P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.27 No.2 April - June 2022
Page: [98-114]
Original research article

Predicting the Mortality in the Patients Hospitalized in Intensive Care Units (ICU) Based on Machine Learning Techniques

Khadijeh Moulaei[1], Kambiz Bahaadinbeigy[2,*], Fahimeh Ghasemian[3], Zahra Mohamadi Taghiabad[4]

[1]Student Research Committee, Kerman University of Medical Sciences, Kerman 7616913555, Iran
[2]Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman 7616913555, Iran
[3]Department of Computer Engineering, Faculty of Engineering, Shahid Bahonar University of Kerman, Kerman 7616913555, Iran
[4]School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran 1996713883, Iran

Received 31 December 2020; Received in revised form 6 May 2021
Accepted 1 June 2021; Available online 29 June 2022

ABSTRACT

The patients hospitalized in an Intensive Care Unit (ICU) are always at risk of death. Use of machine learning for predicting the mortality of ICU patients is essential to assess the severity of the disease and to judge the value of new treatments, interventions, and health care policies. Therefore, this study aims to predict the mortality in patients hospitalized in the ICU based on machine learning techniques. The present study was conducted in four steps: data understanding, preprocessing, modeling, and evaluation. First, the mortality factors in the patients hospitalized in the ICU were identified and were extracted from 800 records of patient data. Next, different prediction algorithms including decision tree (J48), MLP, KNN, random forest, and SVM were developed based on 19 confirmed factors. The random forest algorithm worked appropriately based on receiver operating characteristic (ROC), accuracy, precision, sensitivity, and specificity. The SVM algorithm was the weakest algorithm according to ROC, sensitivity, accuracy and precision. Also, ventilation was the most effective factor for predicting the patients' death. Our study shows that machine learning algorithms can appropriately predict the mortality rate in patients hospitalized in the ICU with high accuracy and precision, extracting deeper specificities. Since diagnosis by humans is a time-consuming process with a higher probability of mistakes, using these algorithms can decrease treatment decision mistakes made by the ICU physicians and increase the chance of successful treatment.

Keywords: Intensive Care Unit (ICU); Mortality; Machine learning; Prediction algorithms

***Corresponding author:** kambizb321@gmail.com

1. Introduction

The intensive care unit (ICU) is one of the most stressful places in the hospital [1] in which acute patients are managed by the most expert staff and developed equipment [2]. The patients hospitalized in the ICU suffer from acute disease or injury and are always at risk of death [3]. The longer a patient stays in the ICU, the more likely the mortalities increase, because the immune system is weakened, accordingly, and vital organs such as the lungs, heart, kidneys, liver, and brain are involved. Therefore, this challenge can increase mortality [4]. The mortality rate among patients in the ICU has been reported as 20 to 30%, constituting 20 to 50% of the total hospital mortality [5].

Predicting mortality as one of the most important clinical outcomes for patients hospitalized in the ICU can help determine the severity of the disease, identify new treatment methods, provide efficient hospital equipment, and adoption of health care policies [6]. To predict the mortality in the ICU, we are always faced with an important question of how to predict the risk of mortality and the factors affecting it for hospitalized patients and minimize the mortality rate in ICU based on available data [7]. To answer this question, it should be said that, in recent years, algorithms such as Decision Tree (J48), Random Forest (RF), K Nearest Neighbor (KNN), Multi-layer Perceptron (MLP), and Support Vector Machines (SVM) have been increasingly used to predict mortalities [8, 9]. Considering the irregular knowledge of the ICU patient, machine learning algorithms have the highest potential to predict mortality in the ICU [10].

These knowledge discovery algorithms can help predict deaths and events in a way that the human brain simply cannot process and analyze [11]. Also, these knowledge discovery tools can automatically analyze raw data and extract high-level information for health decision-makers to make better treatment decisions [12-13]. In their study, Nemati et al. showed that using real-time data in ICU, machine learning algorithms can accurately predict the sepsis start in a patient in the ICU 4 to 12 hours before clinical diagnosis [14].

In the last 30 years, various machine learning algorithms and scoring systems such as Acute Physiology and Chronic Health Evaluation (APACHE), Simplified Acute Physiology Score (SAPS), Mortality Probability Model (MPM), Naïve Bayes, and Sequential Organ Failure Assessment (SOFA) have been developed to predict mortality[15, 16]. These algorithms and scoring systems have limitations. The most comprehensive and accurate of these scoring systems, APACHE, is a proprietary tool that requires a license and depends heavily on choosing the correct acceptance recognition. MPM and SAPS also consider only a few variables, resulting in ease of use, but are too simple models that may ignore important physiological measurements [17].

Some studies have shown that the use of J48, RF, KNN, MLP, and SVM algorithms for predicting mortality is very effective compared to older algorithms [16, 18, 19]. Comparison of j48 and Naive Bayes (NB) algorithms showed that the efficiency and accuracy of j48 are better than that of Naive Bayes [20]. Awad et al. [21] proposed RF ensemble learning for mortality prediction of hospitalized patients and concluded that the introduced algorithm works better than other prediction algorithms. Also, Moridani et al. [17] developed an SVM algorithm to predict the risk of dying of cardiovascular patients in the ICU. They concluded that SVM performed better than the ANN algorithm.

Based on a large amount of the available raw data about patients hospitalized in the ICU the importance of machine learning algorithms to predict patient mortality accurately, and the powerful potential of these algorithms to change the future of medical science, this study aimed to predict mortality in patients hospitalized in the ICU based on machine learning techniques. In this study, after identifying and introducing the effective factors to predict the mortality in patients hospitalized in the ICU and classifying them, five algorithms, decision tree (J48), MLP, KNN, random forest, and SVM have been used. According to the main purpose of this study, it can be said that early prediction of the mortality for patients hospitalized in the ICU can save the patient's life, allocate resources to more acute patients [10], and help physicians to decide about the patient's stay time in hospital [22].

Besides identifying the patients who are at risk of death, it is possible to help make critical decisions such as stopping or continuing the treatment, studying the treatment risks, pursuing the resources and equipment needed for patients in the ICU, and decreasing the stay period in ICU [23].

2. Survey on the Previous Works

In recent decades, the data collected in ICUs has increased and has been used in data analysis and machine learning studies [24]. There is a wealth of data to improve the care of ICU patients which is still unused or in need of further analysis. However, automated detection tools based on different predicting models can analyze raw data and extract information to enable decision makers to make better decisions [12, 24, 25]. In mortality prediction models in the ICU, we are always looking for an answer to the critical question of how to identify the mortality risk and associated factors in order to minimize the mortality rate of patients [25]. Various studies have discussed and compared mortality prediction models for ICU patients [26-28].

Calvert et al. [29] used electronic health record collected clinical variables to predict intensive care unit mortality. Strong predictive accuracy was obtained in this study (80%). However, we question the practical application of this tool, which predicts a 12-hour sampling point. This study does not specify at what stage of the development of a critical care episode this tool should be used to have the best effect. In contrast, the algorithms presented in our study attempt to predict in-hospital mortality shortly after ICU admission.

In another study, Sadeghi et al. [30] tried to predict early hospital mortality. This study presents a new method for predicting mortality using 12 features extracted from patients' heart signals during the first hour of ICU admission using the MIMIC-III database. The extracted features are fed into eight classifiers: decision tree, linear discriminant, logistic regression, SVM, random forest, boosted trees, Gaussian SVM, and K-NN. The random forest algorithm estimates both accuracy and interpretability better than other algorithms. This study shows that heart rate (HR) signals can be used to predict mortality in patients in the ICU. In addition, Crawford et al. [31] concluded that a decision tree (DT) algorithm provides a clinically acceptable mining result in predicting lymph node spread in men with clinically localized prostate carcinoma. Yakovlev et al. [32] used the algorithms of k-nearest neighbors, Random Forest, and Naive Bayes to predict the in-hospital mortality and length of stay (LoS) in patients with acute coronary syndrome (ACS). Among the algorithms used, the accuracy of the Naive Bayes algorithm was the highest (90.0%).

Although in our study the algorithms introduced in the above studies have been used, the proposed framework of our study for mortality prediction is composed of physiological and clinical data. These physiological and clinical data include chart variables, laboratories, vital signs, and patient demographics, and are not necessarily related to a specific organ/diagnosis because the ICU is a very complex environment and patients with several diseases are typically hospitalized [28].

Also, some studies show that customized algorithms perform better and more efficiently than traditional scoring systems [33]. Lee and Maslove [34] performed a retrospective analysis using local clinical data captured from hospital electronic medical records (EMRs). The aim of their study was to customize a severity of illness score using local electronic medical record data. The results of this study showed that customized and trained algorithms based on ICU-specific data provide better mortality predictions than traditional SAPS scoring using the same predictor variables. In our study, we also used trained and customized ICU-specific data-based algorithms to predict mortality in patients hospitalized in the ICU, not traditional scoring such as SAPS.

At the end of this section it should be said that the ICU is a very complex environment and patients admitted to it may suffer from more than one disease or have different signs and symptoms, so it is very difficult to determine which custom model to use [28]. However, there is a need for mortality prediction algorithms in the ICU, which is the focus of this study.

3. Material and Methods

The present study was conducted based on the proposed model in Fig. 1. According to this model, the present study was conducted in four stages of data understanding, preprocessing, modeling, and evaluation. These steps are described below.

3.1 Step 1: Data understanding

As shown in Fig. 1, this step involves the final identification and confirmation of the factors affecting the mortality in patients hospitalized in the ICU. To identify the factors affecting the mortality of patients hospitalized in the ICU, a comprehensive literature review was conducted in Web of Science, PubMed, and Scopus databases. Data in this step were collected using the data extraction form. This form consisted of fields about the factors affecting the prediction of patient mortality in the ICU. Reviewing the relevant research and consulting with three medical informatics experts and a physician specializing in the ICU, content validity was also confirmed.

In the next step, a researcher-made questionnaire was used to approve finally the identified effective factors. This questionnaire consisted of three parts: demographic information of specialists, determining the importance degree of each effective factor by specialists, and other factors. It is worth saying that the other factors section was included to consider the forgotten factors in the questionnaire. The face and content validity of the questionnaire was confirmed by four specialist physicians of the ICU. Moreover, the questionnaire reliability was calculated using Cronbach's alpha formula. Cronbach's alpha was calculated 0.897. To analyze the data about the effective factors, each part of the questionnaire was classified into five levels: "very high", "high", "medium", "low" and "very low". Therefore, the scoring scale was considered one to five for each of the factors.

To approve the factors, 12 physicians working in the ICU in Rasoul Akram and Shahid Hasheminejad hospitals were called. One week later, six people stressed their readiness to participate in the study. In the next step, the questionnaire was distributed among these six people. After collecting the data, they were analyzed using SPSS software version 23. After calculating the standard deviation and the mean of the effective factors, based on the participating specialist physicians' opinions, the factors that had a mean of less than 3.5 were excluded from the study. According to the experts' opinions, these factors were not very significant to predicate the mortality in patients hospitalized in the ICU.

After confirming the factors, a form was designed to extract the data from the files of patients hospitalized in the ICU. This form included fields with patient file numbers and factors approved in the previous step. Its validity was also confirmed based on the three specialist physicians' opinions working in the ICU. To extract and collect the records values in the patients' records, the researchers referred to the records archive of two Hazrat Rasoul Akram and Shahid Hasheminejad hospitals. A total of 800 records (559 alive and 241 dead) were studied. Data were collected from July 1 to September 15, 2020.

The criterion for selecting the records included those recorded in the initial 24 hours for patients hospitalized in the ICU in the last two years. In the last step of this phase, all collected data were entered into an Excel file.

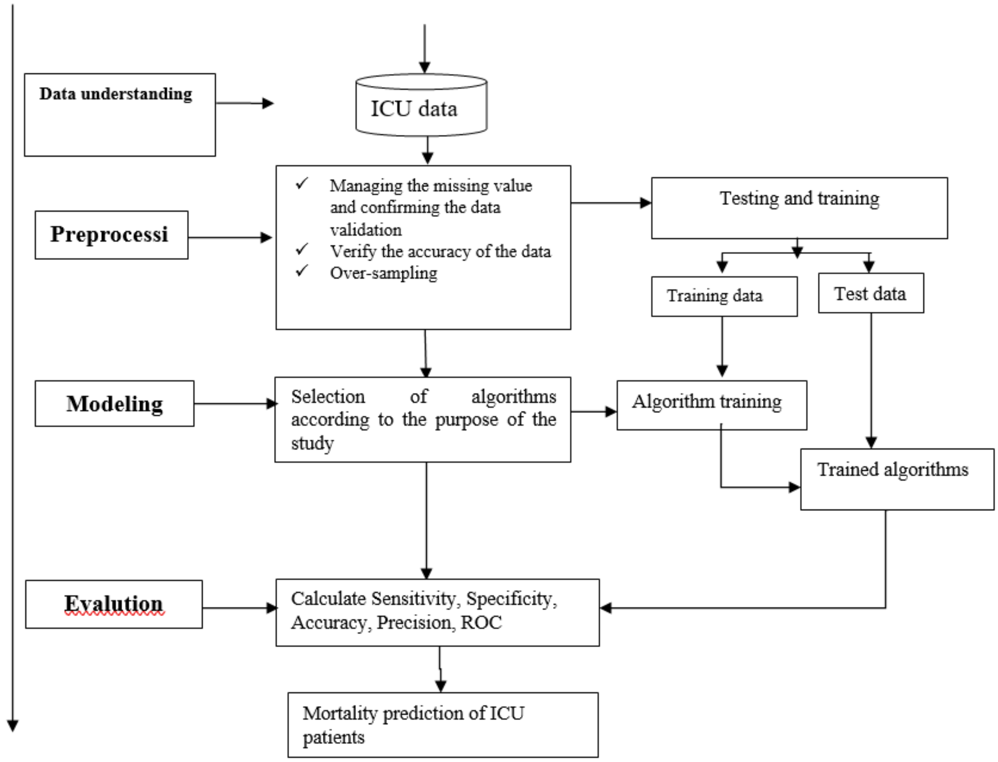

Fig. 1. Proposed model for research.

3.2 Step 2: Preprocessing

After collecting data, they were sorted and managed based on the type of the data. The records with 85% or higher values missing were excluded from the study. Moreover, the variable of patients' file number was excluded. The patients' names were anonymously added to the database (using numbers one to 800). Subsequently, the data collection created for data preprocessing, i.e., managing the missing value and confirming the data validation, was re-examined more accurately. Furthermore, the number of blank records was calculated for each of the identified factors affecting predicting the mortality of patients in the ICU. The blank records and data entered outside of the normal or possible range were reviewed and corrected if needed.

Furthermore, since the initial sample size consisted of 241 dead patients and 559 alive, the data collection was unbalanced. This meant that the number of records in one category was higher than the others (the number of samples in a larger category could be a maximum twice more than the other categories). As a result, predicting the classified models would have a bias towards having a larger sample size. Therefore, the values of the functional indicators in the classified models will be decreased. To solve this problem, two sampling methods are defined: under-sampling and over-sampling. In under-sampling, the high-class size is decreased to be equal to the low-class size, and in over-sampling, low class size is increased to be equal with high-class size. In this study, over-sampling (SMOTE) was used to balance the class records for dead with alive ones. Therefore, the classes are balanced. Data mining models were then performed on the balanced sample [35]. After using SMOTE, the sample size was changed to 559 dead patients and 482 alive.

3.3 Step 3: Implementing the selected data mining techniques (modeling)

In this step, to select machine learning techniques, the different studies conducted in these fields were examined [9, 36-38]. Subsequently, according to the studies, the type, and quality of data, the appropriate techniques were selected. SPSS version 23 was used to analyze the data. Weka v3.9.2 software was used to analyze and calculate the curves, criteria, and draw the confusion matrix. This academic software was used due to its various data preprocessing capabilities, including all models in the area of artificial intelligence and having the open-source capability [39].

3.4 Step 4: Evaluating and validating the models's performance

Sensitivity, specificity, accuracy, precision, and ROC area under curve indices were used to evaluate the performance of predicting models. These criteria will be defined and calculated using the confusion matrix components (Table 1). In this study, accuracy refers to the ratio of the number of alive and dead people who have been diagnosed accurately. Precision means the number of people who have died and the model has accurately identified them. Sensitivity refers to the ratio of dead people that the model has accurately identified as dead. Therefore, the higher the value is, the more accurate the diagnosis of the dead is. Specificity is a ratio of people who are alive and the model has accurately identified them as alive [40]. ROC is also used as an indicator to determine a model's power. Furthermore, in the medical field, the area under the ROC curve is used to evaluate the accuracy of diagnostic tests [41].

Table 1. Confusion matrix.

Output		Prediction value	
		Death(+)	**Living(-)**
Real value	**Death(+)**	TP	FN
	Living(-)	FP	TN

*True positive (TP): The number of deaths that the model has correctly identified.

*False positive (FP): The number of living people but the model has incorrectly identified them as dead.

*True negative (TN): The number of people who are living and the model correctly identified them as living.

*False negative (FN): The number of people who are dead but the model has incorrectly identified them as living [42].

Finally, for each of the selected models, the performance was reported. The model that had better results than other models was reported as the best model for predicting the mortality in patients hospitalized in the ICU.

4. Ethical Considerations

To conduct this study, the code of ethics with the number IR.KMU.REC.1399. 338 was obtained from the ethics committee of Kerman University of Medical Sciences. The specialist physicians participated voluntarily and there was the opportunity to leave the study at any time and without any consequences. All patients' data were collected by the research team confidentially and without registering their identity data. These data were not made available to individuals outside the research team. Furthermore, according to the research type, using data did not pose any risk to patients.

5. Results

The research findings are following based on the research steps.

5.1 Step 1: Data understanding

According to the reviewed literature and after removing duplicates, total of 49 effective factors involved in the mortality in patients hospitalized in the ICU were extracted from the studies. These 49 factors, along with the mean and standard deviation obtained by the physicians who are specialists in the ICU and anesthesia major are listed in Table 3. The demographic information of these physicians is also shown in Table 2.

Table 2. Physicians' demographics.

Variable		N	percentage of frequency
Gender	Male	3	50
	Female	3	50
Age	25-40	3	50
	41-55	2	33.33
	>55	1	15.66
Education degree	Specialist	3	50
	Sub-specialist	3	50
Type of specialty	Intensive care	4	66.64
	Anesthesia	2	33.33
Employment history (year)	1-15	4	66.64
	16-30	2	33.33

According to Table 3, two factors of Ventilate and Glasgow Coma Scale (GCS) with a mean of 4.59 and then two factors of age and vasoconstrictor drugs with a mean of 4.47 were introduced as the most effective predicting factors of mortality of patients hospitalized in the ICU. Cardiopulmonary resuscitation factors before admission, CRF, CKD, Serum lactate, existence of chronic diseases, surgical procedures during the ICU hospitalization period, SGOT, SGPT, PaO2, type of hospital admission, BMI, chronic organ failure and C-reactive protein factors according to the specialists' opinion with a mean less than 3.5 were excluded. In addition, two suggested factors "addiction" and "diabetes" were added to the collection of effective factors by the physicians participating in the study. A total of 36 effective factors were approved by specialists to predict the mortality in patients hospitalized in the ICU.

The ICU patient database was collected from 800 patients hospitalized (241 dead and 559 alive) in two Hazrat Rasoul Akram and Shahid Hasheminejad hospitals. Dead people and living population in this study were 30.1%, 69.9%, respectively. In terms of gender, 58.4% were male and 41.6% were female. The frequency of dead men (18.6) was higher than women (11.5). These statistical details are presented in Table 4.

Table 4. Percentage of the frequency of prediction factors based on gender.

Gender	Percentage of living people	Percentage of dead people	Total
Male	39.8	18.6	58.4
Female	30.1	11.5	41.6

Table 3. Predicting factors for the mortality in patients hospitalized in the ICU.

Row	Effective factors	Specialists' perspectives	
		Mean	±SD
1	Ventilate	4.59	±0.87
2	Glasgow Coma Scale(GCS)	4.59	±0.87
3	Age	4.47	± 1.00
4	Vasoconstrictor drugs	4.47	± 0.73
5	Diastolic pressure	4.46	±1.04
6	Systolic pressure	4.44	± 1.00
7	Heart rate	4.44	± 0.73
8	Estimated Glomerular Filtration Rate (eGFR)	4.35	±0.48
9	Respiratory rate	4.1	±0.81
10	Creatinine(Cr)	4.1	±0.57
11	Partial Pressure of Carbon Dioxide (PCO2)	4.1	±0.37
12	The potential of Hydrogen (PH)	4.1	±0.57
13	ICU length of stay	4.1	±0.96
14	Fraction of Inspired Oxygen (FiO2)	4.1	±0.81
15	Serum sodium	3.89	±0.37
16	Urine output	3.89	±0.37
17	(Glucose (Glu	3.89	±0.58
18	Partial Pressure Of Oxygen (PO2)	3.89	±0.69
19	Albumin (Alb)	3.86	±0.69
20	Hematocrit (HCT)	3.73	±0.95
21	Hemoglobin (Hb)	3.73	±0.95
22	Patient status after surgery	3.73	±0.48
23	Bicarbonate (HCO3)	3.73	±0.48
24	Serum potassium	3.73	±0.48
25	White blood cells (WBCs)	3.64	±0.75
26	Total Bilirubin	3.64	±0.75
27	Mean Arterial Pressure (MAP)	3.63	±0.72
28	Blood urea nitrogen (BUN)	3.6	±0.78
29	Gender	3.6	±0.78
30	Metastatic cancer	3.6	±0.78
31	Partial Thromboplastin Time (PTT)	3.6	±0.53
32	Rectal temperature	3.57	±0.97
33	Platelet Count (PLT)	3.55	±0.53
34	Infection	3.55	±0.53
35	Serum cortisol	3.55	±0.53
36	Readmission	3.50	±0.53
37	Cardiopulmonary resuscitation before admission	*	*
38	Chronic Renal Failure(CRF)	*	*
39	Chronic Kidney Disease (CKD)	*	*
40	Serum lactate	*	*
41	Existence of chronic diseases	*	*
42	Surgical procedures during the ICU hospitalization period	*	*
43	SGOT	*	*
44	SGPT	*	*
45	Partial Pressure of Oxygen (PaO2)	*	*
46	Type of hospital admission	*	*
47	Body mass index (BMI)	*	*
48	Chronic organ failure	*	*
49	C-reactive protein	*	*

Note:* These factors were removed by specialist physicians due to a mean score less than 3.5.

5.2 Step 2: Preprocessing

FiO2, eGFR, Serum cortisol, readmission, stay time length in the ICU, urine output and, infection factors with missing values above 85% were excluded from the data collection. The list of these factors along with the number of missing value and their percentage is presented in Table 5. Removing the factors with high missing values, 31 effective factors of morality prediction (29 factors from Table 3 along with two suggested factors of diabetes and addiction) were moved to the next step.

Table 5. The missing values of each effective factor in the mortality in patients hospitalized in ICU in the database.

Factors name	Missing value number for each factor	%
FiO2	800	100
eGFR	800	100
Serum cortisol	682	85.25
Readmission	800	100
ICU length of stay	790	98.85
Urine output	785	98.62
Infection	700	87.05

In total, according to specialists' opinions and removing the factors with missing values, 31 factors were finally used to predict the mortality in patients hospitalized in the ICU. Table 5 shows the importance of these 31 factors. The importance level of these factors was calculated using Weka v 3.9.2 software. The most effective factor used in the data collection of predicting the mortality of patients was ventilation. The factors in items 20 to 31 were also excluded from the study since they were less important to predict the mortality in patients hospitalized in the ICU.

According to Table 6, a total of 19 factors were used to predict the mortality in patients hospitalized in the ICU.

Table 6. Effective factors used in the data collection of predicting the mortality in patients hospitalized in the ICU.

Row	Factor's name	Values of each effective factor
1.	Ventilate	0.1915
2.	Glasgow Coma Scale(GCS)	0.1787
3.	Patient status after surgery	0.1604
4.	Gender	0.1496
5.	Rectal temperature	0.1153
6.	Creatinine(Cr)	0.0975
7.	Diabetics	0.08
8.	Age	0.0602
9.	Addictions	0.0519
10.	Vasoconstrictor drugs	0.0509
11.	Partial Thromboplastin Time (PTT)	0.039
12.	Blood urea nitrogen (BUN)	0.0298
13.	The potential of Hydrogen (PH)	0.0298
14.	Glucose (Glu)	0.0263
15.	Albumin(Alb)	0.0259
16.	Heart rate	0.0216
17.	Partial Pressure Of Oxygen (PO2)	0.0157
18.	Hematocrit(HCT)	0.0156
19.	Total Bilirubin	0.0134
20.	Metastatic cancer	0.0
21.	Hemoglobin(Hb)	0.0
22.	Mean Arterial Pressure (MAP)	0.0
23.	Diastolic pressure	0.0
24.	Systolic pressure	0.0
25.	Respiratory rate	0.0
26.	Partial Pressure of Carbon Dioxide (PCO2)	0.0
27.	Platelet Count (PLT)	0.0
28.	Serum potassium	0.0
29.	Serum sodium	0.0

30.	White blood cells(WBCs)	0.0
31.	Bicarbonate(HCO3)	0.0

5.3 Step 3: Implementing the selected data mining techniques

In this study, decision tree (J48), MLP, KNN, random forest, and SVM algorithms were used to make prediction models. It is worth noting that to decrease the minimum distance between query instance and the training samples, the KNN model was applied to the data collection six times [43-44]. All these algorithms were implemented using Weka v3.9.2 software.

There were 67 leaves used to make the decision tree. Finally, a tree-sized 133 was formed. The test option selected for the decision tree was Validation-Fold Cross-10. Experiments have shown that the best choice to estimate most accurately is Validation-Fold Cross-10 [45]. To create the neural network, its usual structure, MLP was used. Effective factors to predict the mortality and the target variable or patient death were neural network inputs and outputs, respectively. Therefore, the neural network used in the present study consisted of 31 input neurons, 10 hidden neurons, and one output (death or survival). Samples of data mining studies in the clinical area were cited to determine the hidden layers [12, 46, 47]. In these studies, the ability of two-layer neural networks was emphasized. It should be noted that one of these layers was the output layer and the other was the hidden layer. For KNNs, 1, 3, 5, 7, 9, and 11 neighbors were used. To analyze random forest, bagging with 100 iterations and base learner were used. In this model, similar to the decision tree algorithm, Validation-Fold Cross-10 was used to evaluate the algorithm and estimate most accurately [45]. The most common kernel, the Gaussian kernel or RBF, was used to create the SVM [48]. Validation-Fold Cross-10 was used to evaluate this algorithm and obtain the most accurate estimate [45]. It should be noted that to create each of these models, the studies using these models in different clinical areas were examined [12, 47, 49].

5.4 Step 4: Evaluating and validating the performance of the algorithms

At this step, the performance of decision tree (J48), multilayer perceptron, KNN, random forest, and SVM algorithms were evaluated. Subsequently, their sensitivity, specificity, accuracy, precision, and ROC indices were calculated. The comparison results of these indices for each of the algorithms are presented in Table 7. Sensitivity, specificity, accuracy, precision, and ROC of random forest were higher than other algorithms. SVM was the weakest algorithm according to ROC, sensitivity, accuracy, and precision. The specificity of the MLP algorithm was weaker than other algorithms.

Among the six KNN algorithms, KNN-7 performed better than the other five algorithms in terms of specificity, accuracy, precision, and ROC. Regarding the sensitivity criterion, the KNN-1 algorithm performed better than other algorithms.

Table 7. Algorithms used for predicting the mortality in the patients hospitalized in ICU.

Algorithm	Sensitivity	Specificity	Accuracy	Precision	ROC
MLP	68.67%	74.77%	71.95%	70.12%	0.7872
KNN-1	69.70%	78.35%	74.35%	73.52%	0.7395
KNN-3	67.84%	80.14%	74.44%	74.65%	0.8028
KNN-5	64.73%	83.36%	74.73%	77.03%	0.8072
KNN-7	64.31%	84.61%	75.28%	78.28%	0.8133
KNN-9	62.24%	83.18%	73.48%	76.14%	0.8128
KNN-11	62.44%	81.03%	72.43%	73.95%	0.8083
J48	71.16%	75.84%	73.67%	71.75%	0.7435
Random Forest	76.14%	86.76%	81.84%	83.21%	0.8884
SVM	56.43%	79.98%	65.32%	64.30%	0.7056

6. Discussion

This study investigated predicting the mortality in the patients hospitalized in the ICU based on machine learning algorithms using the collected data in the first 24 hours of admission to the ICU. Totally, 19 factors were used to predict the mortality in the patients in the ICU. Among these factors, Ventilate, GCS, gender, age, rectal temperature, and creatinine were considered significant as in other studies [50-54]. Among these factors, ventilate and the GCS were the most important. In the study of Awad et al., the GCS factor was introduced as one of the important factors to predict in-hospital mortality [21]. Ergan et al., in their study, emphasized the important role of ventilate condition to predict the mortality from High-Risk Pulmonary Embolism in the ICU [55].

Mohamed et al., [56] also showed that in predicting the mortality in infected patients hospitalized in the ICU, in addition to platelet count and CRP level in serum, a ventilated condition is also very important. Ghanem et al. and Lu et al. studies showed that low level of consciousness and GCS score have a relationship with predicting the mortality in patients hospitalized in the ICU [57, 58]. Brown et al. study's findings showed that age, blood urea nitrogen, respiration rate, and ventilation status are important indicators to predict in-hospital mortality [59]. The importance of the present study compared to other above-mentioned studies was to consider the specialists' opinion in the final confirmation of the effective factors. Other studies were limited to extracting the factors from similar studies or patient files.

In this study, two algorithm groups were used to develop the mortality prediction models for patients hospitalized in the intensive care units. Clear or interpretable models and unclear or black-box models, and clear classifications (such as decision trees) using linear patterns explain the hidden clinical consequences and analyze data. Therefore, decision tree-based algorithms provide the output knowledge in the form of a tree of different states, values, and variables. Demonstrating knowledge visually and in the form of a tree caused the categories based on the decision tree fully be interpretable and understandable [35].

In contrast, unclear categories such as random forest, KNN, MLP, and SVM are in the black box group. The most important criticism for this group, especially for the artificial neural network, is to understand and interpret the problem about their results for humans and how to calculate and obtain the output. Unclear classifications are less effective facing noisy data [35, 60]. In this study, the black-box algorithm of random forest based on ROC (0.8884%), accuracy (81.84%), accuracy (83.21%), sensitivity (86.76%) and, specificity (76.14%) had better performance to predict the mortality than other algorithms. According to ROC, sensitivity, accuracy, precision, SVM was the weakest algorithm.

Kim et al. [12] used a decision tree, artificial neural network, SVM, and APACHE III algorithms to predict mortality in the ICU. The models were developed based on 15 factors. According to the ROC, the decision tree-based model, artificial neural network, SVM, and APACHE III had better power, respectively. In the present study, however, the SVM-based model had the weakest performance. Besides, the accuracy obtained for all predicting models in the Kim et al. study was significantly higher than the present study. This result may be due to the large population (23,446 patients) under study or the factor number used to predict mortality. Xia et al. in their study developed models for predicting mortality in the ICU using an artificial neural network and two factor sets [61].

The results of this study show that neural networks with few variables had better performance compared with an artificial neural network with more variables [61]. Among the conducted studies, the study of

Hsieh et al. was more similar to the present study. This study developed mortality prediction models in a special group of the ventilated patients based on random forest, artificial neural network, and SVM. The result of this study, as in the present study, showed that the random forest algorithm had the best performance (91% ROc) to predict the death of these patients in the ICU [62]. Findings from Kandhasamy et al., study similar to the present study to predict diabetes mellitus recognized the performance of the KNN algorithm better than the J48 algorithm [63].

Finally, in this study, the ROC curve was used to compare the extracted knowledge from the mortality prediction models of the patients hospitalized in intensive care units.

For all used algorithms, sensitivity and specificity indices were not the same in terms of importance and specificity was more important than sensitivity which is why ROC was used for comparison. Since the medical staff exaggerates expressing the condition of the disease based on their principles, and in the end, the incorrect prediction of the death of the patient who finally survives is more acceptable compared with the incorrect prediction of the survival of the person who finally dies. According to the presented studies and the result of the present study to evaluate the performance of the algorithms performance and analyze their results, the performance of algorithms can be different from one database to another based on the number and type of factors [12]. Another point is that some factors do not affect predictions. For example, the factors items 20 to 31 in Table 5 did not affect predicting the patient mortality although from the treatment staff's point of view determining these factors is very important for the successful treatment of the patients and predicting their mortality. However, in machine learning techniques and data mining, many of these factors can be ignored, and the mortality rate can be predicted using fewer factors.

7. Study Limitations

Since it was impossible to collect data about patients hospitalized in intensive care units from all hospitals in Iran, the data was collected from only two hospitals. It is suggested that in another study, a large sample size should be collected or generalize it to several provinces. Since some physicians refused to cooperate in completing the questionnaire, only six specialist physicians' opinions to approve the effective factors were used. This challenge can affect the research results; therefore, it is suggested to use more specialists in future studies. As mentioned, like all machine learning models, these models can be viewed as "black box" models, which means that there is little insight into how factors play a role in predicting patient mortality. This problem can be decreased by ranking predictors after their contribution to the total AUC.

8. Conclusion

In this study, five machine learning models were used to predict the mortality of patients hospitalized in the intensive care units. A total of 19 important factors for predicting mortality were identified. Ventilation was the most effective factor used to predict patients' death. Random forest was the best algorithm for predicting mortality in patients hospitalized in the ICU. Also, the SVM algorithm was the weakest algorithm.

Based on these predicting models, the raw data can be best used in the ICU and useful data can be obtained by extracting their hidden patterns. As a result, it could help physicians in the process of patient treatment and decrease medical mistakes due to boredom and long working hours. It also could save patients' lives, reduce mortality, compensate for specialist shortages in the ICU, decrease the number of hospital admissions, and decrease the long-term treatment cost for patients, hospitals, and the insurance industry. Finally, it is suggested in addition to predicting the mortality rate in the patients hospitalized in ICU, the assessing models and classifying the patients' severity are also provided. These models can be used to compare the health status of the patients to each other, identify the most critically ill patients and, as a result, allocate the specialized ICU equipment and, facilities to the most critically ill patients.

Acknowledgments

The authors would like to thank all patients and experts who freely participated in this study.

Conflict of interest statement

Authors declare that there are no conflicts of interests.

References

[1] Rodriguez, A. M.; Gregorio, M. A.; Rodriguez, A. G., Psychological repercussions in family members of hospitalised critical condition patients. *Journal of psychosomatic research* 2005, 58, (5), 447-51.

[2] Ratanarat, R.; Thanakittiwirun, M.; Vilaichone, W.; Thongyoo, S.; Permpikul, C., Prediction of mortality by using the standard scoring systems in a medical intensive care unit in Thailand. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* 2005, 88, (7), 949-55.

[3] Lee, J.; Dubin, J. A.; Maslove, D. M., Mortality Prediction in the ICU. In *Secondary Analysis of Electronic Health Records*, Springer: 2016; pp 315-24.

[4] Ramon, J.; Fierens, D.; Güiza, F.; Meyfroidt, G.; Blockeel, H.; Bruynooghe, M.; Van Den Berghe, G., Mining data from intensive care patients. *Advanced Engineering Informatics* 2007, 21, (3), 243-56.

[5] Yang, S.; Wang, Z.; Liu, Z.; Wang, J.; Ma, L., Association between time of discharge from ICU and hospital mortality: a systematic review and meta-analysis. *Crit Care* 2016, 20, (1), 390.

[6] Purushotham, S.; Meng, C.; Che, Z.; Liu, Y., Benchmarking deep learning models on large healthcare datasets. *Journal of Biomedical Informatics* 2018, 83, 112-34.

[7] Silva, A.; Cortez, P.; Santos, M. F.; Gomes, L.; Neves, J., Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artificial intelligence in medicine* 2006, 36, (3), 223-34.

[8] Hsieh, M.-H.; Lin, S.-Y.; Lin, C.-L.; Hsieh, M.-J.; Hsu, W.-H.; Ju, S.-W.; Lin, C.-C.; Hsu, C. Y.; Kao, C.-H., A fitting machine learning prediction model for short-term mortality following percutaneous catheterization intervention: a nationwide population-based study. *Ann Transl Med* 2019, 7, (23), 732.

[9] Ibrahim, Z. M.; Wu, H.; Hamoud, A.; Stappen, L.; Dobson, R. J. B.; Agarossi, A., On classifying sepsis heterogeneity in the ICU: insight using machine learning. *J Am Med Inform Assoc* 2020, 27, (3), 437-43.

[10] Xu, J.; Zhang, Y.; Zhang, P.; Mahmood, A.; Li, Y.; Khatoon, S., Data mining on ICU mortality prediction using early temporal data: A survey. *International Journal of Information Technology & Decision Making* 2017, 16, (01), 117-59.

[11] Falini, S.; Angelotti, G.; Cecconi, M., ICU management based on big data. *Current opinion in anaesthesiology* 2020, 33, (2), 162-9.

[12] Kim, S.; Kim, W.; Park, R. W., A comparison of intensive care unit mortality prediction models through the use of data mining techniques. *Healthcare informatics research* 2011, 17, (4), 232.

[13] Thorsen-Meyer, H.-C.; Nielsen, A. B.; Nielsen, A. P.; Kaas-Hansen, B. S.; Toft, P.; Schierbeck, J.; Strøm, T.; Chmura, P.

J.; Heimann, M.; Dybdahl, L., Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *The Lancet Digital Health* 2020.

[14] Nemati, S.; Holder, A.; Razmi, F.; Stanley, M. D.; Clifford, G. D.; Buchman, T. G., An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. *Crit Care Med* 2018, 46, (4), 547-53.

[15] Pirracchio, R.; Petersen, M. L.; Carone, M.; Rigon, M. R.; Chevret, S.; van der Laan, M. J., Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med* 2015, 3, (1), 42-52.

[16] Ghorbani, R.; Ghousi, R.; Makui, A.; Atashi, A., A New Hybrid Predictive Model to Predict the Early Mortality Risk in Intensive Care Units on a Highly Imbalanced Dataset. *IEEE Access* 2020, 8, 141066-79.

[17] Moridani, M. K.; Setarehdan, S. K.; Nasrabadi, A. M.; Hajinasrollah, E., New algorithm of mortality risk prediction for cardiovascular patients admitted in intensive care unit. *Int J Clin Exp Med* 2015, 8, (6), 8916-26.

[18] Blom, M. C.; Ashfaq, A.; Sant'Anna, A.; Anderson, P. D.; Lingman, M., Training machine learning models to predict 30-day mortality in patients discharged from the emergency department: a retrospective, population-based registry study. *BMJ Open* 2019, 9, (8), e028015.

[19] Pourhomayoun, M.; Shakibi, M., Predicting mortality risk in patients with COVID-19 using artificial intelligence to help medical decision-making. *MedRxiv* 2020.

[20] Saritas, M. M.; Yasar, A., Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *International Journal of Intelligent Systems and Applications in Engineering* 2019, 7, (2), 88-91.

[21] Awad, A.; Bader-El-Den, M.; McNicholas, J.; Briggs, J., Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *International journal of medical informatics* 2017, 108, 185-95.

[22] Sharma, A.; Shukla, A.; Tiwari, R.; Mishra, A. In *Mortality Prediction of ICU patients using Machine Leaning: A survey*, Proceedings of the International Conference on Compute and Data Analysis, 2017; 2017; pp 49-53.

[23] Bhattacharya, S.; Rajan, V.; Shrivastava, H. In *Icu mortality prediction: A classification algorithm for imbalanced datasets*, Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017; 2017; pp 1288-94.

[24] Yun, K.; Oh, J.; Hong, T. H.; Kim, E. Y., Prediction of Mortality in Surgical Intensive Care Unit Patients Using Machine Learning Algorithms. *Front Med (Lausanne)* 2021, 8, 621861.

[25] Silva, A.; Cortez, P.; Santos, M. F.; Gomes, L.; Neves, J., Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artificial intelligence in medicine* 2006, 36, (3), 223-34.

[26] Le Gall, J.-R.; Loirat, P.; Alperovitch, A.; Glaser, P.; Granthil, C.; Mathieu, D.; Mercier, P.; Thomas, R.; Villers, D., A simplified acute physiology score for ICU patients. *Critical care medicine* 1984, 12, (11), 975-7.

[27] Lemeshow, S.; Teres, D.; Klar, J.; Avrunin, J. S.; Gehlbach, S. H.; Rapoport, J., Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *Jama* 1993, 270, (20), 2478-86.

[28] Awad, A.; Bader-El-Den, M.; McNicholas, J.; Briggs, J.; El-Sonbaty, Y., Predicting hospital mortality for intensive care unit patients: time-series analysis. *Health informatics journal* 2020, 26, (2), 1043-59.

[29] Calvert, J.; Mao, Q.; Hoffman, J. L.; Jay, M.; Desautels, T.; Mohamadlou, H.; Chettipally, U.; Das, R., Using electronic health record collected clinical variables to predict medical intensive care unit mortality. *Ann Med Surg (Lond)* 2016, 11, 52-7.

[30] Sadeghi, R.; Banerjee, T.; Romine, W., Early hospital mortality prediction using vital signals. *Smart Health* 2018, 9, 265-74.

[31] Crawford, E. D.; Batuello, J. T.; Snow, P.; Gamito, E. J.; McLeod, D. G.; Partin, A. W.; Stone, N.; Montie, J.; Stock, R.; Lynch, J.; Brandt, J., The use of artificial intelligence technology to predict lymph node spread in men with clinically localized prostate carcinoma. *Cancer* 2000, 88, (9), 2105-9.

[32] Yakovlev, A.; Metsker, O.; Kovalchuk, S.; Bologova, E., Prediction of in-hospital mortality and length of stay in acute coronary syndrome patients using machine-learning methods. *Journal of the American College of Cardiology* 2018, 71, (11S), A242.

[33] Awad, A.; Bader–El–Den, M.; McNicholas, J., Patient length of stay and mortality prediction: a survey. *Health services management research* 2017, 30, (2), 105-20.

[34] Lee, J.; Maslove, D. M., Customization of a severity of illness score using local electronic medical record data. *Journal of intensive care medicine* 2017, 32, (1), 38-47.

[35] Han, J.; Pei, J.; Kamber, M., *Data mining: concepts and techniques*. Elsevier: 2011.

[36] Cheng, F. Y.; Joshi, H.; Tandon, P.; Freeman, R.; Reich, D. L.; Mazumdar, M.; Kohli-Seth, R.; Levin, M.; Timsina, P.; Kia, A., Using Machine Learning to Predict ICU Transfer in Hospitalized COVID-19 Patients. *Journal of clinical medicine* 2020, 9, (6).

[37] Huang, L.; Yin, Y.; Fu, Z.; Zhang, S.; Deng, H.; Liu, D., LoAdaBoost: Loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data. *PLoS One* 2020, 15, (4), e0230706.

[38] Feretzakis, G.; Loupelis, E.; Sakagianni, A.; Kalles, D.; Martsoukou, M.; Lada, M.; Skarmoutsou, N.; Christopoulos, C.; Valakis, K.; Velentza, A.; Petropoulou, S.; Michelidou, S.; Alexiou, K., Using Machine Learning Techniques to Aid Empirical Antibiotic Therapy Decisions in the Intensive Care Unit of a General Hospital in Greece. *Antibiotics (Basel, Switzerland)* 2020, 9, (2).

[39] Azuaje, F., Witten ih, frank e: Data mining: Practical machine learning tools and techniques 2nd edition. In BioMed Central: 2006.

[40] Ramsay, M. A.; Usman, M.; Lagow, E.; Mendoza, M.; Untalan, E.; De Vol, E., The accuracy, precision and reliability of measuring ventilatory rate and detecting ventilatory pause by rainbow acoustic monitoring and capnometry. *Anesthesia & Analgesia* 2013, 117, (1), 69-75.

[41] Krzanowski, W. J.; Hand, D. J., *ROC curves for continuous data*. Crc Press: 2009.

[42] Evans, J. R.; Fisher, R. P., Eyewitness memory: Balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Applied Cognitive Psychology* 2011, 25, (3), 501-8.

[43] Mansor, M. N.; Yaacob, S.; Nagarajan, R.; Che, L. S.; Hariharan, M.; Ezanuddin,

M. In *Detection of facial changes for icu patients using knn classifier*, 2010 International Conference on Intelligent and Advanced Systems, 2010; IEEE: 2010; pp 1-5.

[44] Daberdaku, S.; Tavazzi, E.; Di Camillo, B., A Combined Interpolation and Weighted K-Nearest Neighbours Approach for the Imputation of Longitudinal ICU Laboratory Data. *Journal of Healthcare Informatics Research* 2020, 1-15.

[45] Liu, J.; Lan, H.; Fu, Y.; Wu, H.; Li, P., Analyzing electricity consumption via data mining. *Wuhan University Journal of Natural Sciences* 2012, 17, (2), 121-5.

[46] Ryan, D. P.; Daley, B. J.; Wong, K.; Zhao, X. In *Prediction of ICU in-hospital mortality using a deep Boltzmann machine and dropout neural net*, 2013 Biomedical Sciences and Engineering Conference (BSEC), 21-23 May 2013, 2013; 2013; pp 1-4.

[47] Rezaeian, A.; Nasimi, F.; Pooralizadeh Moghadam, F., Predicting Mortality Rate of Preterm Infants in Neonatal Intensive Care Unit Using Artificial Neural Network Model. *J-Mazand-Univ-Med-Sci* 2016, 26, (140), 85-94.

[48] Ring, M.; Eskofier, B. M., An approximation of the Gaussian RBF kernel for efficient classification with SVMs. *Pattern Recognition Letters* 2016, 84, 107-13.

[49] Ding, Y.; Wang, Y.; Zhou, D., Mortality prediction for ICU patients combining just-in-time learning and extreme learning machine. *Neurocomputing* 2018, 281, 12-9.

[50] VijayGanapathy, S.; Karthikeyan, V. S.; Sreenivas, J.; Mallya, A.; Keshavamurthy, R., Validation of APACHE II scoring system at 24 hours after admission as a prognostic tool in urosepsis: A prospective observational study. *Investig Clin Urol* 2017, 58, (6), 453-9.

[51] Lee, S. H.; Shin, D. S.; Kim, J. R.; Kim, H., Factors associated with mortality risk in critical care patients treated with veno-arterial extracorporeal membrane oxygenation. *Heart & lung : the journal of critical care* 2017, 46, (3), 137-42.

[52] Pietraszek-Grzywaczewska, I.; Bernas, S.; Łojko, P.; Piechota, A.; Piechota, M., Predictive value of the APACHE II, SAPS II, SOFA and GCS scoring systems in patients with severe purulent bacterial meningitis. *Anaesthesiology intensive therapy* 2016, 48, (3), 175-9.

[53] Skrifvars, M. B.; Varghese, B.; Parr, M. J., Survival and outcome prediction using the Apache III and the out-of-hospital cardiac arrest (OHCA) score in patients treated in the intensive care unit (ICU) following out-of-hospital, in-hospital or ICU cardiac arrest. *Resuscitation* 2012, 83, (6), 728-33.

[54] Juneja, D.; Nasa, P.; Singh, O.; Javeri, Y.; Uniyal, B.; Dang, R., Clinical profile, intensive care unit course, and outcome of patients admitted in intensive care unit with dengue. *Journal of critical care* 2011, 26, (5), 449-52.

[55] 55. Ergan, B.; Ergün, R.; Çalışkan, T.; Aydın, K.; Tokur, M. E.; Savran, Y.; Koca, U.; Cömert, B.; Gökmen, N., Mortality Related Risk Factors in High-Risk Pulmonary Embolism in the ICU. *Can Respir J* 2016, 2016, 2432808.

[56] Mohamed, A. K. S.; Mehta, A. A.; James, P., Predictors of mortality of severe sepsis among adult patients in the medical Intensive Care Unit. *Lung India* 2017, 34, (4), 330-5.

[57] Ghanem, S.; Lamine, M.; Alaali, Y.; Almutawa, E.; Al Balooshi, M., Factors Affecting Mortality in Severe Traumatic Brain Injury. *Bahrain Medical Bulletin* 2017, 158, (5881), 1-4.

[58] Lu, H. Y.; Li, T. C.; Tu, Y. K.; Tsai, J. C.; Lai, H. S.; Kuo, L. T., Predicting long-term outcome after traumatic brain injury using repeated measurements of Glasgow Coma Scale and data mining methods. *Journal of medical systems* 2015, 39, (2), 14.

[59] Brown, L. M.; Calfee, C. S.; Matthay, M. A.; Brower, R. G.; Thompson, B. T.; Checkley, W.; National Institutes of Health Acute Respiratory Distress Syndrome Network, I., A simple classification model for hospital mortality in patients with acute lung injury managed with lung protective ventilation. *Crit Care Med* 2011, 39, (12), 2645-51.

[60] Xia, H.; Keeney, N.; Daley, B. J.; Petrie, A.; Zhao, X. In *Prediction of ICU In-Hospital Mortality Using Artificial Neural Networks*, ASME 2013 Dynamic Systems and Control Conference, 2013; American Society of Mechanical Engineers Digital Collection: 2013.

[61] Xia, H.; Keeney, N.; Daley, B. J.; Petrie, A.; Zhao, X. In *Prediction of ICU In-hospital Mortality Using Artificial Neural Networks*, Dynamic Systems and Control Conference, 2013; American Society of Mechanical Engineers: 2013; p V003T43A001.

[62] Hsieh, M. H.; Hsieh, M. J.; Chen, C.-M.; Hsieh, C.-C.; Chao, C.-M.; Lai, C.-C., Comparison of machine learning models for the prediction of mortality of patients with unplanned extubation in intensive care units. *Sci Rep* 2018, 8, (1), 1-7.

[63] Kandhasamy, J. P.; Balamurali, S., Performance Analysis of Classifier Models to Predict Diabetes Mellitus. *Procedia Computer Science* 2015, 47, 45-51.