Computer Methods and Programs in Biomedicine 233 (2023) 107492

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

CVD22: Explainable artificial intelligence determination of the relationship of troponin to D-Dimer, mortality, and CK-MB in COVID-19 patients

Kevser Kübra Kırboğa[a,b,*], Ecir Uğur Küçüksille[c], Muhammet Emin Naldan[d], Mesut Işık[a], Oktay Gülcü[e], Emrah Aksakal[e]

[a] Bilecik Seyh Edebali University, Bioengineering Department, 11230, Bilecik, Turkey
[b] Informatics Institute, Istanbul Technical University, Maslak, Istanbul, 34469, Turkey
[c] Süleyman Demirel University, Engineering Faculty, Department of Computer Engineering, Isparta 32260, Turkey
[d] Bilecik Seyh Edebali University, Faculty of Medicine, Department of Anaesthesiology and Reanimation, 11230, Bilecik, Turkey
[e] Health Sciences University, Erzurum City Hospital, Department of Cardiology, Erzurum, Turkey

ARTICLE INFO

Article history:
Received 4 November 2022
Revised 6 February 2023
Accepted 15 March 2023

Keywords:
COVID-19
Coronavirus, Troponin
creatine kinase
explainable artificial intelligence

ABSTRACT

Background and purpose: COVID-19, which emerged in Wuhan (China), is one of the deadliest and fastest-spreading pandemics as of the end of 2019. According to the World Health Organization (WHO), there are more than 100 million infectious cases worldwide. Therefore, research models are crucial for managing the pandemic scenario. However, because the behavior of this epidemic is so complex and difficult to understand, an effective model must not only produce accurate predictive results but must also have a clear explanation that enables human experts to act proactively. For this reason, an innovative study has been planned to diagnose Troponin levels in the COVID-19 process with explainable white box algorithms to reach a clear explanation.

Methods: Using the pandemic data provided by Erzurum Training and Research Hospital (decision number: 2022/13-145), an interpretable explanation of Troponin data was provided in the COVID-19 process with SHApley Additive exPlanations (SHAP) algorithms. Five machine learning (ML) algorithms were developed. Model performances were determined based on training, test accuracies, precision, F1-score, recall, and AUC (Area Under the Curve) values. Feature importance was estimated according to Shapley values by applying the SHApley Additive exPlanations (SHAP) method to the model with high accuracy. The model created with Streamlit v.3.9 was integrated into the interface with the name CVD22.

Results: Among the five-machine learning (ML) models created with pandemic data, the best model was selected with the values of 1.0, 0.83, 0.86, 0.83, 0.80, and 0.91 in train and test accuracy, precision, F1-score, recall, and AUC values, respectively. As a result of feature selection and SHApley Additive exPlanations (SHAP) algorithms applied to the XGBoost model, it was determined that DDimer mean, mortality, CKMB (creatine kinase myocardial band), and Glucose were the features with the highest importance over the model estimation.

Conclusions: Recent advances in new explainable artificial intelligence (XAI) models have successfully made it possible to predict the future using large historical datasets. Therefore, throughout the ongoing pandemic, CVD22 (https://cvd22covid.streamlitapp.com/) can be used as a guide to help authorities or medical professionals make the best decisions quickly.

1. Introduction

More than 239.4 million COVID-19 cases had been officially confirmed as of October 15, 2021, with more than 4.8 million deaths [1]. The SARS-CoV-2 virus is the infectious agent that causes acute SARS-CoV-2 respiratory illness, which belongs to the *Coronaviridae* family. As of August 2022, an epidemic still exists. The COVID-19 disease started with symptoms such as fever, muscle, sore throat, hoarseness, and cough, and then caused pneumonia by keeping the lungs serious. Pneumonia was the cause of 91% of hospitalizations during the pandemic period [2]. In addition to the respiratory system, the cardiovascular system was also affected, and studies showed that the group with the highest mortality rate was

* Corresponding author.
E-mail address: kubra.kirboga@yahoo.com (K.K. Kırboğa).

https://doi.org/10.1016/j.cmpb.2023.107492
0169-2607/

those with known cardiovascular diseases (CVD) [3]. The death rate was 2.3% for those aged 30-79, most of whom were aged 30-79, and 10.5% for those with concomitant cardiovascular disease [4].

In the follow-up of COVID-19, the following three effects were observed on CVD:

1. In patients with concomitant CVD disease, the clinical course has worsened, and the risk of death has increased.
2. COVID-19 is associated with myocardial injury, arrhythmia, and venous thromboembolism. The drugs used in the treatment of COVID-19 have various side effects on CVD.

High-sensitivity cardiac troponin (hs-cTn) levels were increased in most COVID-19 patients [4]. The rate of hs-cTn >28 in cases that ended in death was 46%, while it was 1% in surviving circumstances [5]. This is crucial in showing the prognostic importance of hs-cTn in hospitalized cases. While signs of myocardial damage (such as higher blood cardiac Troponin) have been linked to increased mortality, particularly in individuals with underlying cardiovascular illness, pulmonary symptoms are primarily COVID-19 manifested [6]. Blood usually does not contain Troponin. Instead, Troponin enters the circulation when the heart muscles are damaged. The higher amounts of Troponin released into the circulation, the more heart damage occurs.

In this research, we propose a white-box model to determine the effects behind the higher or lower-than-average troponin value with the help of explainable artificial intelligence (XAI) based on data from patients suffering from COVID-19 in Turkey Erzurum Regional Education and Research Center in March 2020 and March 2021 (decision number: 2022/13-145). Our first attempt was to organize the data between 0 and 0.04 ng/mL and above by collecting the data and evaluating the troponin levels. Secondly, we labeled those with troponin levels higher than 0.04 ng/mL as one and those with less than 0 ng/mL as 0. The third was to determine the importance of the feature affecting the troponin value using the test and training sets. Finally, in the fourth stage, testing and training were to determine the accuracy, F-1 score, recall score, precision score, and AUC (Area Under the Curve) values on the five models and select the forecast model with the highest accuracy rate. In the final stage, XAI was applied to this model.

1.1. Machine learning for COVID-19

Artificial intelligence (AI) -based models are one of the approaches that give great hope to doctors in detecting COVID-19-positive cases earlier. Models that effectively increase the accuracy estimation and reduce the workload contribute positively to the definitive treatment process. AI-based models are used in many areas, such as identifying viruses, diagnosing diseases, and repositioning drugs [8]. Various studies have been done on disease prediction with feature selection [7,8], mortality prediction by cardiac indicators [9–11], and prediction of SARS-CoV-2 infection with various biomarkers [12].

1.2. Explainable artificial intelligence (XAI)

Especially in medical situations, the explainability and transparency of AI-based methods have significant limitations. First, it is important for medical professionals to trust the models and know how predictive models make their decisions. For example, in health, "What makes this prediction reliable?" And how did this smart model arrive at that conclusion? XAI methods, based on post-hoc methods, emphasize the steps of ML models in the decision-making process and thus the features that contribute to model accuracy and prediction result during model training, and also aim to provide a traceable and reliable intelligent model [13]. While some prediction techniques, such as decision trees, are transparent, most AI applications that use deep learning (DL) techniques in medicine are black boxes and therefore have no explanation for their predictions. As every ML model in medical interventions needs diagnosis and advice [14], a rapidly growing field, a justifiable, transparent, interpretable, reliable, and traceable smart model provides human-understandable explanations for the behavior of XAI models. Studies aiming to contribute to the interpretability of results with visual graphics [15], XAI allows a clearer explanation of the relationship and correlation between cause and effect in the diagnosis and prediction process.

There are several limitations on how effective explanations may be when used in global AI processes like model creation and knowledge discovery. The first of these restrictions is that readers must interpret each sentence for themselves. When a model is imprecise or even wholly untrained, it might be challenging to believe that a technique still has the power to explain something. The model is trained to recognize the prototype elements of each class and then measure how much of each component it describes for the choice made in generally interpretable ways like SHAP. Another disadvantage is that complicated and irregular input data can significantly alter the output and produce entirely different conclusions, which can have significant clinical medical ramifications [13,16]. For these reasons, more careful research and improvement of existing methods are necessary. Currently, these techniques cannot guarantee that a specific choice was made correctly, but they also cannot boost user trust or support the adoption of AI suggestions in therapeutic practice [16].

1.3. Related works

Predictions combining various features have been created to predict the risk of infection in the COVID-19 process. For example, a model that predicts COVID-19 test results with high accuracy has recently been created using eight binary features. Zoabi et al. have developed a model that detects cases of COVID-19 with simple features accessed by asking basic questions based on data made public by the Israeli Ministry of Health nationwide [17]. Comprehensive survey studies have been conducted to inspire the COVID-19 process, especially in research on the combination of health and medicine with AI. The relationship of AI with the pandemic process has been examined from drug and vaccine development, fetal head ultrasound image, disease detection and diagnosis, epidemiology, pathogenesis, epidemic and transmission prediction data [18–21].

Tostmann et al. presented their findings from a large group of symptomatic healthcare workers who had been tested for SARS-CoV-2 infection as part of the hospital program. They aimed to identify the symptoms associated with a positive test of the COVID-19 virus and to identify the symptoms associated with SARS-CoV-2 based on early signs. In addition, they aimed to develop a diagnostic model to predict infection [22].

Dipto et al. have presented a fast and effective solution to detect COVID-19 using CT scan images. They first collected data using various architectures, built an ensemble model with the three best classifiers, and found the model with the best overall accuracy of 90.89% from COV19EXAI V1 and from COV19EXAI V2. Then, they integrated XAI into their model to better understand classification [23].

Zou et al. used a new image annotation method called Ensemble XAI based on SHAP and Grad-CAM++ methods for their study on COVID-19 and XAI. In this way, they provided a visual explanation for the DL predictive model that predicts the risk of death of patients with community-acquired pneumonia and COVID-19 respiratory tract infections [24]. Similarly, in our study, troponin level estimation of infected patients who encountered COVID-19 in Turkish society was made.

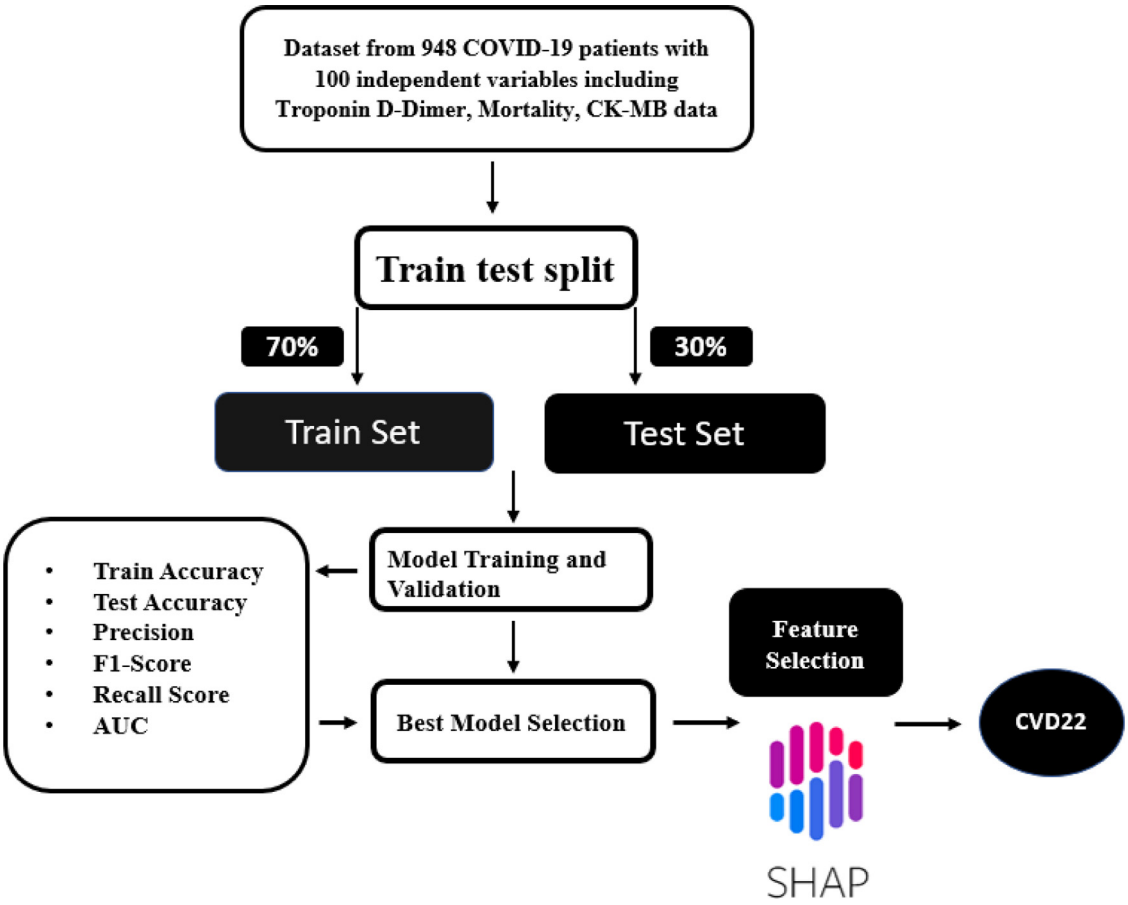


Fig. 1. Visualization of the method applied in this study.

2. Methods

Five classification ML models were created to determine the situation where the Troponin value is above the average in the data obtained from patients with COVID-19 and the relationship between it and other biomarkers. Selection among these models was determined by accuracy scores (Table 2). The interpretability and explainability of black box models were also addressed using the SHapley Additive exPlanations (SHAP) approach. Python version 3.9.7 was employed to carry out the investigation. Matplotlib version 3.4.3, Sklearn version 0.24.2, Pandas version 1.3.4, and NumPy version 1.23.1 were the libraries used throughout the implementation. Fig. 1

2.1. Random forest

A random forest classifier (RF) combines several essential decision trees and is an ensemble strategy. In regression problems, it decides with individual trees, while in classification problems, it decides by the average of the results of individual trees [25]. It calculates the average of many decision trees and uses bagging eight bootstrapping aggregations to mitigate variance. Breiman observed that trees are deeply correlated in a collection of decision trees when the algorithm can select any available features in the tree-growing procedure [26].

2.2. XGBoost

A distributed gradient boosting library, XGBoost was created to be exceptionally efficient, flexible, and portable. XGBoost, commonly known as GBDT or GBM, is a parallel tree extension technique that can quickly and accurately handle various data science problems. XGBoost offers superior results than other ML methods. Since it was initially created, it has matured into the "state-of-the-art" ML method for managing structured data. The base score is the first stage in XGBoost. This estimate can be any value because it will converge with the actions made in the subsequent phases to produce the right outcome.

2.3. Support vector machine (SVM)

Statistical learning theory provided the fundamental concepts for constructing support vector machines (SVM) [27]. The SVM delivers a decision boundary from the classes that is the ideal geometric margin hyperplane, which has the maximum generalization ability when considering a binary linearly separable classification issue. By using a technique known as the "kernel trick," this idea may be expanded to a nonlinear separable issue. To estimate the loss function, an uncertainty value rises as the sample approaches the decision boundary [28].

2.4. Decision tree classifier (DTC)

A Decision Tree (DT) is a supervised learning method that may be used to solve classification and regression issues. However, it works best when dealing with categorization issues [29]. It is a tree-structured classifier where core nodes represent the properties of a dataset, branches represent judgment calls, and each leaf node represents the classification outcome. The two decision tree nodes are the decision node and the leaf node. Decision nodes

Table 1
Computational Complexity of Machine Learning algorithms in our study.

Algorithm	Application	Training	Classification	Ref
Decision Tree (DT)	Classification	O(n2p)	O(p)	[31,32]
RandomForest (RF)	Regression/Classification	O(n2pntrees)	O(pntrees)	[31,32]
XGBoost	Regression/Classification	O(npntrees)	O(pntrees)	[31,32]
Multi-layer Perceptron Classifier (MLP)	Regression/Classification	-	O(pnl1+nl1nl2+...)	[33]
Support Vector Machine (SVM)	Regression/Classification	O(n2p+n3)	O(nsvp)	[31,34]

n:number of training sample. **p**:number of features. **ntrees**: the number of trees. **nsv**: the number of support vectors. **nli**: the number of neurons at layer **i** in a neural network.

Table 2
Accuracy scores were used in the selection of models.

Model	Train Accuracy	Test Accuracy	Precision	F1-score	Recall Score	AUC
Random Forest (RF)	1.0	0.81	0.86	0.80	0.76	0.90
XGBoost	1.0	0.83	0.86	0.83	0.80	0.91
Support Vector Machine (SVM)	0.74	0.70	0.78	0.67	0.58	0.79
Decision Tree Classifier (DTC)	1.0	0.71	0.77	0.69	0.62	0.72
Multi-layer Perceptron Classifier (MLP)	0.64	0.64	0.85	0.50	0.36	0.64

are used to make decisions and have many branches, whereas leaf nodes are the outcomes of decisions and do not have any more branches. The provided dataset's characteristics are used to execute the test or make the judgments.

2.5. Multi-layer perceptron classifier (MLP)

The multi-layer perceptron (MLP) is used in addition to the feed-forward neural network. Its three layers are the hidden layer, the output layer, and the input layer. The input layer is where the signal is received for processing at the input layer. The output layer handles the required work, such as prediction and categorization. The MLP's exact computing engine comprises an undetermined number of hidden layers between the input and output layers. The backpropagation learning process trains the neurons of the MLP. Since MLPs are designed to approximate any continuous function, they may be used to address problems that are not linearly separable [30].

2.6. Computational complexity

Computational complexity measures how fast or slow the algorithm will run for the input size and the amount of extra memory you need to run your algorithm in ML applications. The computational complexity calculations of the algorithms we used are shown in Table 1.

2.7. Feature importance

The ability of input characteristics to predict a target variable is rated using feature significance techniques. Examples of the many diverse types and sources of feature importance scores include statistical correlation scores, coefficients produced as a component of linear models, decision trees, and permutation significance scores. In a predictive modeling project, feature significance scores are crucial because they serve as the foundation for dimensionality reduction and feature selection, which can improve a predictive model's performance on the data, model insight, and issue. When feature importance is applied, its function in a problem of predictive modeling, how feature importance is derived from linear models and decision trees, and how permutation feature importance scores are computed and evaluated are discussed [35].

2.8. SHapley additive exPlanations (SHAP)

The measurement of Shapley values was first proposed in a game theory [36]. In this theory, it has been shown that the payment concepts correspond to the model estimate and that these values represent the mean of all associated contributions [37,38]. Shapley values, which have an essential role in improving the explainability of ML models, enable the model to be evaluated without specifying the functional forms of the models. The functions and property variables in an ML model are as follows:

$$f(f(Xi)) \equiv \phi 0 + \sum Kk = 1 \phi k(Xi), \forall i = 1, \ldots, n \quad (2.1)$$

Where k denotes a single property variable, K denotes the total number of explanatory variables available; n is the total number of units that should be. $\phi \in RK$; $\phi k \in R$. Φk (Xi) is the Shapley values of local functions.

As a general definition, Shapley values respond by adding credibility to the model's complex decision-making process [39]. This way, it reveals more clearly how the model uses its features. SHAP aims to explain individual explanations based on Shapley values' cooperative game theory. For the prediction to be explained by the supposition that each feature value of the instance is a "player," Shapley values are utilized. These numbers inform the user of how evenly the "players" in the game are distributed. The average marginal contribution of a feature value overall potential coalitions is known as the Shapley value [39]. These attributes serve as "players," which are then utilized to determine if the payout distribution is equitable. Therefore, it must satisfy the local accuracy, missingness, and consistency properties of making predictions [38]. SHAP explains the black-box model's output by demonstrating how the model defines an instance's prediction and calculates each feature's contribution to the prediction. Mathematically, SHAP then explains each prediction as it provides the local accuracy of the represented characteristics, as shown in (2.1).

3. Experiments

The main stages of the experiments consisted of data collection, data preprocessing, model tuning, model evaluation, and interpretation and prediction of models. All tests were conducted using a free Jupyter Notebook/Anaconda system with a Windows operating system, an Intel i5 8300H 2.30GHz quad-core CPU, and 25GB of RAM. Using PIP packages, the methods were put into practice in Python 3.7.

3.1. Data acquisition

The dataset included probable COVID-19 patients registered at the Erzurum Regional Research Center in Turkey between March 2020 and March 2021 (decision number: 2022/13-145). It is important to remember that the data has significant restrictions because only suspected COVID-19 instances were used to construct it. People also typically exhibit symptoms or have come into touch with other sick individuals directly. As a result, it is unlikely that asymptomatic individuals will be detected in this data set.

3.2. Data preprocessing

The raw dataset includes more than five thousand patients, including the results of various examination types and biomarkers. The literature mentions that several trials with missing values may result from the doctor's conclusion that specific examinations are unnecessary because the patient did not show any apparent abnormalities [40,41]. Therefore, although we had reference values for all regular value features, we avoided data loading, excluded samples with missing values, and filtered the final dataset to 948 patient data. Finally, due to the lack of a control group of healthy individuals in this cohort and because every patient in the dataset has the potential to have COVID-19 symptoms, it is difficult to estimate mean or average values for these patients.

As for the target class, we used the Troponin test, whose value was very high in some patients and less than expected (between 0 and 0.04 ng/mL). This way, values higher than a particular value (greater than 0.04 ng/mL) and less than a certain value (less than 0 ng/mL) are considered 0. The total value of 1 for our current dataset is roughly 474, making it a balanced classification. To test the generalization capacity of the model, the final datasets were randomly divided into 70% training and 30% testing.

3.3. Model training and validation

Train accuracy, test accuracy, precision, F1-score, and recall score criteria were used in this study, which includes deciding on the diagnosis of troponin level through binary classification. At the same time, AUC and ROC analysis was performed to evaluate the performance in true positive, false positive, and risk groups and to reach balanced average accuracy values [42].

4. Results and discussion

The overall results in this study are derived from data from actual COVID-19 patients. Our methodology offers the first COVID-19 test screening based on straightforward clinical symptoms and indications. Improving clinical priorities might lessen the strain on health systems by enabling optimal administration of health resources throughout upcoming waves of the SARS-Cov-2 pandemic. Particularly crucial in resource-constrained emerging nations. All symptoms are self-reported. Therefore, a negative number for a symptom might indicate that it wasn't mentioned. Assessing the model's performance is crucial when there are more missing or unreported data than negative ones. We deleted negative values from our prospective test set and randomly picked negative reports for each of the five symptoms to mimic a less biased condition. The model continued to produce encouraging results when applied to these simulated test sets, boosting our faith in the approach.

To choose the most accurate prediction model from a pool of 5 models, test and train accuracy, precision, F1-score, and recall scores were calculated. To comprehend and interpret the prediction model more clearly, we included XAI into the XGBoost algorithm. For each model, a ROC curve and confusion matrix were created to enable the creation of an exhaustive sensitivity/specificity report. The XGBoost model provided a ratio closer to 1 than the other models, as evidenced by the area under the ROC curves (AUC) (0.91). Table 3 displays the confusion matrix and ROC curves for each of the five models.

4.1. Feature selection

The classification accuracies provided by the first experiment, based on accurate data, allow for a quantitative comparison of the applied methods and to test for statistically significant differences between results in the complete feature set compared to sub-selected datasets. The concept of "feature significance" refers to techniques for rating every input feature for a particular model; the ratings represent the "importance" of each feature. A higher rating indicates that the trait will influence the forecasting model utilized for a certain variable. Examining the impact of randomly rearranging each predictor variable on the model's accuracy may determine each variable's relative relevance.

After the feature importance process, the features that have a high effect on the troponin level are higher or lower than the average group. According to the results (Fig. 2), D-Dimer Average, creatine kinase isoenzyme (CK-MB), mortality, Atrial Fibrillation (AF), Total Protein-1, angiotensin-converting enzyme (ACE) inhibitor, and an angiotensin receptor blocker (ARB) (ACEI/ARB), Asthma, Platelet Distribution Width (PDW), and Procalcitonin levels are essential in changing troponin levels.

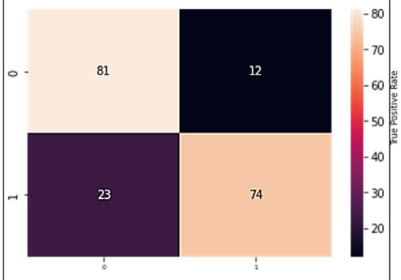
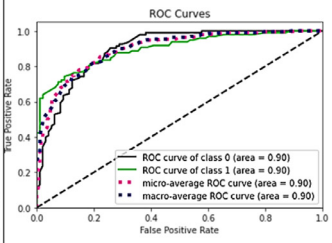
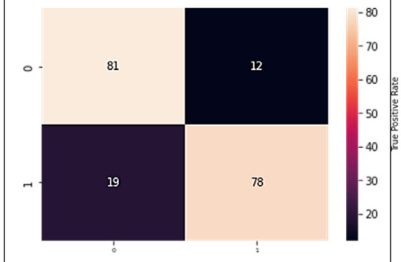
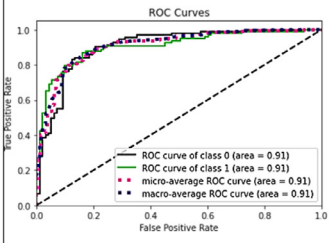
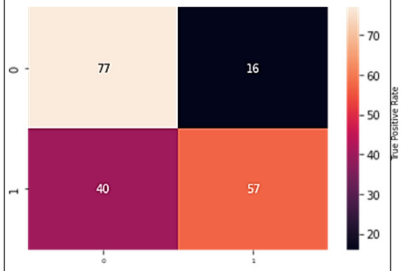
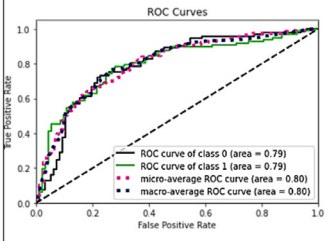
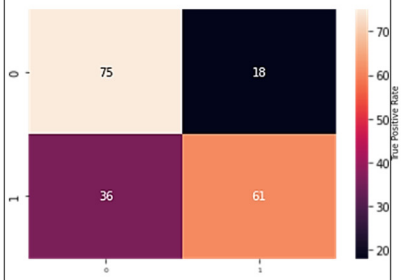
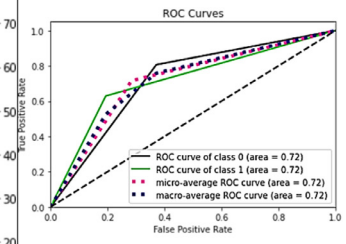
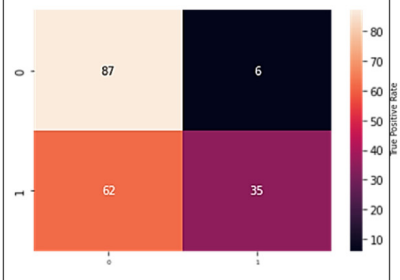
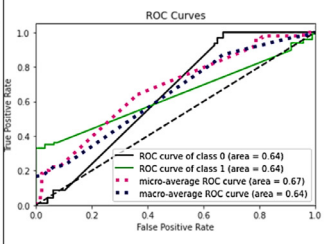
4.2. Interpretation and explanations

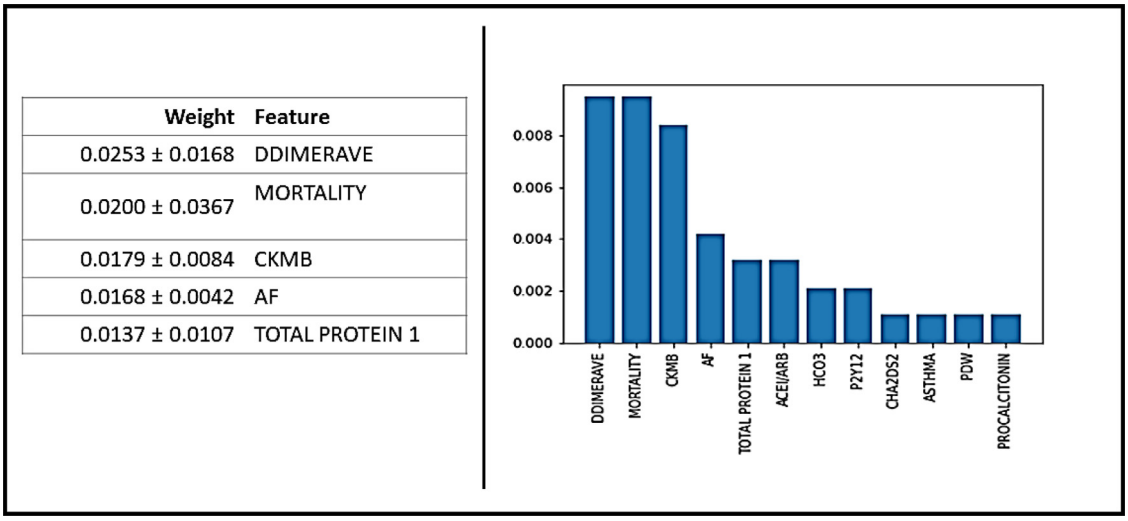
The objective of XAI is to provide a collection of interpretable models and approaches that produce models that are easier to comprehend while yet having high forecasting performance [41]. The explanation and interpretation of models were characterized by Samek et al. [43] as converting abstract ideas into a realm that people can comprehend. They also asserted that explanation is a group of characteristics of the interpretable environment that results in a decision being made in a certain situation. The idea of explanation in this study is consistent with Samek et al. Regional explanations reveal what causes the model to produce its current results while illuminating to scientists which features from the obtained data could more accurately represent the model. We may interpret these weight coefficients for these models as the weight the model accords a feature. Weights influence a positive diagnosis with positive values, and characteristics with negative values influence a negative diagnosis.

D-Dimer is one of the most important features that affect the troponin level to be higher than the upper limit of normal. As can be seen in the Waterfall graph, $E[f(x)] = 0.504$ has a more significant role to play when it comes to a higher value $(f(x)) = 0.79$ than the expected value. D-Dimer levels are elevated in patients with acute myocardial infarction, indicating the severity of the disease. However, the relationship between structural markers of myocardial damage and D-Dimer levels has not been well-studied. The correlation between myocardial markers, most clinical markers, and D-Dimer and Troponin markers can be predicted thanks to XAI and its graphics. However, as a result of the support of the waterfall graph, it is clear that the measurement of D-Dimer level in patients with acute myocardial infarction is correlated with Troponin [44].

In research on forecasting, force plot images are used as a general indication. Additionally, it backs up the Waterfall chart (Fig.3). D-Dimer average, CK-MB, and Glucose significantly impact troponin levels that are greater than usual. In hospitalized COVID-19 patients, the prevalence of myocardial damage increases and causes significant increases in some cardiac biomarkers [45]. A

Table 3
Confusion matrix and ROC curves of the five prediction models.

Models	Confusion Matrix	ROC Curve
Random Forest (RF)		
XGBoost		
Support Vector Machine (SVM)		
Decision Tree Classifier (DTC)		
Multi-layer Perceptron Classifier (MLP)		



Weight	Feature
0.0253 ± 0.0168	DDIMERAVE
0.0200 ± 0.0367	MORTALITY
0.0179 ± 0.0084	CKMB
0.0168 ± 0.0042	AF
0.0137 ± 0.0107	TOTAL PROTEIN 1

Fig. 2. The feature importance graph establishes the process of selecting a subset of the relevant features for use in model building. [DDIMERAVE: Average of DDimer values. MORTALITY: 0: died|1: lived. CKMB: Creatine phosphokinase. AF: Atrial fibrillation. TOTAL PROTEIN 1: Total protein value. ACEI/ARB: Angiotensin-converting enzyme (ACE)) inhibitors and angiotensin receptor blockers (ARB). HCO3: Bicarbonate concentration. P2Y12: Persistent aggregation receptor. CHA2DS2: Atrial fibrillation guideline stroke risk marker. ASTHMA: Asthma. RDW: Erythrocyte Distribution Width. PROCALCITONIN: Procalcitonin].

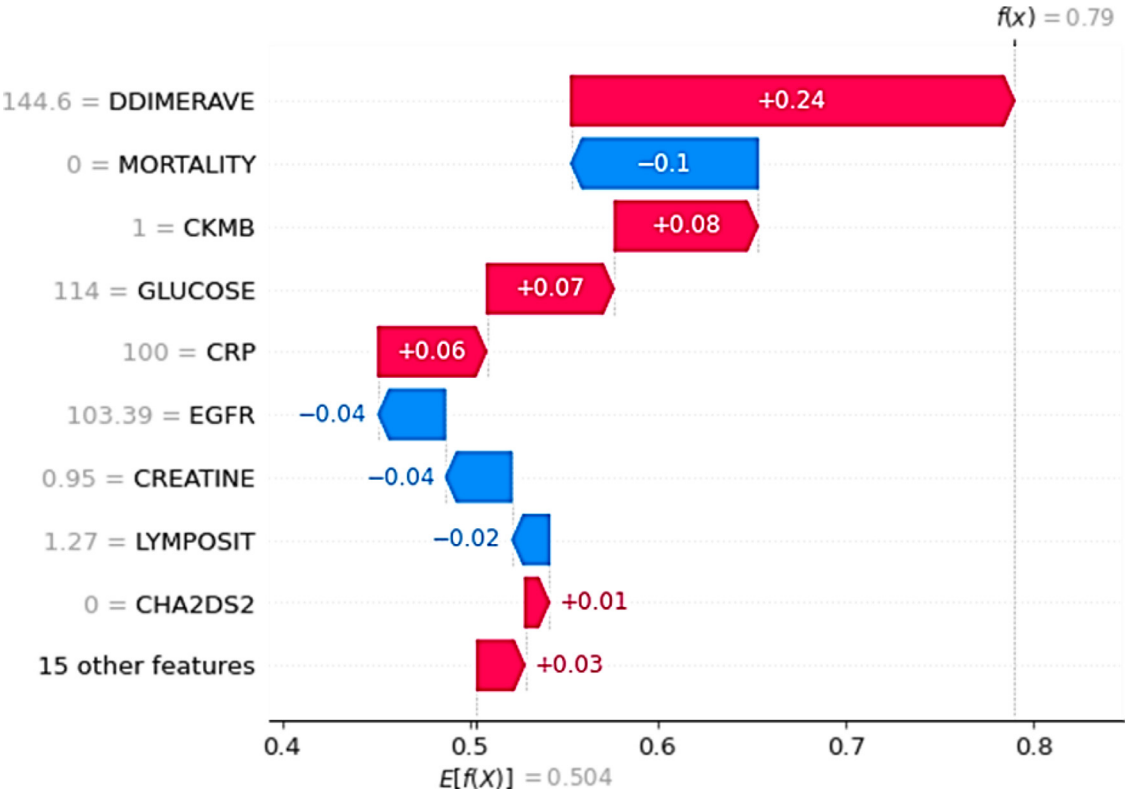


Fig. 3. XAI's waterfall graph integrated into the XGBoost model.

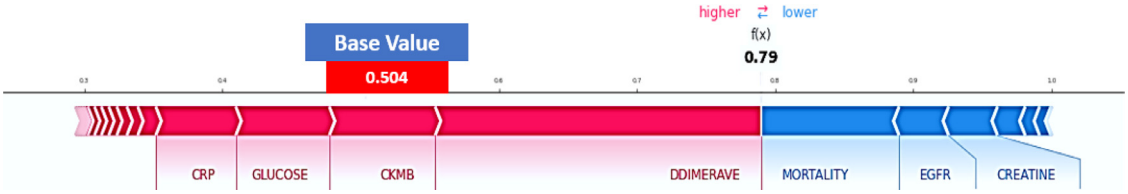


Fig. 4. XAI's force plot integrated into the XGBoost model. The features marked with red affect the troponin level to be lower than the normal level, while the features marked with blue affect the higher-than-normal level.

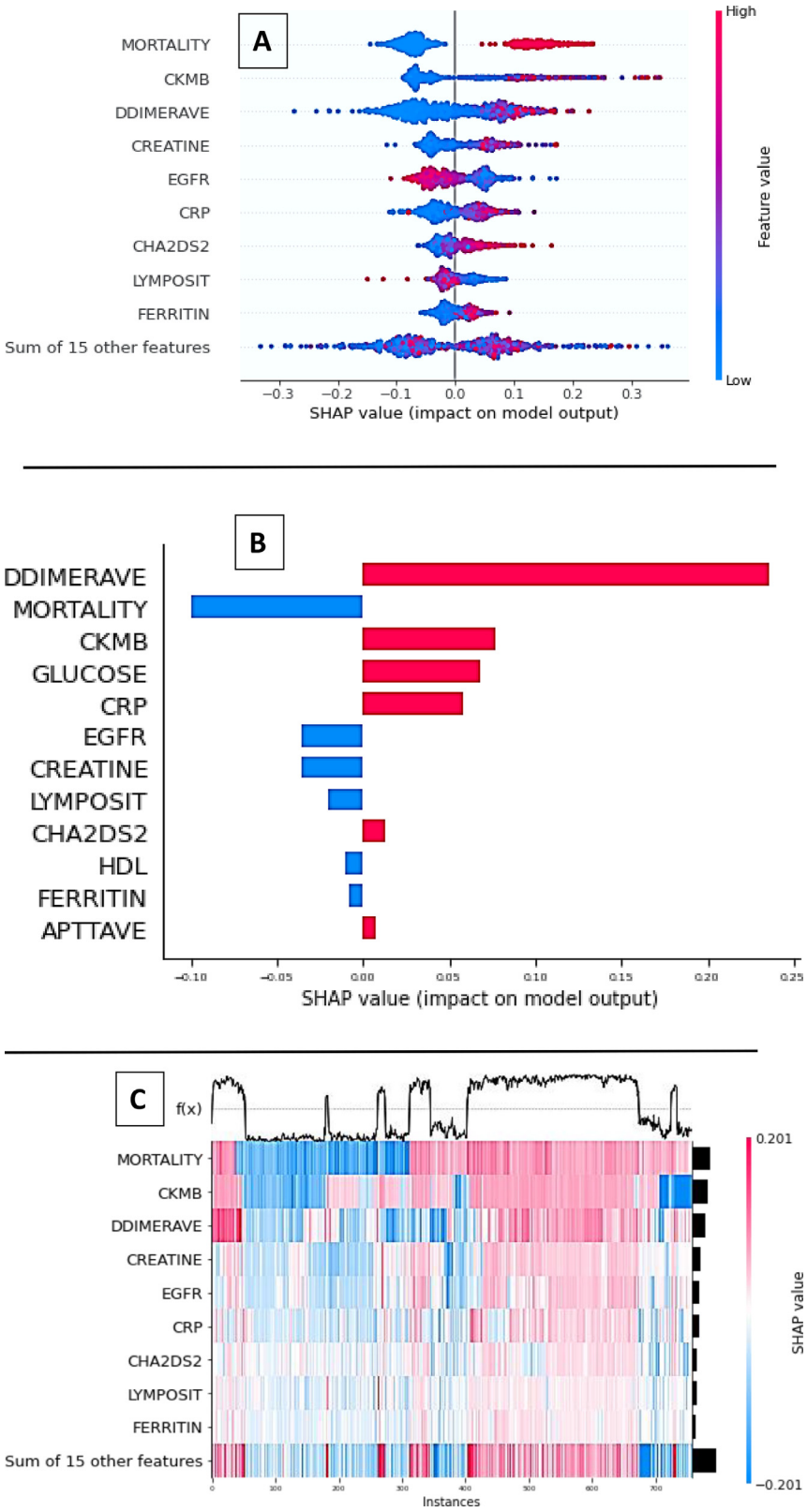


Fig. 5. XAI's beeswarm (A), boxplot (B), and heatmap (C) integrated into the XGBoost model.

study determined that laboratory findings such as D-dimer, C-reactive protein (CRP), total bilirubin, aspartate aminotransferase, and creatine increased significantly in the death and non-recovery groups. In addition, several cardiac biomarkers such as myoglobin, CK-MB, and high-sensitivity troponin T were significantly elevated in the groups; it was also stated that myoglobin was positively correlated with both CK-MB and high-sensitivity troponin T [46]. In another study, creatinine, sodium, Glucose, CRP, procalcitonin, AST, CK-MB, D-Dimer, LDH, white blood cells (WBC), and ferritin levels were higher in COVID-19 patients, determined to have cardiac

damage [47]. Among these biomarkers, high-sensitivity troponin T is one of the most important biomarkers showing the severity of myocardial damage, and numerous studies in the literature support this finding [48]. The studies showed that D-Dimer and CK-MB levels changed significantly in the death and non-recovery groups, and the study's findings support the results in the literature. Therefore, cardiac biomarkers may have effective potential in predicting COVID-19 patient prognosis.

An excellent measure for raising troponin levels is the CK-MB level. High CK-MB levels are strongly linked to a higher risk of death in COVID-19-infected individuals, according to meta-analysis research [49]. However, a low troponin level significantly impacts the decline in the mortality rate. Elevated serum troponin, a marker of mortality, is typically associated with acute respiratory tract infections and sepsis.

Additionally, the number of COVID-19 datasets in this study was not large enough for CVD22 to work optimally and still needs improvement to perform troponin value estimation more accurately. Therefore, we will continue to collect COVID-19 data, expand the dataset, and optimize CVD22 to improve prediction accuracy.

According to Fig. 5, a troponin level below normal also impacts the mortality rate. Kaura et al. provided more evidence that a positive troponin test, even if marginally over routine, is linked with clinically significant increased mortality regardless of age. In their research, the first few weeks saw a sharp increase in the severe mortality rate associated with elevated troponin levels [50].

The model developed in this study is based on an important innovation. Here are the elements that make it outperform many other prediction methods:

1. This model, which uses intravenous blood collection approaches and is created according to the method developed by expensive equipment, expert personnel, and certified laboratories, is a much faster, accessible, open interface, cheaper, and cost-effective routine computational approach.
2. This retrospective study proposes an AI decision system to provide doctors, nurses and other clinical workers with a simple and human-interpretable set of rules for diagnosing Troponin level in COVID-19 cases. It has provided a variety of XGBoost powered visuals to provide transparency and explainability for any clinician.
3. In many previous COVID-19 prediction models, the final prediction was made by a single classifier, but this study was done on multiple classifications. Therefore, the ability to generalize is superior.
4. Unrelated features in ordinary methods strongly affect the performance and diagnostic ability of the models. All irrelevant and irrelevant data should be removed so that computational complexity is reduced in an intelligent and explainable forecast. In our proposed CVD22 model, an additional feature selection step can be included and the Troponin diagnostic performance of the final model can be increased. As a result, the model can constantly update itself.
5. SHAP, the important algorithm we used in the study, performs all the details such as dependency and interaction between features consistently and quickly due to global model comments and a fast calculation.

Although the XGBoost Troponin level detection model has high performance, it also has some limitations. The first of these limitations is that its performance is affected by sparse and unstructured data. However, due to the big data feature of blood and COVID-19 data, it did not cause any problems in this study. Another limitation is its sensitivity to outliers. Because of this sensitivity, it constantly tries to correct the errors and has difficulty. It is common for outliers to appear in COVID-19 data. However, since identifying outliers in data entries is not very difficult, this should be considered when updating model data. The last limitation is that it is possible to deliberately create misleading comments with the SHAP method, which can affect prejudices. This situation may cause uncertainty about the explanations for healthcare professionals who read Shapley values.

In conclusion, this study showed that XAI is valuable in predicting the COVID-19 outbreak. However, we know that additional reliable data is needed to predict the COVID-19 process in Turkey accurately. Therefore, continuous data collection and exchange is necessary to prevent biases from affecting the ongoing epidemic. Using data from the Erzurum Regional and Research Center, we tried five alternative models to predict the low or high troponin level that experts have difficulty estimating. If testing resources are limited, and with the available data and the use of CVD22, our intended model can predict CVD, Troponin, and COVID-19. Python's Streamlit 3.9 was used to estimate the value of Troponin in medicine and health via CVD22. The address https://cvd22covid.streamlitapp.com/ was assigned to the proposed model.

Availability of data and materials

The datasets used and/or analysed during the current study can be made available from the corresponding author on reasonable request.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Statements and declarations

The authors have no relevant financial or non-financial interests to disclose.

The authors have no competing interests to declare relevant to this article's content. The authors state that no conflict of interest exists relevant to this article's content.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We are grateful to Turkey Erzurum Training and Research Hospital, which provided pandemic data with the approval of the ethics committee (decision number: 2022/13-145) within the scope of this study.

References

[1] WHO. "WHO coronavirus disease (COVID-19) dashboard." https://covid19.who.int/. 2021b. (accessed 2022).
[2] W.-j. Guan, et al., Clinical characteristics of coronavirus disease 2019 in China, N. Engl. J. Med. 382 (18) (2020) 1708–1720, doi:10.1056/NEJMoa2002032.
[3] C. Huang, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet North Am. Ed. 395 (10223) (2020) 497–506, doi:10.1016/S0140-6736(20)30183-5.
[4] D. Wang, et al., Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China, JAMA 323 (11) (Mar 17 2020) 1061–1069, doi:10.1001/jama.2020.1585.
[5] F. Zhou, et al., Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, Lancet North Am. Ed. 395 (10229) (2020) 1054–1062, doi:10.1016/S0140-6736(20)30566-3.
[6] T. Guo, et al., Cardiovascular implications of fatal outcomes of patients with coronavirus disease 2019 (COVID-19), JAMA Cardiol 5 (7) (2020) 811–818, doi:10.1001/jamacardio.2020.1017.

[7] K. Zhou, et al., Eleven routine clinical features predict COVID-19 severity uncovered by machine learning of longitudinal measurements, Comput. Struct. Biotechnol. J. 19 (2021) 3640–3649, doi:10.1016/j.csbj.2021.06.022.

[8] W. Cai, et al., CT quantification and machine-learning models for assessment of disease severity and prognosis of COVID-19 patients, Acad. Radiol. 27 (12) (2020) 1665–1678, doi:10.1016/j.acra.2020.09.004.

[9] M. Pourhomayoun and M. Shakibi, "Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making," (in eng), *Smart Health (Amst),* vol. 20, p. 100178, Apr 2021, doi:10.1016/j.smhl.2020.100178.

[10] C.D.K. Wungu, et al., Meta-analysis of cardiac markers for predictive factors on severity and mortality of COVID-19, Int. J. Infect. Dis. 105 (2021) 551–559, doi:10.1016/j.ijid.2021.03.008.

[11] M.M. Banoei, R. Dinparastisaleh, A.V. Zadeh, M. Mirsaeidi, Machine-learning-based COVID-19 mortality prediction model and identification of patients at low and high risk of dying, Crit. Care 25 (1) (2021) 1–14.

[12] V. Bayat, et al., A severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) prediction model from standard laboratory tests, Clin. Infect. Dis. 73 (9) (2021) e2901–e2907, doi:10.1109/access.2018.2870052.

[13] A. Holzinger, "The next frontier: AI we can really trust," 2021, pp. 427-440.

[14] M. Rostami, M. Oussalah, A novel explainable COVID-19 diagnosis method by integration of feature selection with random forest, Inform. Med. Unlocked 30 (2022) 100941 2022/01/01/, doi:10.1016/j.imu.2022.100941.

[15] D. Shin, The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI, Int. J. Hum. Comput. Stud. 146 (2021) 102551, doi:10.1016/j.ijhcs.2020.102551.

[16] M. Ghassemi, L. Oakden-Rayner, A.L. Beam, The false hope of current approaches to explainable artificial intelligence in health care, The Lancet Digital Health 3 (11) (2021) e745–e750 2021/11/01/, doi:10.1016/S2589-7500(21)00208-9.

[17] Y. Zoabi, S. Deri-Rozov, N. Shomron, Machine learning-based prediction of COVID-19 diagnosis based on symptoms, NPJ Digit. Med. 4 (1) (2021) 3 2021/01/04, doi:10.1038/s41746-020-00372-6.

[18] J. Chen, K. Li, Z. Zhang, K. Li, P.S. Yu, A survey on applications of artificial intelligence in fighting against COVID-19, ACM Comput. Surv. 54 (8) (2022) 1–32, doi:10.1145/3465398.

[19] Q.-V. Pham, D.C. Nguyen, T. Huynh-The, W.-J. Hwang, P.N. Pathirana, Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the State-of-the-Arts, IEEE Access 8 (2020) 130820–130839, doi:10.1109/access.2020.3009328.

[20] I. G. Pereira et al., "Epidemiology forecasting of COVID-19 using AI—a survey," *Comput. Intell. COVID-19 Future Pandemics,* 2022.

[21] L. Zhao, K. Li, B. Pu, J. Chen, S. Li, X. Liao, An ultrasound standard plane detection model of fetal head based on multi-task learning and hybrid knowledge graph, Future Generat. Comput. Syst. 135 (2022) 234–243 2022/10/01/, doi:10.1016/j.future.2022.04.011.

[22] A. Tostmann, et al., Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020, Euro Surveill. 25 (16) (2020) 2000508, doi:10.2807/1560-7917.ES.2020.25.16.2000508.

[23] S. M. Dipto, I. Afifa, M. Sagor, M. T. Reza, and M. A. Alam, "Interpretable COVID-19 classification leveraging ensemble neural network and XAI," 2021, pp. 380-391.

[24] L. Zou, et al., Ensemble image explainable AI (XAI) algorithm for severe community-acquired pneumonia and COVID-19 respiratory infections, IEEE Trans. Artif. Intell. (2022) 1, doi:10.1109/TAI.2022.3153754.

[25] L. Breiman, Random forests, Int. J. Mach. Learn. Cybern. 45 (1) (2001) 5–32, doi:10.1023/a:1010933404324.

[26] L. Breiman, Bagging predictors, Int. J. Mach. Learn. Cybern. 24 (2) (1996) 123–140, doi:10.1007/bf00058655.

[27] V.N. Vapnik, Methods of pattern recognition, in: V.N. Vapnik Ed (Ed.), The Nature of Statistical Learning Theory, Springer New York, New York, NY, 2000, pp. 123–180.

[28] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol. (TIST),* vol. 2, no. 3, pp. 1-27, 2011, doi:10.1145/1961189.1961199.

[29] P.H. Swain, H. Hauska, The decision tree classifier: design and potential, IEEE Trans. Geosci. Electron. 15 (3) (1977) 142–147, doi:10.1109/TGE.1977.6498972.

[30] S. Abirami and P. Chitra, "Chapter Fourteen - Energy-efficient edge based real-time healthcare support system," in *Advances in Computers*, vol. 117, P. Raj and P. Evangeline Eds.: Elsevier, 2020, pp. 339-368.

[31] H. Younes, M. Alameh, A. Ibrahim, M. Rizk, M. Valle, Efficient algorithms for embedded tactile data processing (2020) 113–138.

[32] Prashant, "Computational complexity of ML algorithms," ed, 2021.

[33] E. Mizutani and S. E. Dreyfus, *On complexity analysis of supervised MLP-learning for algorithmic comparisons*. 2001, pp. 347-352 vol.1.

[34] S. Ray, An analysis of computational complexity and accuracy of two supervised machine learning algorithms—K-nearest neighbor and support vector machine, in: Data Management, Analytics and Innovation: Proceedings of ICDMAI 2020, Volume 1, Springer, 2021, pp. 335–347.

[35] A. Altmann, L. Toloşi, O. Sander, T. Lengauer, Permutation importance: a corrected feature importance measure, Bioinformatics 26 (10) (2010) 1340–1347, doi:10.1093/bioinformatics/btq134.

[36] J. Derks, H. Peters, A shapley value for games with restricted coalitions, Int. J. Game Theory 21 (4) (1993) 351–360 [Online]. Available: https://EconPapers.repec.org/RePEc:spr:jogath:v:21:y:1993:i:4:p:351-60 .

[37] A. Joseph, *Shapley regressions: a framework for statistical inference on machine learning models*. 2019.

[38] S. Lundberg and S.-I. Lee, *A unified approach to interpreting model predictions*. 2017.

[39] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, W. Samek, in: Explainable AI Methods - A Brief Overview, Springer International Publishing, 2022, pp. 13–38.

[40] D. Schmidt, M. Niemann, and G. L. v. Trzebiatowski, "The handling of missing values in medical domains with respect to pattern mining algorithms," presented at the CS&P, 2015. [Online]. Available: https://www.semanticscholar.org/paper/The-Handling-of-Missing-Values-in-Medical-Domains-Schmidt-Niemann/cac63a50b6de96c2facf9ddcaf3b571421d32004.

[41] A. Adadi, M. Berrada, Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), IEEE Access 6 (2018) 52138–52160, doi:10.1109/access.2018.2870052.

[42] A.M. Carrington, et al., Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation, IEEE Trans. Pattern Anal. Mach. Intell. 45 (1) (2023) 329–341, doi:10.1109/TPAMI.2022.3145392.

[43] W. Samek, G. Montavon, A. Vedaldi, L.K. Hansen, K.-R. Müller, Explainable AI: interpreting, explaining and visualizing deep learning, Explainable AI: Interpret. Expl. Visual. Deep Learn. (2019), doi:10.1007/978-3-030-28954-6_1.

[44] M. İ, Correlation of D-dimer levels with troponin I in patients with acute myocardial infarction, Presented at the Correlation of D-Dimer levels with troponin I in patients with acute myocardial infarction, 2020 [Online]. Available: https://abstracts.isth.org/abstract/correlation-of-d-dimer-levels-with-troponin-i-in-patients-with-acute-myocardial-infarction/ .

[45] M. Madjid, P. Safavi-Naeini, S.D. Solomon, O. Vardeny, Potential effects of coronaviruses on the cardiovascular system: a review, JAMA Cardiol 5 (7) (2020) 831–840, doi:10.1001/jamacardio.2020.1286.

[46] J. Yang, et al., Elevated cardiac biomarkers may be effective prognostic predictors for patients with COVID-19: a multicenter, observational study, Am. J. Emerg. Med. 39 (Oct 13 2020) 34–41, doi:10.1016/j.ajem.2020.10.013.

[47] M. Tahir Huyut, Z. Huyut, F. İlkbahar, C. Mertoğlu, What is the impact and efficacy of routine immunological, biochemical and hematological biomarkers as predictors of COVID-19 mortality? Int. Immunopharmacol. 105 (Apr 2022) 108542, doi:10.1016/j.intimp.2022.108542.

[48] M. Aboughdir, T. Kirwin, A. Abdul Khader, B. Wang, Prognostic value of cardiovascular biomarkers in COVID-19: a review, Viruses 12 (5) (May 11 2020), doi:10.3390/v12050527.

[49] L. Shi, Y. Wang, Y. Wang, G. Duan, H. Yang, Meta-analysis of relation of creatine kinase-MB to risk of mortality in coronavirus disease 2019 patients, Am. J. Cardiol. 130 (Sep 1 2020) 163–165, doi:10.1016/j.amjcard.2020.06.004.

[50] T.M. Welsh, G.D. Kukes, L.M. Sandweiss, Differences of creatine kinase MB and cardiac troponin I concentrations in normal and diseased human myocardium, Ann. Clin. Lab. Sci. 32 (1) (2002) 44–49 Winter[Online]. Available https://pubmed.ncbi.nlm.nih.gov/11848617/ .