leukocytes and neutrophils have also been described as predictors of severity,<sup>69</sup> as well as thrombocytopenia described in critically ill patients with COVID-19.<sup>70</sup> The recent systematic review by Bottino *et al*<sup>44</sup> concludes, as does our study, that among the predictors most associated with mortality are age and CRP and LDH levels.



## XGB as a predictive model of mortality

XGB is the easiest binary classification method to implement and train, which means that as more data become available this algorithm will improve with respect to predictive performance.<sup>44</sup> Similarly, Sánchez-Salmerón *et al*<sup>71</sup> in their systematic review highlight the XGB method as one of the models that achieve the highest level of predictive accuracy and can be a good tool to aid the triage process of patients with COVID-19. Wan *et al*<sup>72</sup> in their recent study used the random forest classifier with very similar characteristics to the XGB and obtained similar results with both.

Comparative studies have revealed that ML methods can be more accurate and efficient than traditional logistic regression analysis, especially when the sample size is limited.<sup>73</sup> Including data from other modalities, such as genomic profiling and medical imaging, could further improve the predictive performance of the presented model. Since the length of hospital stay for most patients was greater than 1 week, our model can predict patients' outcome more than 1 week in advance.

## CONCLUSION

ML techniques are the most sophisticated and accurate tools for predicting events of interest in general and COVID-19 mortality prediction in particular. Of the five ML methods studied and validated, the XGB method obtained the highest accuracy in predicting hospital mortality due to COVID-19, with the following predictors of hospital mortality having the highest weight: nine biomarkers (CRP, PCT, GOT, GPT, neutrophils, D-dimer, creatinine, lactic acid and ferritin), days of NIV, septic shock and age. Other variables of interest were white cell counts and platelets. None of the pharmacological treatments included in the study had sufficient weight in predicting mortality for any of the models used.

The XGB method achieves a prediction value of 92%, improving by 6.95% the results shown by KNN, the second better ML method. The XGB method will help health-care professionals in the process of stratifying cases and in making decisions about resource allocation and optimizing treatment for patients with COVID-19. This study will lay the groundwork for future multicenter studies with large inpatient and home-based populations. The results of this work will facilitate implementation of optimal economic and socio-health policies.

**Contributors** AR, AMT, JMi, JC, PB and JMa contributed to the design of the work, and to the acquisition, analysis and interpretation of data for the work. AR, AMT, JMi, JC, PB and JMa contributed to the drafting of the work and revising it critically for important intellectual content. AR, AMT, JMi, JC, PB and JMa agreed on all aspects of the work related to the accuracy or integrity of any part of the work. AR, AMT, JMi, JC, PB and JMa contributed to the final approval of the version to be published. AR is responsible for the overall content as guarantor.

**Funding** This work was sponsored by the General University Hospital of Valencia (Spain), Fondo Europeo de Desarrollo Regional (FEDER) and Instituto de Salud Carlos III (P120/01363; JMi), Centro de Investigaciones Biomédicas en Red de Enfermedades Respiratorias (CIBERES) (CB06/06/0027; JMi), and the Institute of Technology (University of Castilla-La Mancha).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study involves human participants and was approved by the General University Hospital of Valencia. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request.

This article is made freely available for personal use in accordance with BMJ's website terms and conditions for the duration of the covid-19 pandemic or until otherwise determined by BMJ. You may download and print the article for any lawful, non-commercial purpose (including text and data mining) provided that all copyright notices and trade marks are retained.

## ORCID iD

Jorge Mateo <http://orcid.org/0000-0003-2453-3058>

## REFERENCES

- Chen N, Zhou M, Dong X, *et al*. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507–13.
- Liu Y, Gayle AA, Wilder-Smith A, *et al*. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med* 2020;27:taaa021.
- Wang D, Hu B, Hu C, *et al*. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020;323:1061–9.
- Coronavirus W. Dashboard— who coronavirus (COVID-19) Dashboard with vaccination data, 2022. Available: <https://covid19.who.int/>
- Ji D, Zhang D, Xu J, *et al*. Prediction for progression risk in patients with COVID-19 pneumonia: the call score. *Clin Infect Dis* 2020;71:1393–9.
- Liang W, Liang H, Ou L, *et al*. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med* 2020;180:1081–9.
- Yang HJ, Zhang YM, Yang M. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2. *ERJ* 2020;56.
- Yadaw AS, Li Y-C, Bose S, *et al*. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health* 2020;2:e516–25.
- Vollmer S, Mateen BA, Bohner G, *et al*. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ* 2020;368:l6927.
- Motwani M, Dey D, Berman DS, *et al*. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;38:500–7.
- Rajkomar A, Oren E, Chen K, *et al*. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018;1:1–10.
- Pourghasemi HR, Pouyan S, Heidari B, *et al*. Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *Int J Infect Dis* 2020;98:90–108.
- Mollalo A, Rivera KM, Vahedi B. Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States. *Int J Environ Res Public Health* 2020;17:4204.
- Sakagiani A, Feretakis G, Kalles D, *et al*. Setting up an easy-to-use machine learning pipeline for medical decision support: a case study for COVID-19 diagnosis based on deep learning with CT scans. *Stud Health Technol Inform* 2020;272:13.
- McRae MP, Simmons GW, Christodoulides NJ, *et al*. Clinical decision support tool and rapid point-of-care platform for determining disease severity in patients with COVID-19. *Lab Chip* 2020;20:2075–85.
- Gao Y, Cai G-Y, Fang W, *et al*. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun* 2020;11:1–10.
- Wang T, Paschalidis A, Liu Q, *et al*. Predictive models of mortality for hospitalized patients with COVID-19: retrospective cohort study. *JMIR Med Inform* 2020;8:e21788.
- Sánchez-Montañés M, Rodríguez-Belenguier P, Serrano-López AJ, *et al*. Machine learning for mortality analysis in patients with COVID-19. *Int J Environ Res Public Health* 2020;17:8386.
- Vaid A, Somani S, Russak AJ, *et al*. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York City: model development and validation. *J Med Internet Res* 2020;22:e24018.
- Guan X, Zhang B, Fu M, *et al*. Clinical and inflammatory features based machine learning model for fatal risk prediction of hospitalized COVID-19 patients: results from a retrospective cohort study. *Ann Med* 2021;53:257–66.
- Syeda HB, Syed M, Sexton KW, *et al*. Role of machine learning techniques to tackle the COVID-19 crisis: systematic review. *JMIR Med Inform* 2021;9:e23811.



- 22 Chen C, Dong D, Qi B, *et al.* Quantum ensemble classification: a sampling-based learning control approach. *IEEE Trans Neural Netw Learn Syst* 2017;28:1345–59.
- 23 Chang W, Liu Y, Wu X, *et al.* A new hybrid XGBSVM model: application for hypertensive heart disease. *IEEE Access* 2019;7:175248–58.
- 24 Chen W, Fu K, Zuo J, *et al.* Radar emitter classification for large data set based on weighted-xgboost. *IET Radar, Sonar & Navigation* 2017;11:1203–7.
- 25 Mateo J, Rius-Peris JM, Marañón-Pérez AI, *et al.* Extreme gradient boosting machine learning method for predicting medical treatment in patients with acute bronchiolitis. *BioCybern Biomed Eng* 2021;41:792–801.
- 26 Que Z, Xu Z. A data-driven health prognostics approach for steam turbines based on xgboost and dtw. *IEEE Access* 2019;7:93131–8.
- 27 Rivera-Lopez R, Canul-Reich J. Construction of near-optimal axis-parallel decision trees using a differential-evolution-based approach. *IEEE Access* 2018;6:5548–63.
- 28 Sharmila A, Geethanjali P. DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers. *IEEE Access* 2016;4:7716–27.
- 29 Das BK, Dutta HS. GFNB: Gini index-based fuzzy naive Bayes and blast cell segmentation for leukemia detection using multi-cell blood smear images. *Med Biol Eng Comput* 2020;58:2789–803.
- 30 Zhang S, Li X, Zong M, *et al.* Efficient knn classification with different numbers of nearest neighbors. *IEEE Trans Neural Netw Learn Syst* 2018;29:1774–85.
- 31 Xing W, Bei Y. Medical health big data classification based on KNN classification algorithm. *IEEE Access* 2019;8:28808–19.
- 32 Yu S, Li X, Zhang X, *et al.* The OCS-SVM: an objective-cost-sensitive SVM with sample-based misclassification cost invariance. *IEEE Access* 2019;7:118931–42.
- 33 Kafai M, Eshghi K. CROification: accurate kernel classification with the efficiency of sparse linear SVM. *IEEE Trans Pattern Anal Mach Intell* 2019;41:34–48.
- 34 Zhou X, Obuchowski NA, McClish DK. *Statistical methods in diagnostic medicine*. 2nd ed. John Wiley and Sons, 2011.
- 35 Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006;27:861–74.
- 36 Li X, Ge P, Zhu J, *et al.* Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables. *PeerJ* 2020;8:e10337.
- 37 Xie J, Covassin N, Fan Z, *et al.* Association between hypoxemia and mortality in patients with COVID-19. *Mayo Clin Proc* 2020;95:1138–47.
- 38 Wynants L, Van Calster B, Collins GS, *et al.* Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328.
- 39 Han J, Pei J, Kamber M. *Data mining: concepts and techniques*. 3rd ed. Morgan Kaufmann is an imprint of Elsevier, 2016.
- 40 Azevedo A. Data mining and knowledge discovery in databases. In: *Advanced methodologies and technologies in network architecture, mobile computing, and data analytics*, 2019: 502–14.
- 41 Mahdavi M, Choudhary H, Zabe H, *et al.* A machine learning based exploration of COVID-19 mortality risk. *PLoS One* 2021;16:e0252384.
- 42 Pourhomayoun M, Shakibi M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health* 2021;20:100178.
- 43 Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. *Comput Struct Biotechnol J* 2021;19:2833–50.
- 44 Bottino F, Tagliente E, Pasquini L, *et al.* COVID mortality prediction with machine learning methods: a systematic review and critical appraisal. *J Pers Med* 2021;11:893.
- 45 Chen T, Guestrin C. Xgboost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on knowledge discovery and data mining; Aug 2016, 2016:785–94.
- 46 Ma B, Meng F, Yan G, *et al.* Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. *Comput Biol Med* 2020;121:103761.
- 47 Torlay L, Perrone-Bertolotti M, Thomas E, *et al.* Machine learning-XGBoost analysis of language networks to classify patients with epilepsy. *Brain Inform* 2017;4:159–69.
- 48 Sodmann P, Vollmer M, Nath N, *et al.* A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms. *Physiol Meas* 2018;39:104005.
- 49 Ye C, Fu T, Hao S, *et al.* Prediction of incident hypertension within the next year: prospective study using statewide electronic health records and machine learning. *J Med Internet Res* 2018;20:e22.
- 50 Ogunleye A, Wang Q-G, Qing-Guo W. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinform* 2020;17:2131–40.
- 51 Yu B, Qiu W, Chen C, *et al.* SubMito-XGBoost: predicting protein submitochondrial localization by fusing multiple feature information and eXtreme gradient boosting. *Bioinformatics* 2020;36:1074–81.
- 52 Zhong J, Sun Y, Peng W, *et al.* XGBFEMF: an XGBoost-based framework for essential protein prediction. *IEEE Trans Nanobioscience* 2018;17:243–50.
- 53 Rechtman E, Curtin P, Navarro E, *et al.* Vital signs assessed in initial clinical encounters predict COVID-19 mortality in an NYC hospital system. *Sci Rep* 2020;10:1–6.
- 54 Izquierdo JL, Ancochea J, *et al.* Savana COVID-19 Research Group. Clinical characteristics and prognostic factors for intensive care unit admission of patients with COVID-19: retrospective study using machine learning and natural language processing. *J Med Internet Res* 2020;22:e21801.
- 55 Zheng Z, Peng F, Xu B, *et al.* Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect* 2020;81:e16–25.
- 56 Abdulaal A, Patel A, Charani E, *et al.* Prognostic modeling of COVID-19 using artificial intelligence in the United Kingdom: model development and validation. *J Med Internet Res* 2020;22:e20259.
- 57 Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Mod Pathol* 2021;34:522–31.
- 58 Lippi G, Wong J, Henry BM. Hypertension and its severity or mortality in coronavirus disease 2019 (COVID-19): a pooled analysis. *Pol Arch Intern Med* 2020;130:304–9.
- 59 Guo W, Li M, Dong Y, *et al.* Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes Metab Res Rev* 2020;36:e3319.
- 60 Cheng Y, Luo R, Wang K, *et al.* Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int* 2020;97:829–38.
- 61 Zhou F, Yu T, Du R, *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
- 62 Li Z, Wu M, Yao J, *et al.* Caution on kidney dysfunctions of COVID-19 patients. *medRxiv* 2020.
- 63 Jiang X, Coffee M, Bari A, *et al.* Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Comput Mater Contin* 2020;62:537–51.
- 64 Connelly KG, Moss M, Parsons PE, *et al.* Serum ferritin as a predictor of the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 1997;155:21–5.
- 65 Wu C, Chen X, Cai Y, *et al.* Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med* 2020;180:934–43.
- 66 Henry BM, de Oliveira MHS, Benoit S, *et al.* Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis. *Clin Chem Lab Med* 2020;58:1021–8.
- 67 Guan W-J, Ni Z-Y, Hu Y, *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708–20.
- 68 Haimovich AD, Ravindra NG, Stoytchev S, *et al.* Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. *Ann Emerg Med* 2020;76:442–53.
- 69 Sun L, Song F, Shi N, *et al.* Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19. *J Clin Virol* 2020;128:104431.
- 70 Lippi G, Plebani M, Henry BM. Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a meta-analysis. *Chimica Acta* 2020;506:145–8.
- 71 Sánchez-Salmerón R, Gómez-Urquiza JL, Albendín-García L, *et al.* Machine learning methods applied to triage in emergency services: a systematic review. *Int Emerg Nurs* 2022;60:101109.
- 72 Wan T-K, Huang R-X, Tulu TW, *et al.* Identifying predictors of COVID-19 mortality using machine learning. *Life* 2022;12:547.
- 73 Fernández-Delgado M, Cernadas E, Barro Senén. Do we need hundreds of classifiers to solve real world classification problems? *JMLR* 2014;15:3133–81.
- 74 Shi H, Wang H, Huang Y, *et al.* A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Comput Methods Programs Biomed* 2019;171:1–10.