Exp Biomed Res 2025; 8(2): 58-71 https://doi.org/ 10.30714/j-ebr.2025.237

Experimental Biomedical Research *Original article*

Predicting mortality risk and determining critical factors in intensive care patients: A preliminary study on covid-19 patients

Fatma Latifoğlu[1*], Aynur Karayol Akin[2], Firat Orhan Bulucu[3], Gamze Kalin Ünüvar[4], Ramis İleri[1], Burcu Baran Ketencioğlu[5], Şahin Temel[6]

[1]*Department of Biomedical Engineering, Faculty of Engineering, Erciyes University, Kayseri, Türkiye*
[2]*Department of Anesthesiology and Reanimation, Faculty of Medicine, Erciyes University, Kayseri, Türkiye*
[3]*Department of Biomedical Engineering, Inonu University, Malatya, Türkiye*
[4]*Department of Infectious Diseases, Faculty of Medicine, Erciyes University, Kayseri, Türkiye*
[5]*Department of Pulmonary Disease, Faculty of Medicine, Erciyes University, Kayseri, Türkiye*
[6]*Division of Intensive Care, Department of Internal Medicine, Faculty of Medicine, Erciyes University, Kayseri, Türkiye*

ABSTRACT

Aim: To predict the mortality risk of COVID-19 patients in the intensive care unit (ICU) using clinical parameters and machine learning approaches.

Methods: Data from 307 ICU patients at Erciyes University Hospital (2021–2022) were analyzed. Particle swarm optimization (PSO) and least absolute shrinkage and selection operator (LASSO) methods were utilized for feature selection. Four machine learning algorithms support vector machine (SVM), K-nearest neighbors (KNN), ensemble methods, and artificial neural network (ANN) were applied to the selected parameters.

Results: The top 10 predictive parameters, common to both LASSO and PSO, included sodium, nucleated red blood cell (NRBC) count, magnesium, mean corpuscular hemoglobin (MCH), and lymphocyte count. The best prediction performance was achieved using PSO feature selection and ANN (AUC: 86.77%, sensitivity: 85.12%, specificity: 77.44%, F1-score: 81.10%).

Conclusions: This study identifies critical parameters for predicting ICU COVID-19 patient mortality risk, employing two feature selection methods and comparing their performance with four machine learning algorithms. These results offer valuable insights for specialized physicians regarding disease progression and mortality risk prediction, but further research is needed.

Keywords: Intensive care unit, COVID-19, PSO, LASSO, machine learning.

*Prof. Dr. Fatma Latifoğlu**
Department of Biomedical Engineering,
Faculty of Engineering, Erciyes University, Kayseri, Türkiye
E-mail: flatifoglu@erciyes.edu.tr
Received: 2024-10-24 / Revisions: 2025-01-01
Accepted: 2025-01-12 / Published: 2025-03-16

1. Introduction

As of March 2020, COVID-19, defined as a worldwide pandemic by the World Health Organization, has affected the whole world and caused millions of deaths [1]. Symptoms of this pandemic may vary in patients and control of the pandemic may be difficult [2]. The most common symptoms include flu-like cough and fever, respiratory distress, and loss of smell and taste [1-3]. These symptoms generally affect elderly people and those with chronic diseases more. It has been observed that these groups are more likely to be admitted to intensive care unit

(ICU) than others [4]. Various treatment and vaccine development studies for COVID-19 are ongoing. However, a definitive treatment method has not been determined due to new variants [5]. Early decision-making is critical for the timely intervention of the patient. For this reason, it has become important to predict the lethal risk of the epidemic in advance. At this point, the importance of machine learning (ML) algorithms, frequently preferred in health studies, emerges. It has been observed that ML algorithms provide sufficient accuracy in predicting COVID-19 mortality [6].

ML models are statistical and mathematical algorithms capable of analyzing facts and making decisions in complex problem domains. These models can offer critical insights for rapid, evidence-based decision-making by utilizing laboratory tests and clinical data [3, 7]. ML algorithms have been used in many studies, such as detecting COVID-19 outbreak outbreaks, rapid diagnosis, diagnosis and classification of disease diseases from medical images, and admission of patients to ICU [8]. Recently, ML studies have gained significant importance in developing predictive models related to the COVID-19 outbreak, utilizing clinical data such as demographic characteristics and blood samples [1, 7, 9]. While most of these studies focus on detecting the presence of COVID-19, research on mortality and recovery prediction has remained insufficient until recently [10]. The effects of the COVID-19 pandemic are inherently challenging to predict [11]. Consequently, many researchers have directed their efforts toward predicting patient survival following COVID-19 infection [12]. One of the primary concerns, particularly for ICU patients and physicians, is determining the likelihood of survival [8]. In this context, an early warning system that classifies COVID-19 patients based on their mortality risk upon ICU admission is crucial for assisting physicians in managing the disease. Machine learning methods play a pivotal role in studies focused on intensive care [9]. Data in electronic health records related to clinical tests routinely obtained from patients can be evaluated together with ML models to help physicians identify individuals at high risk of death. In addition, while predicting the mortality rate of COVID-19 patients using ML models, identifying important clinical data that are effective in these predictions effectively reduce mortality rates. For this reason, this study aims to predict the risk of death based on ML algorithms using clinical data according to the 7-day periodic follow-up of COVID-19 patients hospitalized in the intensive care unit. In addition, two different feature selection methods were applied to determine the most important parameters affecting the mortality risk. For this purpose, the following questions were sought in the study:

1. The question of what are the most determinant features in the risk of death of COVID-19 patients hospitalized in the ICU was investigated, and the most determinant clinical data were determined by applying two different feature selection methods.

2. The question of which ML algorithm performs the most successfully in predicting the risk of death as a result of 7-day periodic controls of COVID-19 patients admitted to ICU was investigated. Within the scope of this research, the results of four different ML algorithms were obtained and their performances were compared in terms of classification criteria.

Studies utilizing ML methods to predict the risk of mortality based on the clinical data of patients diagnosed with COVID-19 and admitted to the ICU have been examined. These recent studies are detailed in Section 1.1.

1.1. Related Works: Since the outbreak of the COVID-19 pandemic, many studies have been conducted on this topic. The use of ML for these studies is an essential research area [9, 13]. Many studies have used various ML algorithms to evaluate clinical values, hospitalization, and treatment times. It is seen that most of these studies are on COVID-19 outbreak detection and mortality risk [2, 8, 10-15]. In studies on intensive care, it has been observed that there is a need for intensive care of patients, admission to intensive care, prediction of discharge time from ICU, and death/recovery rate [4, 9, 16-23]. Research indicates that studies on the survival of individuals infected with COVID-19 remain insufficient [10, 18]. In particular, there is a recognized need for research focusing on the progression of the disease in COVID-19 patients hospitalized in the ICU [9, 21]. A review of the literature reveals a notable gap in studies addressing the mortality and recovery prediction of COVID-19 patients, especially those admitted to the ICU. To address this gap, this study focuses on mortality prediction for COVID-19 patients admitted to the ICU, as opposed to those hospitalized in general wards. In line with this objective, a summary of studies employing ML models to predict the survival of COVID-19 patients based on clinical data in the ICU is presented below.

- Elhazmi et al. estimated the mortality rate using several clinical characteristics obtained from 1468 COVID-19 patients in intensive care in critical condition with 28-day periodic measurements. They used the decision tree (DT) algorithm to predict this rate and obtained an accuracy rate of 73.1% [9].
- Nazir and Ampadu proposed a new model in their study on ICU mortality of COVID-19 patients. Their suggested model, estimated mortality at an AUC rate of 96.3% using features such as demographic information, symptoms, and laboratory tests [18].
- In their study, Aznar-Gimeno et al. identified important parameters by reducing 165 different variables to 20 in the pre-processing stage. Then, they evaluated these parameters in different ML algorithms and obtained 82.1% AUC, 71% sensitivity, and 78% specificity rates of the XGBoost model [20].
- Jamshidi et al. conducted a study on the early prediction of mortality using ML models based on laboratory results. In their study, they examined 797 patients admitted to ICU and diagnosed with COVID-19. As a result of the examination, mortality was predicted at 70% sensitivity and 75% specificity rates with the RF algorithm [21].
- Ryan et al. conducted an analysis study on mortality according to various periods including 114 patients who were admitted to the ICU and tested positive for COVID-19. Their study, obtained the most successful AUC result of 91% in the XGBoost algorithm [23].
- Vaid et al. analyzed the data of 4029 COVID-19 patients from five hospitals in their study. As a result of their analyses, they predicted the mortality rate within 7 days with the Multi-Layer Perceptron (MLP) algorithm with an AUC of 82.2% [24].
- Cheng et al. aimed to predict mortality in their study based on clinical data of COVID-19 patients hospitalized in the ICU. For this purpose, they obtained 69.7% AUC, 65.7% accuracy, 67.4% sensitivity, 64.4% specificity, and 63.8% F1-Score using the deep learning model they proposed [25].

2. Materials and methods

2.1. Proposed Study: This study, aimed to determine the most successful model by testing ML models on death/recovery prediction based on clinical values obtained periodically (7 days) from COVID-19 patients hospitalized in ICU and to propose a support system to help physicians. For this study, original data collected from COVID-19 patients admitted to the intensive care unit were utilized. Clinical test results, including parameters such as sodium, calcium, white blood cell count, mean corpuscular hemoglobin and mean corpuscular volume, were preprocessed and analyzed using four different ML algorithms for classification. The performance of these algorithms was evaluated based on classification metrics. Additionally, two different feature selection methods were applied to identify and analyze significant clinical parameters. Detailed information about the dataset, preprocessing steps, and classification processes is provided below. A general flowchart summarizing the proposed methodology is presented in Figure 1.

conducted by the Declaration of Helsinki and the Erciyes University Faculty of Medicine Ethics Committee accepted the study protocol. Some preprocessing stages were applied since there were blank values, missing information, etc. in the data set. The preprocessing and feature selection stages applied to the dataset are described in sections 2.3-2.4. A general flow diagram summarizing these stages is shown in Figure 2.

2.3. Pre-processing: Preprocessing is the stage where various techniques such as merging data and discarding missing values are performed. The application of preprocessing to data is important for developing ML models [18, 26]. In this study, the data of COVID-19 patients admitted to the intensive care unit between 2021-2022 were used. Data from those under 18 years of age were eliminated by excluding them from

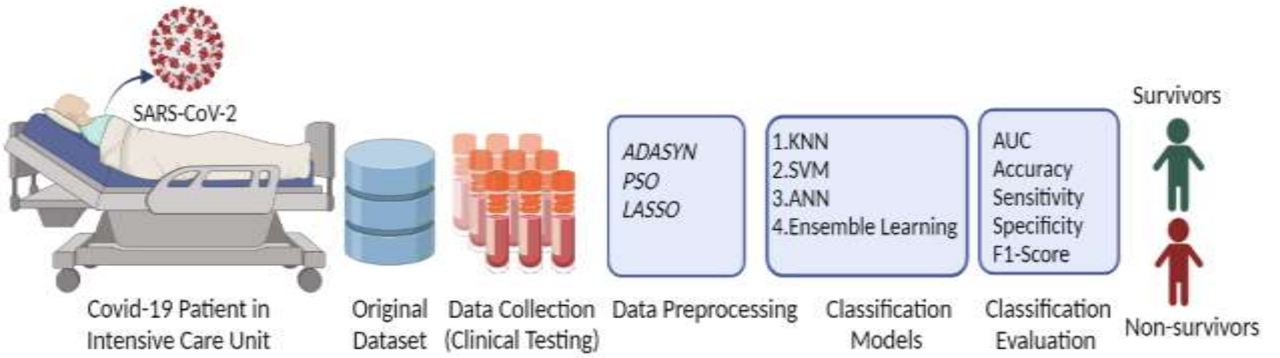

Figure 1. Flow chart outlining the proposed work.

2.2. Data Collection: This study, used data from COVID-19 patients hospitalized in the intensive care unit of Erciyes University Hospital between 2021-2022. Before the data were used, the consent of the subjects included in the study was obtained and information was provided. Data from those under 18 years of age were not included in the study. All of the data used were taken from people aged 18 years and over. In the study, clinical parameters were measured periodically (7 days), and total of 307 COVID-19 patients were evaluated. Our study was

the study. As a result of discarding empty values from these data, data belonging to 307 people were included in the study. Among 307 COVID-19 patients, 217 people were discharged from the intensive care unit, while 90 people died. As a result of discarding the empty values, the number of clinical parameter measurements taken from these patients was 244. When the standard measured parameters were selected, 29 parameters were included in the study. Before feature selection and data balancing, 307 data and 29 parameters were selected for this study.

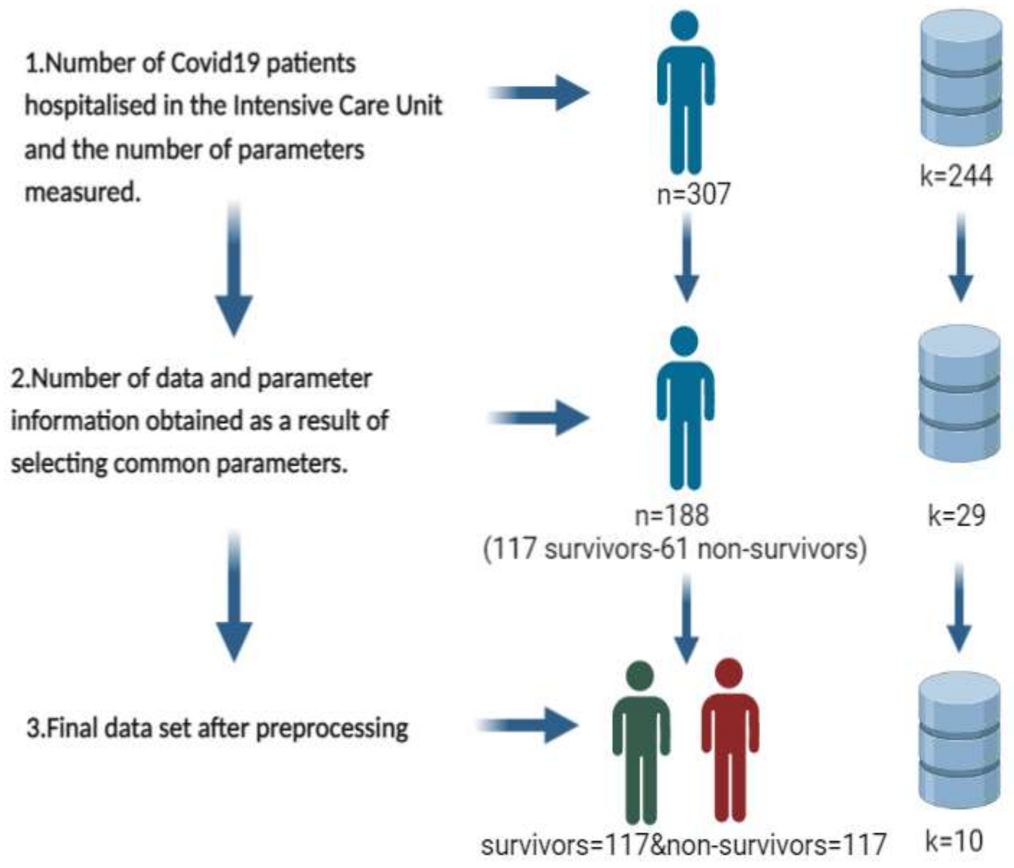

Figure 2. Summary diagram of the pre-processing stages.

Unbalanced data distribution before classification is one of the most important obstacles to ML algorithms [6]. This situation occurs when the groups in the classification stage are not equally distributed. In this study, two classes of COVID-19 patients who survived (n=217) and died (n=90) as a result of periodic controls from COVID-19 patients hospitalized in ICU were predicted using ML algorithms. Before this process, the number of data belonging to these two classes showing unbalanced distribution was equalized by applying the Adaptive Synthetic (ADASYN) technique. With this technique, the number of data belonging to the minority class was equalized to that of the majority class. In other words, a total of 434 data were used before the classification process.

ADASYN: It is a method introduced to the literature by He et al., [27] and used for balancing unbalanced distributed data. With this method, more synthetic data is generated for minority class samples that are more difficult to learn compared to minority samples that are easier to learn. In the ADASYN method, a weighted distribution is used for different minority classes according to the level of difficulty in learning [18, 27].

2.4. Feature Selection Methods: The feature selection phase is a technique employed to reduce the number of available features in a dataset, selecting the most relevant ones before constructing ML models [11, 18]. This process simplifies the dataset by removing irrelevant features [6]. Feature selection is an essential step commonly used in COVID-19 studies and has demonstrated its value in problem-solving [28]. The primary goals of feature selection are to identify which data is most suitable for a given classification model, reduce training time, and enhance model performance [29, 30].

In the feature selection phase of this study, PSO and LASSO techniques were applied. Through these methods, the number of clinical parameters, initially 29, was reduced to 10, thereby identifying the most important parameters before the classification process. PSO was chosen for this research due to its proven effectiveness and, to the best of our knowledge, its novelty in studies focused on survival prediction for COVID-19 patients hospitalized in the ICU. To compare the PSO technique, LASSO was also used as an alternative feature selection method in this study.

2.4.1. Particle Swarm Optimization Technique: PSO technique is one of the metaheuristic techniques introduced to the literature by Eberhart and Kennedy [31]. Since it is simple to implement, it is a popular and successfully applied technique [32]. PSO is a swarm-based intelligent stochastic search technique [33]. In PSO, several swarms are assigned to a search space with a position and velocity. Each swarm improves its position within the search space for the best local best position and global position obtained by the whole swarm using the fitness function [32]. Each particle in the swarm is identified by its position and velocity in the multidimensional search space. All particles move to discover their best position in the search space by keeping a record of their previous best positions and trying to continue to improve in finding the best global position. The global best position is governed by the fitness function. PSO tries to optimize the fitness function until the stopping criterion is satisfied [34]. For detailed information, please refer to these sources. In the literature, there are successful studies on medical diseases and COVID-19 using PSO [30-34].

2.4.2. Least Absolute Shrinkage and Selection Operator*:* The LASSO technique was formulated and proposed by Robert Tibshirani in 1996. This technique fulfils two important tasks as regularization and feature selection. The regularization process is performed depending on the coefficients of the regression variables. During the feature selection process, variables with a non-zero coefficient after the regularization process are selected to be part of the model. In this way, the prediction error minimized. In the LASSO technique, as the parameter λ increases, the coefficients are brought closer to zero and the dimensionality can be reduced [35].

2.5. Classification Models: ML algorithms have an important role in the classification of medical data [36]. Routinely obtained data from patients depending on clinical tests can be evaluated together with ML models to help physicians. For this purpose, SVM, K-Nearest KNN, Ensemble learning, and ANN classifiers from ML algorithms were used in the classification phase of this study on predicting the survival of COVID-19 patients in the ICU. In the classification phase, a 10-fold cross-validation technique was applied in the training and test phase of the data. In this case, 90% of the data was reserved for training and 10% for testing. Classification processes were performed using MATLAB 2022 software program. Information about the classifiers applied in this study is explained below.

SVM: The SVM algorithm is used in classification and regression processes. The purpose of this algorithm is to take over how to draw the boundary when categorizing the class of features. It finds the hyperplane that sets the distance between itself and the dataset at the nearest point [29].

KNN: The KNN algorithm, like the SVM algorithm, is used in classification and regression processes. It determines the closest relationship between variables for classification. In the KNN algorithm, the distance between the data is calculated according to the similarity, and the unknown data is labeled according to the values of the nearest neighbors [26].

Ensemble Learning: The purpose of this model is to increase accuracy by significantly reducing classification errors by combining predictions. Bagging and boosting models are generally preferred in studies. In this study, the bagging model is preferred and the classification process is performed. With the bagging model, weak models are trained by reducing the variance of a noisy dataset [36, 37].

ANN: ANN is an algorithm that mimics the performance of the human brain, usually consisting of three layers (input layer, calculation or hidden layer, and output layer) [38]. Each layer in this algorithm contains neurons, and has different tasks [39]. The critical parameter in ANN is the activation function. Thanks to the activation function, non-linear learning can be increased and complex computations can be organized [38, 39]. In this study, a feed-forward ANN algorithm known as a multilayer perceptron is preferred.

2.6. Classification Models Evaluation Metrics: Different performance measures are used to evaluate the performance of classification models. In this study, AUC, sensitivity, specificity, and F1-Score criteria were used to evaluate the performance of the preferred classifiers. Sensitivity, also known as the Recall rate is widely preferred in ML studies. This criterion value expresses the rate of correctly predicted data for the survivor class to the total number of survivors. The specificity rate is calculated as the ratio of the correctly predicted data for the deceased class to the total number of deceased data. F1-Score is used to evaluate the performance of a model by considering both precision and sensitivity. The F1-Score rate is calculated by taking the harmonic mean of these two measures. AUC helps to visualize the performance of machine learning models in predicting classes. Its performance is evaluated depending on the area under the curve. As this area increases, the performance between classes increases [36]. One of the most important criteria for studies with unbalanced data sets is the AUC rate [40]. The formulas for sensitivity, specificity, and F1-Score are given in Equation 1-3.

$$Sensitivity = \frac{tp}{tp+fn} \times 100 \quad (1)$$

$$Specificity = \frac{tn}{tn+fp} \times 100 \quad (2)$$

$$F1 - Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Recall} + \text{Precision}} \times 100 \quad (3)$$

3.Results

This study, aims to predict whether COVID-19 patients hospitalized in the intensive care unit will survive or not according to the clinical measurement parameters taken from COVID-19 patients with the help of ML algorithms. For this purpose, data from 307 COVID-19 patients and 29 clinical measurements were used. PSO and LASSO techniques were used as feature selection methods to determine the important ones from these parameters, and 10 important parameters were determined before classification. Information about the important parameters determined as a result of PSO and LASSO techniques is given in Table 1. In the classification phase of this study, the parameters in Table 1 were applied as input to four different ML algorithms and the prediction of deceased and surviving COVID-19 patients was performed. In the classification phase, KNN, SVM, Ensemble Learning, and ANN models were used and the results were compared in terms of AUC, sensitivity, specificity, and F1-Score evaluation metrics.

In this study, the classification process was performed for three different cases. The classification was performed by applying both

Table 1. Important parameters determined as a result of the application of feature selection methods.

PSO	**LASSO**
Mean corpuscular volume (MCV)	Eosinophilia
Hematocrit	Calcium
Neutrophils	Procalcitonin
White blood cell count (WBC)	Red cell distribution width
Mean platelet volume	Mean corpuscular hemoglobin concentration
MCH	**MCH**
NRBC	**NRBC**
Lymphocyte count	**Lymphocyte count**
Magnesium	**Magnesium**
Sodium	**Sodium**

Table 2. Classification results obtained by applying the PSO feature selection method.

	AUC (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
KNN	80.19	73.55	79.48	76.07
SVM	83.80	80.99	70.94	77.47
Ensemble Learning	86.76	80.17	77.78	79.51
ANN	86.77	85.12	77.44	81.10

Table 3. Classification results obtained by applying the LASSO feature selection method.

	AUC (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
KNN	82.73	80.99	76.07	79.35
SVM	85.38	80.99	75.21	79.03
Ensemble Learning	86.57	85.12	76.92	82.07
ANN	84.55	78.51	76.92	78.19

Table 4. Classification results obtained according to the common parameters of LASSO and PSO feature selection methods.

	AUC (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
KNN	83.95	80.17	76.07	78.87
SVM	80.93	74.38	73.50	74.38
Ensemble Learning	82.81	84.30	72.65	80.00
ANN	81.02	90.91	70.09	82.71

PSO and LASSO feature selection. In addition, the classification process was performed using the common ones of important parameters selected by the application of both LASSO and PSO methods. When Table 1 was analyzed, five important parameters for both feature selection methods were found. The classification results obtained as a result of PSO feature selection are given in Table 2. The classification results obtained as a result of LASSO feature selection are given in Table 3. The classification results of the parameters common to the feature selection methods are given in Table 4.

When Table 1 was analyzed, the number of parameters was reduced to 10 by determining the important ones from the initial 29 parameters before the classification process by applying PSO and LASSO feature selection methods.

Sodium, NRBC, Magnesium, MCH, and Lymphocyte count were determined as important parameters for both feature selection methods. Sodium levels are recommended for fluid therapy and monitoring for prevention of neurological complications. NRBC levels are associated with increased mortality risk and are a biomarker that should be closely monitored. Magnesium is important for preventing cardiac complications. MCH is an important biomarker for the assessment of oxygen carrying capacity in COVID-19 patients. Lymphocyte levels are important for the assessment of patients' immunity.

According to Table 2-4, it is seen that the ML algorithms preferred in this study to determine the mortality risk of COVID-19 patients hospitalized in ICU give close results, especially when evaluated on the AUC success rate. The ANN algorithm performed more successfully than the PSO feature selection method. When the parameters determined according to the LASSO feature selection are evaluated, it is seen that the Ensemble Learning algorithm gives successful results. The KNN algorithm was observed to be more successful in the jointly selected parameters. In this study, both the performances

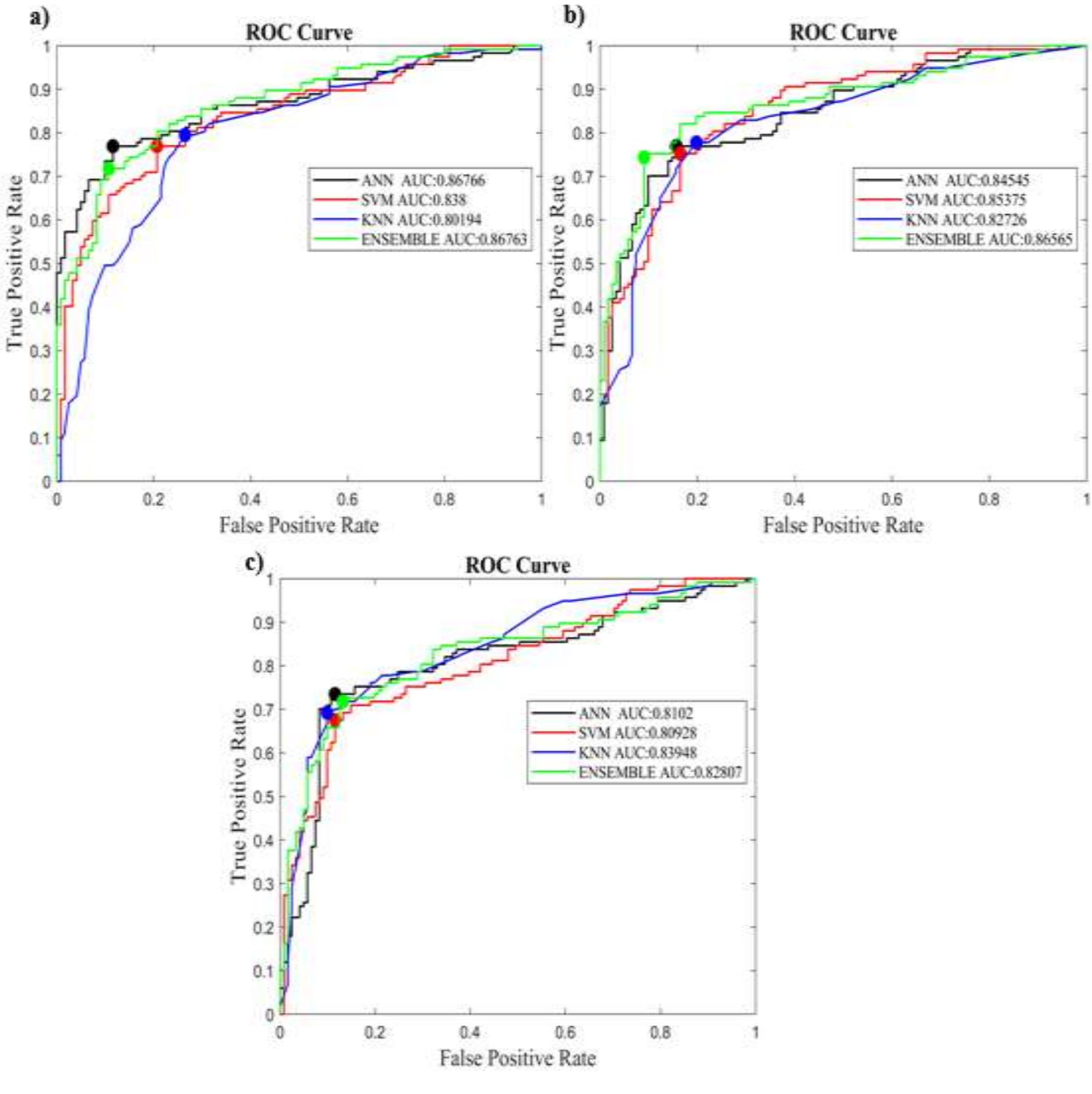

Figure 3. Comparison of area under the receiver operating characteristic curves for classifier models. (a. PSO feature selection method, b. LASSO feature selection method, c. PSO & LASSO feature selection method).

of ML algorithms and the performances of feature selection methods with these algorithms were compared. The ROC curve is used to evaluate and compare the overall performance in diagnostic procedures. The receiver operating characteristic (ROC) curves obtained to see the overall performance of the feature selection methods and ML algorithms used in this study are given in Figure 3.

4. Discussion

This study aimed to predict the survival of COVID-19 patients admitted to the ICU based on certain clinical measurements with ML algorithms. One of the biggest concerns for patients hospitalized in the ICU and physicians responsible for the follow-up of these patients is whether the patients will survive [8]. For this, many clinical measurements are taken from patients periodically. However, making predictions based on these measurements can be both laborious and time-consuming. At this point, predictions based on ML algorithms become important.

This study, examined parameters based on clinical measurements of 307 COVID-19 patients hospitalized in intensive care between 2021-2022. In this examination, LASSO and PSO feature selection methods, which have recently been preferred in COVID-19 studies, were used [31-33]. However, as far as we know, no study has been conducted to estimate the risk of death of patients hospitalized in the ICU using these methods. For this reason, we think that this study is important. The 10 important parameters determined by PSO and LASSO methods were evaluated with four different ML algorithms to predict the survival of COVID-19 patients. As a result of the evaluation, the parameters that are effective in estimating mortality risk and have the highest values were determined and the estimation was carried out to determine the mortality risk.

Among the parameters given in Table 1, "MCV, MCH, WBC, Lymphocyte count" was found to be the parameters with the highest value on mortality risk prediction in a study [21] by using a different feature selection method. In

Table 5. Comparison of the results of the proposed study with present studies.

Reference	Year	Feature Selection	Best Classifier	Evaluation Results (%)
[9]	2020	Statistical Analysis	DT	Accuracy:73.1
[20]	2020-2021	Statistical Analysis	XGBoost	AUC:82.1
[21]	2020	Statistical Analysis	RF	Specificity:75
[23]	2020	-	XGBoost	AUC:91
[24]	-	LASSO	MLP	AUC:82.2
[25]	2020	Statistical Analysis	deep neural network	AUC:69.7
[41]	2020	Statistical Analysis	deep neural network	AUC:84.4
Proposed study	**2021-2022**	**PSO** **LASSO** **PSO & LASSO**	**ANN** **Ensemble Learning** **KNN**	**AUC:86.77** **AUC:86.57** **AUC:83.95**

DT: Decision tree, RF: Random Forest, MLP: Multi-Layer Perceptron, ANN: Artificial Neural Network, KNN: K-Nearest Neighbors, LASSO: Least Absolute Shrinkage and Selection, PSO: Particle Swarm Optimization.

another study [20], MCV and Lymphocyte count were determined as important parameters. The results obtained in this study on the determination of important parameters support the literature. When we examined the studies conducted in the literature, it was seen that very few studies have been conducted on predicting the risk of death of COVID-19 patients and especially examining COVID-19 patients hospitalized in ICU [9, 10, 18, 21]. It has been observed that these studies are not sufficient and additional studies are needed [10, 21]. The results obtained in this study and the studies conducted to estimate the mortality risk of COVID-19 patients hospitalized in the ICU are compared in Table 5.

As seen in Table 5, this study contains diversity in terms of both the comparison of feature selection methods and the comparison of ML algorithms to similar studies. While comparing with the studies in the literature, we tried to look at the AUC rate in particular. Because, as in this study, most of the studies given in Table 5 are analyzed over unbalanced data distribution. AUC rate is widely used to determine the best ML algorithm for unbalanced data [40]. As can be seen from the results, the AUC rate obtained in this study was successful compared to many studies [20, 24, 25, 41]. Considering the time lapse between data collection and various factors that may have influenced the results [23], it is possible to evaluate the potential impact of these factors. Treatment recommendations made to improve COVID-19 patients over time may be one of the factors. It is difficult to make predictions on the results of patients hospitalized in the ICU due to the development of treatments throughout the epidemic. Different results can be obtained in studies conducted at different times to evaluate similar basic clinical measurement parameters [25].

This study has some limitations.

- Since the data set used in the study can only be obtained from one institution, there may not be enough data. As a result, when we consider that the epidemic has spread widely, there may be concerns about generalizing the results. A larger data population is needed to generalize the results.
- Since the data set used in the study covers the period between 2021-2022, patients may have started drug or vaccine treatment. Since there is no information about the treatment in the data received, no evaluation can be made on this issue.

Along with the limitations of this study, there are also positive aspects that we think will contribute to the literature.

- We think that this study will contribute to the literature because there are not enough studies in the literature on the mortality risk of COVID-19 patients hospitalized in the ICU.
- In this study, the most important clinical parameters measured from COVID-19 patients were determined by two different feature selection methods, and the risk of death was estimated using these features. Parameters that can provide important information about the recovery process of COVID-19 patients hospitalized in ICU and support specialist physicians have been identified. In this way, we think that knowing the effective parameters that determine the prediction of the risk of death can help clinicians in periodic controls of patients.
- Four different ML algorithms and two different feature selection methods were used to predict the risk of death and their performances were compared and

diversity was created in the study. In this way, we think that our research can make important contributions in terms of both comparing the performances of feature selection methods and observing the results of ML algorithms together with these methods.

4.1. Conclusion: In conclusion, in this study, we aimed to follow the recovery process and predict the risk of death of COVID-19 patients in the ICU using ML algorithms. For this purpose, important clinical parameters that affect the survival of COVID-19 patients in the intensive care unit were determined. A novel study was carried out using these parameters to automatically predict survival with ML algorithms. As a result of this study, the most successful classification result was obtained with PSO feature selection and ANN classifier with a rate of 86.77% in terms of the AUC rate. We think that the results obtained in this study can help specialist physicians to provide information on the survival of COVID-19 patients lying in the ICU and to make predictions on the risk of death of patients. However, more studies are needed to interpret these results in a generalizing way.

Funding: *The authors received no financial support for the research, authorship, and/or publication of this article.*

Conflict of Interest: *The authors declared no conflict of interest.*

Ethical Statement: *This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Erciyes University Hospital (protocol code 2021/350).*

References

[1]Xiao Z. COVID-19 mortality rate prediction based on machine learning methods. In: Proceedings of the 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI). 2021:169-177.

[2]Jamshidi E, Asgary A, Tavakoli N, et al. Symptom prediction and mortality risk calculation for COVID-19 using machine learning. Front Artif Intell. 2021;4:673527.

[3]Tabata, S, Imai, K, Kawano, S, et al. Clinical characteristics of COVID-19 in 104 people with SARS-CoV-2 infection on the Diamond Princess cruise ship: a retrospective analysis. Lancet Infect Dis. 2020; 20(9): 1043-1050.

[4]Abate SM, Ahmed Ali S, Mantfardo B, et al. Rate of Intensive Care Unit admission and outcomes among patients with coronavirus: a systematic review and Meta-analysis. PloS one. 2020; 15(7): 1-19.

[5]Wise, J. Covid-19: New coronavirus variant is identified in UK. BMJ 2020;371:m4857.

[6]Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z, et al. Comparing machine learning algorithms for predicting COVID-19 mortality. BMC medical informatics and decision making. 2022; 22(1): 1-12.

[7]de Hond AAH, Kant IMJ, Fornasa M, et al. Predicting Readmission or Death After Discharge From the ICU: External Validation and Retraining of a Machine Learning Model. Crit Care Med. 2023;51(2):291-300.

[8] Gao, Y, Cai GY, Fang W, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. Nat Commun. 2020; 11(1): 5033.

[9] Elhazmi A, Al-Omari A, Sallam H., et al. Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. J Infect Public Health. 2022; 15(7); 826-834.

[10] Atef O, Nassir AB, Talib MA, et al. Death/Recovery Prediction for Covid-19 Patients using Machine Learning. IJECS. 2020;14:189-193.

[11] Gupta VK, Gupta A, Sardana A, et al. Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model. Big Data Mining and Analytics. 2021; 4(2): 116-123.

[12] Kumar N, Qi SA, Kuan L, et al. Learning accurate personalized survival models for predicting hospital discharge and mortality of COVID-19 patients. Sci Rep. 2022; 12(1): 1-11.

[13] Alballa N, Al-Turaiki I. Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. Inform Med Unlocked. 2021;24:100564.

[14] Pourhomayoun M, Shakibi M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. Smart Health (Amst). 2021;20:100178.

[15] Aktar S, Ahamad MM, Rashed-Al-Mahfuz M, et al. Machine learning approach to predicting COVID-19 disease severity based on clinical blood test data: statistical analysis and model development. JMIR Med Inform. 2021; 9(4): e25884.

[16] Subudhi S, Verma A, Patel AB, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. NPJ Digit Med. 2021; 4(1): 1-7.

[17] No JT, Timur LB, Sekaki KP. Optimasi Naive Bayes dengan PSO untuk Prediksi Kebutuhan ICU Pasien Covid-19. J Riset Inform. 2018;1(1):43-50.

[18] Nazir A, Ampadu HK. Interpretable deep learning for the prediction of ICU admission likelihood and mortality of COVID-19 patients. Peer J Comput Sci. 2022;8:e889.

[19] Varzaneh ZA, Orooji A, Erfannia L, et al. A new COVID-19 intubation prediction strategy using an intelligent feature selection and K-NN method. Inform Med Unlocked. 2022;28:100825.

[20] Aznar-Gimeno R, Esteban LM, Labata-Lezaun G, et al. A Clinical Decision Web to Predict ICU Admission or Death for Patients Hospitalised with COVID-19 Using Machine Learning Algorithms. Int J Environ Res Public Health. 2021;18(16):8677.

[21] Jamshidi E, Asgary A, Tavakoli N, et al. Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. Front Digit Health. 2022;3:681608.

[22] Banoei MM, Dinparastisaleh R, Zadeh AV, et al. Machine-learning-based COVID-19 mortality prediction model and identification of patients at low and high risk of dying. Crit Care. 2021;25(1):328.

[23] Ryan L, Lam C, Mataraso S, et al. Mortality prediction model for the triage of COVID-19, pneumonia, and mechanically ventilated ICU patients: A retrospective study. Ann Med Surg (Lond). 2020;59:207-216.

[24] Vaid A, Jaladanki SK, Xu J, et al. Federated Learning of Electronic Health Records Improves Mortality Prediction in Patients Hospitalized with COVID-19. medRxiv [Preprint]. 2020:2020.08.11.20172809.

[25] Cheng J, Sollee J, Hsieh C, et al. COVID-19 mortality prediction in the intensive care unit

with deep learning based on longitudinal chest X-rays and clinical data. Eur Radiol. 2022; 32(7): 4446-4456.

[26] Khan, IU, Aslam N, Aljabri M, et al. Computational intelligence-based model for mortality rate prediction in COVID-19 patients. Int J Environ Res Public Health. 2021; 18(12): 6429.

[27] He H, Bai Y, Garcia EA, et al. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In International Joint Conference on Neural Networks (IJCNN 2008). 2008:1322-1328.

[28] Limam H, Hasni O, Alaya IB. A novel hybrid approach for feature selection enhancement: COVID-19 case study. Comput Methods Biomech Biomed Engin. 2023;26(10):1183-1197.

[29] de Freitas Barbosa VA, Gomes JC, de Santana MA, et al. Heg.IA: an intelligent system to support diagnosis of Covid-19 based on blood tests. Res. Biomed. Eng. 2022;38(1):99–116.

[30] Savareh, BA, Alibabaei A, Ahmady S, et al. Predicting Patient Length of Stay in the Neurosurgical Intensive Care Unit of an Academic Medical Center. J Cell Mol Anesth. 2021; 6(2): 132-140.

[31] Eberhart R, Kennedy JA. A New Optimizer Using Particle Swarm Theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, 4-6 October 1995, 39-43. http://dx.doi.org/10.1109/MHS.1995.494215

[32] Makade RG, Chakrabarti S, Jamil B. Real-time estimation and prediction of the mortality caused due to COVID-19 using particle swarm optimization and finding the most influential parameter. Infect Dis Model. 2020;5:772-782.

[33] Pervaiz S, Ul-Qayyum Z, Bangyal WH, et al. A Systematic Literature Review on Particle Swarm Optimization Techniques for Medical Diseases Detection. Comput Math Methods Med. 2021;2021:5990999.

[34] Khan MA, Alhaisoni M, Nazir M, et al. A Healthcare System for COVID19 Classification Using Multi-Type Classical Features Selection. Computers, Materials and Continua. 2022;74(1): 1393-1412.

[35] Fonti V, Belitser E. Feature selection using LASSO. Business, Computer Science, Mathematics. 2017;30: 1-25.

[36] Acer İ, Bulucu FO, Içer S, et al. Early diagnosis of pancreatic cancer by machine learning methods using urine biomarker combinations. Turkish Journal of Electrical Engineering and Computer Sciences. 2023; 31(1): 112-125.

[37] Zhou ZH, Zhou ZH. Ensemble learning. Springer Nature Singapore Pte Ltd. 2021. pp. 181-210. https://doi.org/10.1007/978-981-15-1967-3

[38] Li H, Zhang Z, Liu Z. Application of artificial neural networks for catalysis: a review. Catalysts. 2017; 7(10): 306.

[39] Shanbehzadeh M, Nopour R, Kazemi-Arpanahi H. Design of an artificial neural network to predict mortality among COVID-19 patients. Inform Med Unlocked. 2022;31:100983.

[40] Massoud, Y.M., Abdelzaher, M., Kuhlmann, L. et al. General and patient-specific seizure classification using deep neural networks. Analog Integr Circ Sig Process. 2023;116(3): 205–220.

[41] Li X, Ge P, Zhu J, et al. Q. Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables. PeerJ. 2020; 8: e10337.