



# Comparison of Machine Learning Tools for the Prediction of ICU Admission in COVID-19 Hospitalized Patients

Mostafa Shanbehzadeh <sup>1</sup>, Hamideh Haghiri <sup>2</sup>, Mohammad Reza Afrash <sup>3</sup>, Morteza Amraei <sup>4</sup>, Leila Erfannia <sup>5</sup> and Hadi Kazemi-Arpanahi <sup>6, 7, \*</sup>

<sup>1</sup>Department of Health Information Technology, Faculty of Paramedical, Ilam University of Medical Sciences, Ilam, Iran

<sup>2</sup>Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>3</sup>Student Research Committee, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>4</sup>Department of Health Information Technology, School of Allied Medical Sciences, Lorestan University of Medical Sciences, Khorramabad, Iran

<sup>5</sup>Department of Health Information Management, School of Management and Medical Informatics, Health Human Resources Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>6</sup>Department of Health Information Technology, Abadan University of Medical Sciences, Abadan, Iran

<sup>7</sup>Student Research Committee, Abadan Faculty of Medical Sciences, Abadan, Iran

\*Corresponding author: Department of Health Information Technology, Abadan University of Medical Sciences, Abadan, Iran. Email: h.kazemi@abadanums.ac.ir

Received 2021 July 12; Revised 2021 November 07; Accepted 2021 December 16.

## Abstract

**Background:** The rapid coronavirus disease 2019 (COVID-19) outbreak has overwhelmed many healthcare systems worldwide and put them at the edge of collapsing. As the capacity of intensive care units (ICUs) is limited, deciding on the proper allocation of required resources is crucial.

**Objectives:** This study aimed to create a machine learning (ML)-based predictive model of ICU admission among COVID-19 in-hospital patients at the initial presentation.

**Methods:** This retrospective study was conducted on 1225 laboratory-confirmed COVID-19 hospitalized patients during January 9, 2020 - January 20, 2021. The top clinical parameters contributing to COVID-19 ICU admission were identified based on a correlation coefficient at P-value < 0.05. Next, the predictive models were constructed using five ML algorithms. Finally, to evaluate the performances of models, the metrics derived from the confusion matrix, classification error, and receiver operating characteristic were calculated.

**Results:** Following feature selection, a total of 11 parameters were selected as the top predictors to build the prediction models. The results showed that the best performance belonged to the random forest (RF) algorithm with the mean accuracy of 99.5%, mean specificity of 99.7%, mean sensitivity of 99.4%, Kappa metric of 95.7%, and root mean squared error of 0.015.

**Conclusions:** The ML algorithms, particularly RF, enable a reasonable level of accuracy and certainty in predicting disease progression and ICU admission for COVID-19 patients. The proposed models have the potential to inform frontline clinicians and health authorities with quantitative tools to assess illness severity and optimize resource allocation under time-sensitive and resource-constrained situations.

**Keywords:** Coronavirus, COVID-19, Decision Support Systems, Intensive Care Unit, Machine Learning

## 1. Background

Since December 2019, the coronavirus disease 2019 (COVID-19) has been spreading aggressively worldwide, and the health systems have encountered an increasing shortage of critical care resources, such as personal protective equipment, intensive care unit (ICU) beds, and ventilators (1). So far, this virus has affected almost all countries, with more than two million deaths worldwide (2). Approximately 20% of COVID-19 patients need hospitalization, and 20%-30% of COVID-19 in-hospital patients are admitted to ICU (3). In Iran, the ICU admission rate is esti-

mated to be 32% of hospitalized patients, and the ICU death rate is about 39% (4). The ICU resources are currently restricted, and more than 50% of ICU beds are occupied under normal conditions (5).

The unexpected outbreak, rapid transmission, ambiguous disease course, and prognosis, as well as emerging new mutations of COVID-19, along with health systems being not prepared for large-scale epidemic responses, created a situation in which the management of intensive care resources is crucial (6). Under this condition, it is urgent to construct and test effective clinical risk prediction tools for appropriately triaging critical patients (7, 8). Ac-

cordingly, predicting the individual disease courses and outcomes is essential for triaging patients, customized care service provision, utilizing life-saving resources in the best possible way, and directing them toward vulnerable and at-risk sub-groups deteriorating to critical COVID-19. Furthermore, many problems resulting from the shortage of hospital resources can be overcome by predicting the risk of patient deterioration, determining the length of stay, using hospital resources efficiently, and managing bed turnover (9, 10).

Innovative approaches for the early identification and triaging of patients at the time of admission will be greatly beneficial and effective in estimating which patients are at the high risk of clinical deterioration and have poor outcomes requiring ICU admission (8, 9). In such conditions, the design and implementation of clinical decision support systems equipped with machine learning (ML)-based prediction models will be critical for the optimal use of limited hospital resources and supporting clinical decisions. The ML, as a sub-form of artificial intelligence technologies, provides new insight or knowledge through extracting functional patterns and applicable rules from the large raw datasets (7, 10).

In the previous researches, a large number of ML algorithms were trained for the prediction of COVID-19 disease progression, patient condition deterioration (9, 10), ICU hospitalization (8, 9, 11-13), and death (8, 11, 14-19).

## 2. Objectives

The present study aimed to develop and validate a data-driven framework using five ML techniques to predict the patients who need transfer to ICU and find out the contributing clinical predictors by analyzing the available data at the time of admission.

## 3. Methods

This retrospective single-center cross-sectional study was conducted in 2021 to predict admission to ICU based on selected data-driven ML techniques.

### 3.1. Dataset and Participants

In this study, a COVID-19 hospital-based registry database from Taleghani Hospital, Abadan, Iran, was reviewed retrospectively. Only hospitalized confirmed COVID-19 patients aged  $\geq 18$  years and admitted during January 9, 2020 - January 20, 2021, met our inclusion criteria. During this period, a total of 13885 suspected cases with COVID-19 were referred to Taleghani Hospital, 3350 of which were confirmed as COVID-19 by RT-PCR

test. The exclusion criteria entailed non-COVID-19 cases, non-hospitalized COVID-19 patients, cases with unknown disposition, patients under 18 years old, incomplete case records (missing more than 70%), and admission time before January 9, 2020, or after January 20, 2021. Following applying the exclusion criteria, finally, 1225 records were entered in the study (Figure 1).

### 3.2. Feature Selection

Feature selection is an effective technique for determining the most significant variables, reducing the dimensions of the dataset, and improving the efficiency of ML algorithms (20). The included cases are defined based on 53 primary risk factors. In the current study, the variables with a correlation coefficient value less than 0.05 (P-value < 0.05) were identified as influential risk factors in predicting ICU admission.

### 3.3. Model Development

To predict ICU admission for COVID-19 patients, several ML classification algorithms, including Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), were used. We applied a set of parameters as shown in Table 1.

**Table 1.** Parameters for ML Algorithms

| Model | Parameters                 |
|-------|----------------------------|
| KNN   | K = 1, 3, 5                |
| SVM   | Kernel function = Gaussian |
| RF    | -                          |
| DT    | -                          |
| ANN   | 57-10-5-2                  |

### 3.4. Study Roadmap and Experiment Environment

All experiments on the classification algorithms described in this study were implemented using Python version 3.7.7. The Python experiment environment offers a well-defined framework for researchers and developers to run and assess their ML models. The roadmap of the proposed prediction model in this work is depicted in Figure 2.

### 3.5. Preprocessing

In the preprocessing stage, in order to use the data effectively in classifiers, raw data input was performed utilizing several preprocessing techniques, such as deleting missing values (missing more than 70%), minimum and maximum scalar values, and standard scalar.

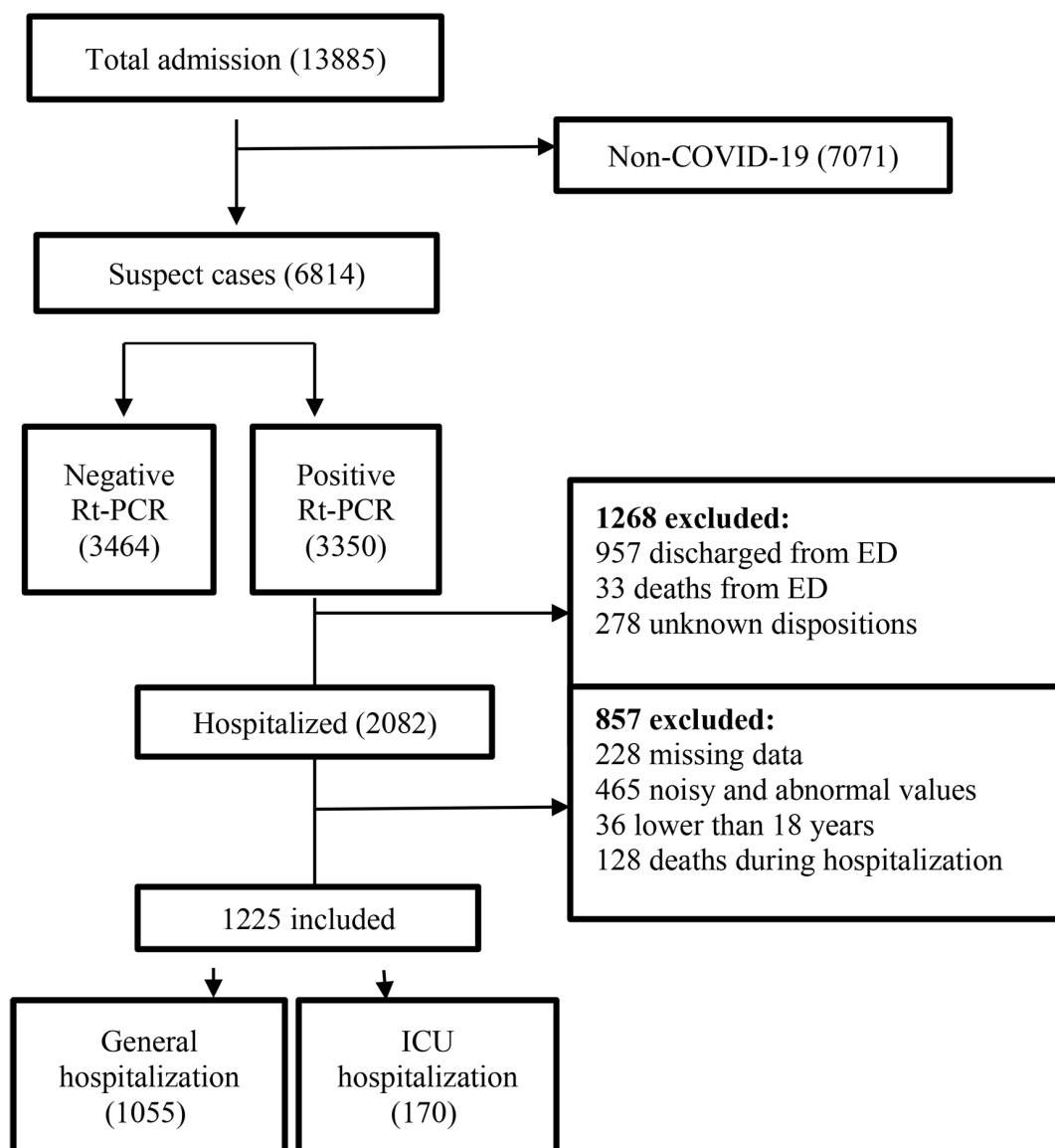


Figure 1. Flowchart describing patient selection

### 3.6. Experiment Evaluation

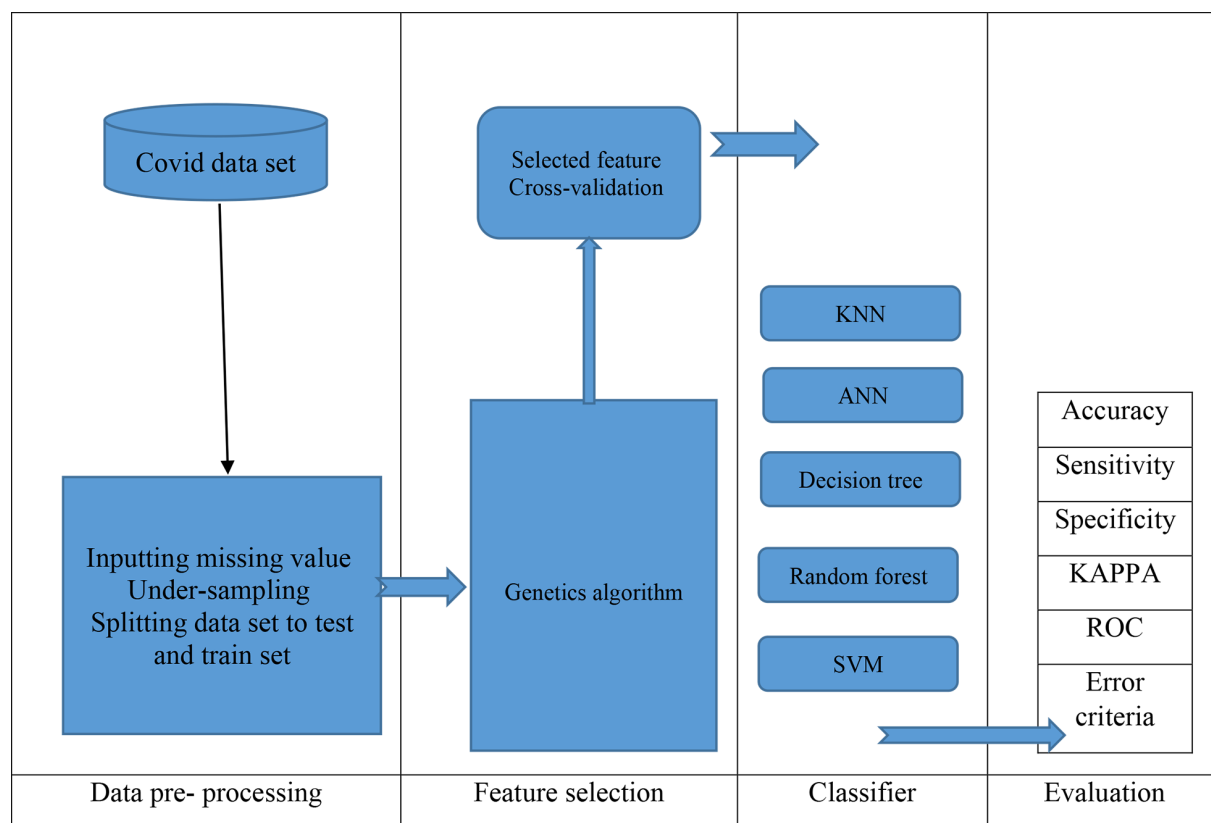
In this study, the performance of ML algorithms was calculated based on the 10-fold cross-validation method. To better compare the performance of algorithms, we assessed the effectiveness of five ML algorithms in terms of accuracy, specificity, sensitivity, error rate (Equations 1-4), receiver operating characteristic (ROC) curve, time to build the model, correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error (MAE), root mean squared error (RMSE), relative absolute error, and root relative squared error (RRSE).

$$\text{Classification accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

$$\text{Classification sensitivity} = \frac{TP}{TP + FN} \times 100 \quad (2)$$

$$\text{Classification specificity} = \frac{TN}{TN + FP} \times 100 \quad (3)$$

$$\text{Classification error} = \frac{FP + FN}{TP + TN + FP + FN} \times 100 \quad (4)$$



**Figure 2.** Block diagram of the proposed system for predicting ICU admission in patients with COVID-19

### 3.7. Ethical Considerations

The present study was approved by the Ethics Committee of Abadan University of Medical Sciences (Ethics code: IR.ABADANUMS.REC.1400.054). In order to protect the privacy and confidentiality of patients, we concealed the unique identification information of all patients in the process of data collection and presentation.

## 4. Results

### 4.1. Demographic and Clinical Characteristics

After applying the exclusion criteria, a total of 1225 patients were eligible (Figure 1). Of 1225 hospitalized COVID-19 cases, 664 (54.2%) were male, and 561 (45.8%) were women. Moreover, the median age of participants was 57.25 years (interquartile 18 - 100 years). We observed that 170 (13.87%) individuals were hospitalized in the ICU, and 1055 (86.13%) cases were hospitalized in general wards. Among the eligible patients, 1136 (92.75%) recovered, and 89 (7.25%) were deceased. Descriptive statistics for the 1225 records in this dataset are summarized in Table 2.

### 4.2. Top Predictors of ICU Admission

The most important predictors affecting ICU admission and disease progression were determined using the correlation coefficient at P-value < 0.05. As shown in Table 3, the top 11 ICU admission predictors were age, white cell count, neutrophil count, and lymphocyte count, blood urea nitrogen (BUN), aspartate transaminase (AST)/alanine transaminase (ALT), lactate dehydrogenase (LDH), cough, dyspnea, and oxygen therapy.

### 4.3. Model Development

In this study, to construct the ICU admission prediction model, we used five classification algorithms, namely ANN, KNN, SVM, DT, and RF, with k-fold ( $k = 10$ ) cross-validation methods. The mean metrics of 10-fold cross-validation methods were measured. Table 4 shows the 10-fold cross-validation results of five classifiers.

### 4.4. ANN Configuration

In this structure, according to Table 5, the input data consisted of 57 variables, two hidden layers with ten and five neurons, and two outputs (configuration: 57-10-5-2).

**Table 3.** Key Diagnostic Criteria at P-Value < 0.05

| Variables                             | Pearson's Correlation | P-Value | Variables                 | Pearson's Correlation | P-Value   |
|---------------------------------------|-----------------------|---------|---------------------------|-----------------------|-----------|
| Age                                   | -0.045                | *0.019  | Blood type                | -0.025128             | 0.274     |
| Height                                | -0.24                 | 0.409   | Gender                    | -0.107                | 0.38      |
| Weight                                | -0.25                 | 0.388   | Cough                     | 0.299                 | *0.041417 |
| Temperature                           | -0.32                 | 0.268   | Contusion                 | -0.122                | 0.342     |
| Creatinine                            | -0.066                | 0.119   | Hypertension              | -0.1744378            | 0.0541    |
| Red cell count                        | 0.029                 | 0.315   | Cardiovascular            | 0.2746594             | 0.125     |
| White cell count                      | -0.054                | *0.047  | Alcohol consumption       | 0.7923469             | 0.218     |
| Hematocrit                            | -0.017                | 0.562   | Smoking                   | 0.3123146             | 0.354     |
| Hemoglobin                            | -0.1                  | 0.724   | Diabetes                  | 0.0980716             | 0.104     |
| Platelet count                        | 0.018                 | 0.532   | Other underline disorders | 0.0904762             | 0.465     |
| Absolute lymphocyte count             | -0.057                | *0.044  | Sore throat               | 0.0591151             | 0.64      |
| Absolute neutrophil count             | 0.061                 | *0.033  | Runny nose                | -0.1446846            | 0.253     |
| Calcium                               | -0.055                | 0.055   | Loss of smell             | 0.0335175             | 0.811     |
| Phosphorus                            | -0.02                 | 0.476   | Loss of taste             | -0.1192558            | 0.414     |
| Magnesium                             | -0.033                | 0.243   | Dyspnea                   | 0.4443414             | *0.017    |
| Sodium                                | -0.015                | 0.59    | Oxygen therapy            | 0.460136              | *0.008    |
| Potassium                             | 0.015                 | 0.607   | Pneumonia                 | 0.2690936             | 0.115     |
| Bun                                   | -0.059                | *0.038  | Fever                     | 0.0241734             | 0.842     |
| Total bilirubin                       | -0.003                | 0.915   | Chill                     | 0.0269847             | 0.824     |
| Asp                                   | 0.054                 | *0.033  | Muscular pain             | 0.0885438             | 0.466     |
| Alt                                   | 0.047                 | *0.027  | GI complications          | -0.20181              | 0.179     |
| Albumin                               | 0.024                 | 0.394   | Headache                  | -0.0297449            | 0.831     |
| Glucose                               | 0.017                 | 0.552   | Vomit                     | 0.0525788             | 0.696     |
| Idh                                   | 0.056                 | *0.049  | Nausea                    | 0.0083105             | 0.949     |
| Activated partial thromboplastin time | -0.036                | 0.213   | ESR                       | 0.04                  | 0.157     |
| Prothrombin time                      | 0.01                  | 0.714   | Hyper sensitive troponin  | -0.2146846            | 0.439     |
| Alkaline phosphatase                  | -0.003                | 0.929   | C-reactive protein        | 0.0258788             | 0.196     |

**Table 4.** Average Performance of Ten Independent Runs of Classifiers Based on 10-Fold Cross-validation

| Evaluation Criteria              | Classifier |       |       |       |       |
|----------------------------------|------------|-------|-------|-------|-------|
|                                  | ANN        | DT    | KNN   | SVM   | RF    |
| Best time to build a model (s)   | 38         | 22    | 27    | 14    | 31    |
| Correctly classified instances   | 1,747      | 1,654 | 1,783 | 1,637 | 1,885 |
| Incorrectly classified instances | 148        | 241   | 112   | 258   | 10    |
| Mean accuracy (%)                | 92.2       | 87.3  | 94.1  | 86.4  | 99.5  |
| Mean specificity (%)             | 96.8       | 85.4  | 88.7  | 87.5  | 99.7  |
| Mean sensitivity (%)             | 87.6       | 89.3  | 99.5  | 85.3  | 99.4  |

**Table 5.** ANN Configuration

| Model      | Parameters                | Values  |
|------------|---------------------------|---|
| ANN        | Grid weights              | Random (between 1 and -1)   |
|            | Network error             | According to MSE  |
|            | Training ratio            | 0.7   |
|            | Validation ratio          | 0.15  |
|            | Test ratio                | 0.15  |
|            | Optimization              | Using Trainlm, which is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization |
|            | Maximum epochs            | 1000  |
|            | Maximum training time     | Inf   |
|            | Performance goal          | 0   |
|            | Minimum gradient          | min_grad: 1e-07   |
|            | Maximum validation checks | max_fail: 6   |
|            | Mu                        | 0.001   |
|            | Mu decrease ratio         | mu_dec: 0.1   |
|            | Mu increase ratio         | mu_inc: 10  |
| Maximum mu | mu_max: 10000000000       |   |

According to the experimental results of evaluating selected ML models in 10-iterations, the KNN algorithm (with  $K = 5$ ) had 94.1% accuracy, 88.7% specificity, and 99.5% sensitivity. After running ANN, the algorithm achieved an accuracy of 92.2%, a specificity of 96.8%, and a sensitivity of 87.6%. The RF algorithm showed high performance with 99.5% accuracy, 99.7% specificity, and 99.4% sensitivity. The SVM algorithm had 87.5%, 85.3%, and 86.4% specificity, sensitivity, and accuracy, respectively. Furthermore, the DT algorithm demonstrated 87.3% accuracy, 85.4% specificity, and 89.3% sensitivity. Based on the analysis of variance, the five selected algorithms were significantly different ( $P < 0.05$ ). The error rate of classifiers and computation time of implemented models on the given dataset are shown in [Figure 3](#).

The results obtained for the error rate of selected classifiers revealed that the RF algorithm had the highest Kappa metric of 95.7%. In addition, it has the lowest MAE (0.02), RMSE (0.015), and RRSE (22%). The results of comparing confusion matrix metrics and area under the curve (AUC)-ROC of different classifiers are shown in [Figures 4 and 5](#).

According to [Figures 4 and 5](#), the RF algorithm was the best classifier for predicting ICU admission based on evaluation criteria. The SVM algorithm has the best computation time of processing with 14 s. The results for ten independent runs of RF are summarized in [Table 6](#).

**Table 6.** Performance of RF in Ten Independent Runs

| Run  | Accuracy | Specificity | Sensitivity |
|------|----------|-------------|-------------|
| 1    | 99.1425  | 99.5817     | 99.77514    |
| 2    | 99.2141  | 99.458      | 99.8541     |
| 3    | 99.3236  | 99.6745     | 99.75114    |
| 4    | 99.2451  | 99.8542     | 99.347      |
| 5    | 99.6173  | 99.5841     | 99.471      |
| 6    | 99.657   | 99.37414    | 99.3445     |
| 7    | 99.4235  | 99.335      | 99.5457     |
| 8    | 99.741   | 99.421      | 99.247      |
| 9    | 99.2351  | 99.9524     | 99.741      |
| 10   | 99.541   | 99.3541     | 99.8242     |
| Std  | 99.41402 | 99.55891    | 99.59008    |
| Min  | 0.212369 | 0.21395     | 0.226258    |
| Max  | 99.1425  | 99.335      | 99.247      |
| Mean | 99.741   | 99.9524     | 99.8541     |

## 5. Discussion

Given the heterogeneity of the clinical manifestations of COVID-19, it is critical to develop models for predicting the likelihood of ICU admission by ML techniques. This study created five ML-based models using the most relevant variables in determining the risk of ICU admission derived from a correlation coefficient analysis. The tech-

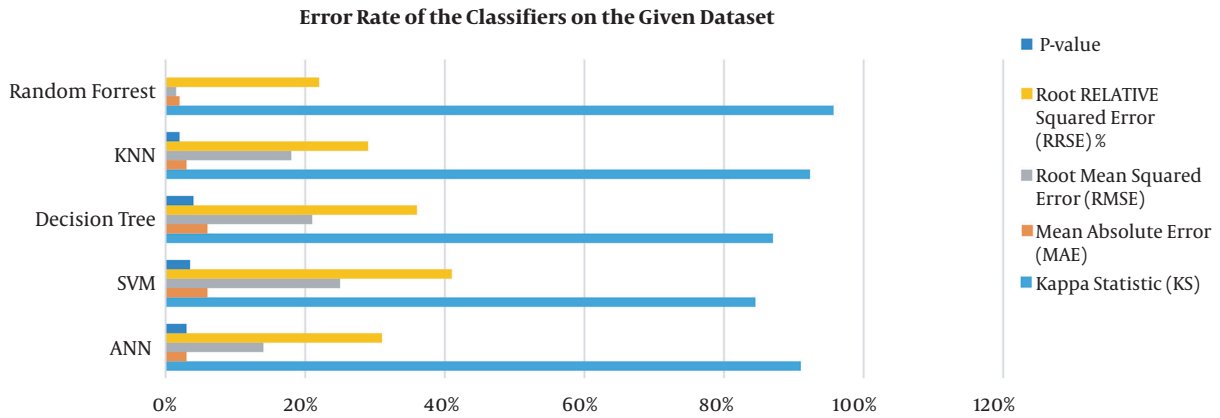


Figure 3. Error rate of the classifiers on the given dataset

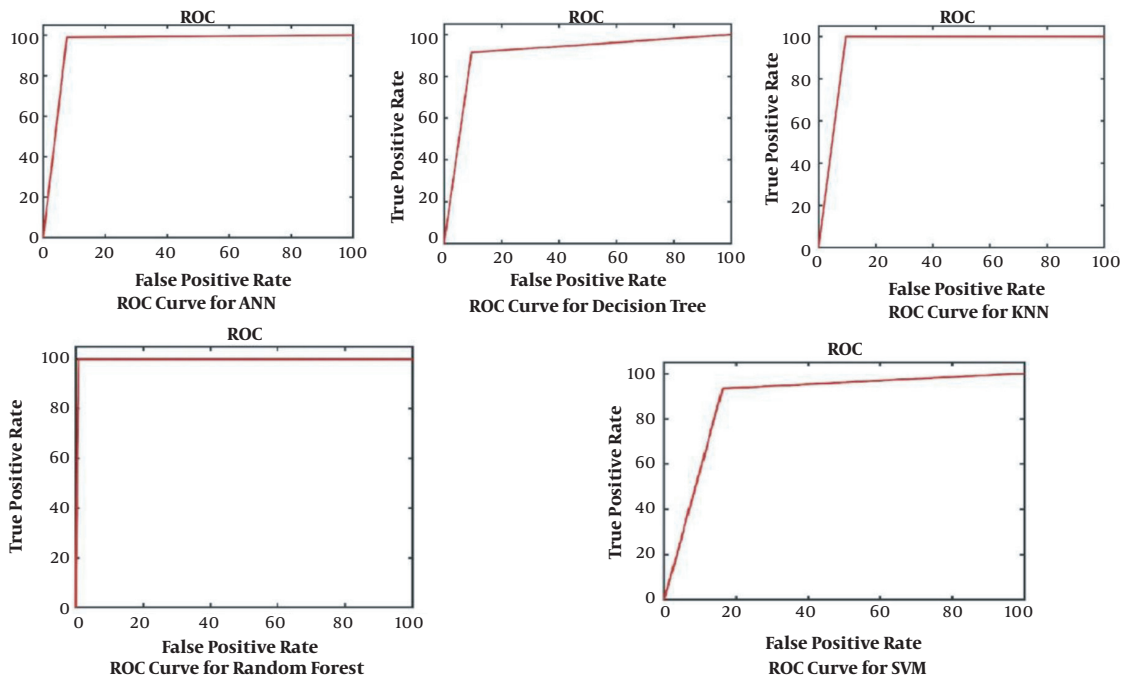


Figure 4. The ROC curve chart of the selected ML algorithms

niques used herein included ANN, DT, KNN, SVM, and RF, which were trained through the most significant predictors from 1225 laboratory-confirmed COVID-19 patients at the time of admission. Finally, based on our analysis of selected algorithms, we found that RF with a mean accuracy of 99.5%, a mean specificity of 99.7%, and a mean sensitivity of 99.4% have better performance than other ML algorithms in predicting the probability of ICU transfer after

hospital admission.

In the ICU, the need for informed decision-making is critical, especially in crisis circumstances, such as the current COVID-19 pandemic, where the healthcare systems encountered an increasing surge of patients and severe shortage in hospital resources (21, 22). The models developed in this study could be simply computerized as an alternative to manual and subjective clinical assessment methods. To

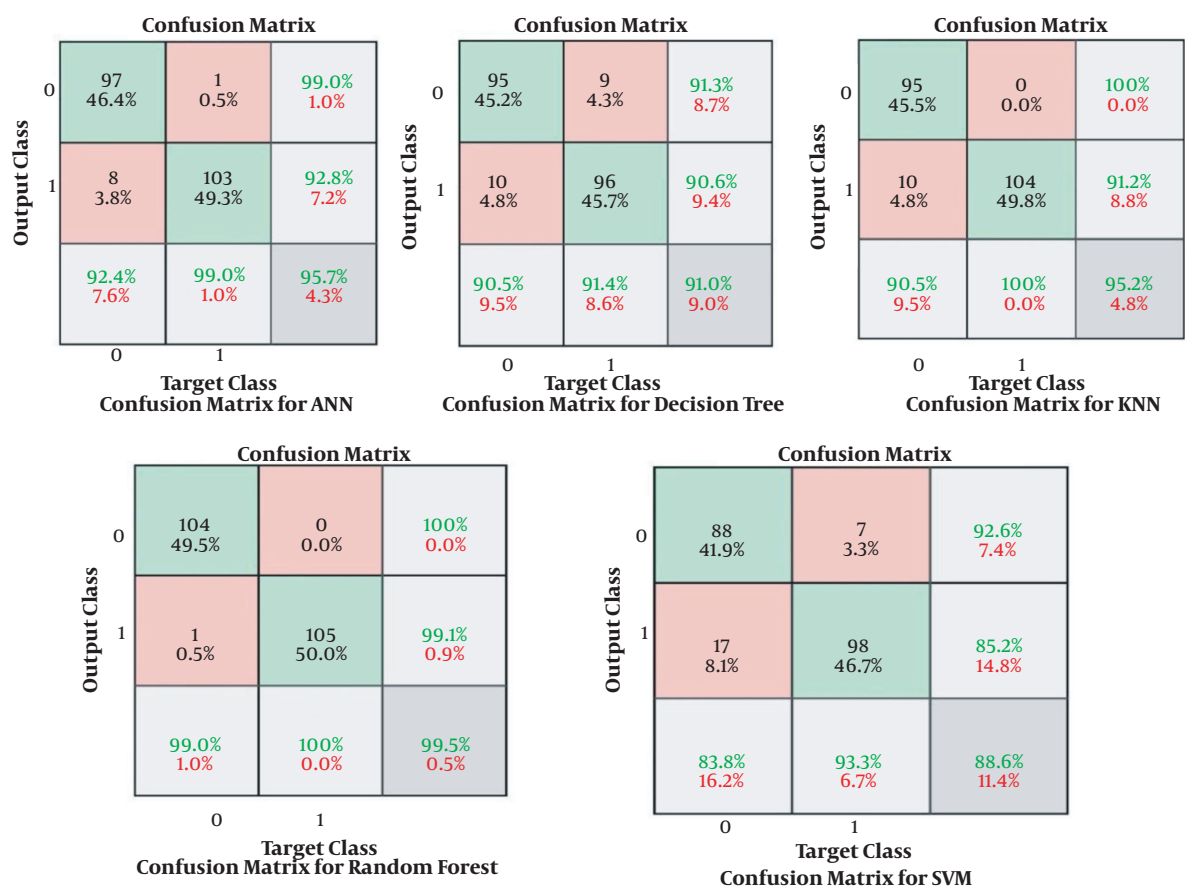


Figure 5. Confusion matrix calculation for the selected ML algorithms

correctly extract clinical predictors for estimating the potential need for ICU services, we evaluated clinical features at the time of admission and not at the progressive/severe course of the disease. In addition, the critical patients at admission time were discarded from the analysis. Thus, if validated, these features could be applied for predicting the likelihood to enter in ICU at the first hospitalization. For this purpose, Feature selection is a significant step to prepare the data before entering into the model (23). Hence, we identified the most important variables ( $n=11$ ) through correlation coefficient at  $P\text{-value} < 0.05$ . The most significant predictors of ICU admission were older age, high creatinine, leukocytosis, increased BUN, elevated ASP/ALT, augmented LDH, dry cough, hypertension, cardiovascular disorders, diabetes, dyspnea, decreased  $\text{SPO}_2$ , pneumonia, and high C-reactive protein.

Many studies have focused on determining the key risk factors for ICU admission (8, 9, 11, 13, 24, 25). The ten top clinical variables predicting ICU risk in reviewed studies en-

compassed age (older age), body temperature (high), oxygen saturation (decreased), neutrophil count and lymphocyte count (raised), C-reactive protein (elevated), D-dimer (increased), ALT and/or AST (augmented), LDH (elevated), loss of consciousness, and hypertension/cardiovascular diseases. In general, high compliance was observed between the results of reviewed studies and the most common variables in the current study.

In general, the developed ML algorithms in this study, similar to those reported in the previous studies (26), have achieved optimum results with an accuracy range of 86.4% - 90.37%. In particular, the experimental findings showed that RF had the best performance compared to the other four ML techniques with the mean accuracy of 99.5%, mean specificity of 99.7%, mean sensitivity of 99.4%, Kappa metric of 95.7%, and RMSE of 0.015. According to the results of the previous studies, the ANN and RF techniques have the most remarkable performance in predicting COVID-19 outcomes, which is consistent with the present study.



As a screening instrument for the development of severe disease, model developed in our study has several opportunities for clinical use. These models decrease the existing uncertainty and ambiguity in COVID-19 clinical practice by presenting measurable, non-subjective, and evidence-based approaches (12, 18). Accurate ICU admission prediction can support the sharing of limited hospital resources and improve the quality of care along with patient survival chance (12). The timely identification of at-risk patients could diminish the need for imminent ICU beds and invasive mechanical ventilators. Moreover, using proposed model in present study can surge the tolls of timely ICU transfers, resulting in reduced mortality and shorter lengths of ICU stay. Designing a scientific and valid ML-based prediction model would assist in early detection and effective supportive intervention to improve patient outcomes, the quality of care, and ultimately a reduction in the mortality rate of COVID-19 patients. Ambiguity declines due to offering quantitative, objective, and evidence-based models for risk stratification, prediction, and care planning (9, 10).

This study had several limitations. First, we retrospectively analyzed a dataset without control over data fields or incomplete data. Second, the dataset was extracted from a single hospital with a low sample size of 1225, making the results ungeneralizable. Third, this study only included 11 clinical features at admission to the hospital. It does not mean that these should be the only criteria for determining ICU admission. Longitudinal changes in these clinical features need to be investigated. Moreover, we only used five ML algorithms for prediction analyses. Finally, the selected dataset lacked some critical clinical variables, such as radiological indicators. In the future, the performance accuracy of our model and its generalizability will be enhanced if we test more ML techniques for larger, multicenter, and prospective datasets equipped with more qualitative and validated data.

### 5.1. Conclusions

We trained and validated different ML algorithms to predict the need for ICU transfer in COVID-19 hospitalized patients based on the data collected easily and routinely at the time of hospital admission. This study first identified the highly ranked clinical predictors that can predict the likelihood of ICU admission more precisely. Second, we developed and compared five ML-driven prediction models based on these selected predictors. It was observed that the RF model performed best on classification accuracy compared to the other ML algorithms. This method has the potential to provide frontline clinicians with an objective instrument to manage COVID-19 patients more efficiently in such time-sensitive, resource-demanding, stressful, and

potentially resource-constrained situations. Finally, the results of comparing the performance of prediction models in this study were satisfactory to some extent, and we believe that further investigations are needed to validate our model for a larger, multi-central, and more qualitative dataset.

### Acknowledgments

We would like to thank the research deputy of Abadan University of Medical Sciences for financially supporting this project (IR.ABADANUMS.REC.1400.054).

### Footnotes

**Authors' Contribution:** M.Sh. and H.K.A. designed the study. M.R.A. and H.H. performed data mining. M.A. performed the statistical analysis. LE and M.SH. reported the results.

**Conflict of Interests:** The authors declare that they have no conflicts of interest.

**Ethical Approval:** The current study was approved by the Ethics Committee of Abadan University of Medical Sciences (Ethics code: IR.ABADANUMS.REC.1400.054). In order to protect the privacy and confidentiality of patients, we concealed the unique identification information of all patients in the process of data collection and presentation.

**Funding/Support:** Abadan University of Medical Sciences

**Informed Consent:** All authors are aware of writing and submitting the article and confirmed the content.

### References

1. Basu S. Approaches to critical care resource allocation and triage during the COVID-19 pandemic: an examination from a developing world perspective. *J Med Ethics Hist Med.* 2021;**14**:5. doi: [10.18502/jmehm.v14i5.5652](https://doi.org/10.18502/jmehm.v14i5.5652). [PubMed: [34849214](https://pubmed.ncbi.nlm.nih.gov/34849214/)]. [PubMed Central: [PMC8595072](https://pubmed.ncbi.nlm.nih.gov/PMC8595072/)].
2. Vidal-Cevallos P, Higuera-De-La-Tijera F, Chavez-Tapia NC, Sanchez-Giron F, Cerda-Reyes E, Rosales-Salyano VH, et al. Lactate dehydrogenase associated with mortality in hospitalized patients with COVID-19 in Mexico: a multi-centre retrospective cohort study. *Ann Hepatol.* 2021;**24**:100338. doi: [10.1016/j.aohep.2021.100338](https://doi.org/10.1016/j.aohep.2021.100338). [PubMed: [33647501](https://pubmed.ncbi.nlm.nih.gov/33647501/)]. [PubMed Central: [PMC7908830](https://pubmed.ncbi.nlm.nih.gov/PMC7908830/)].
3. Smith EMT, Lee ACW, Smith JM, Thiele A, Zeleznik H, Ohtake PJ. COVID-19 and Post-intensive Care Syndrome: Community-Based Care for ICU Survivors. *Home Health Care Manag. Pract.* 2020;**33**(2):117-24. doi: [10.1177/1084822320974956](https://doi.org/10.1177/1084822320974956).
4. Abate SM, Ahmed Ali S, Mantfardo B, Basu B. Rate of Intensive Care Unit admission and outcomes among patients with coronavirus: A systematic review and Meta-analysis. *PLoS One.* 2020;**15**(7). e0235653. doi: [10.1371/journal.pone.0235653](https://doi.org/10.1371/journal.pone.0235653). [PubMed: [32649661](https://pubmed.ncbi.nlm.nih.gov/32649661/)]. [PubMed Central: [PMC7351172](https://pubmed.ncbi.nlm.nih.gov/PMC7351172/)].

5. Sadeghi A, Eslami P, Dooghaie Moghadam A, Pirsalehi A, Shojaee S, Vahidi M, et al. COVID-19 and ICU admission associated predictive factors in Iranian patients. *Caspian J Intern Med.* 2020;**11**(Suppl 1):512-9. doi: [10.22088/cjim.11.0.512](https://doi.org/10.22088/cjim.11.0.512). [PubMed: [33425268](https://pubmed.ncbi.nlm.nih.gov/33425268/)]. [PubMed Central: [PMC7780877](https://pubmed.ncbi.nlm.nih.gov/PMC7780877/)].
6. Elhadi M, Msherghi A, Alkeelani M, Alsuyihili A, Khaled A, Buzreg A, et al. Concerns for low-resource countries, with under-prepared intensive care units, facing the COVID-19 pandemic. *Infect Dis Health.* 2020;**25**(4):227-32. doi: [10.1016/j.idh.2020.05.008](https://doi.org/10.1016/j.idh.2020.05.008). [PubMed: [32631682](https://pubmed.ncbi.nlm.nih.gov/32631682/)]. [PubMed Central: [PMC7274573](https://pubmed.ncbi.nlm.nih.gov/PMC7274573/)].
7. Chin V, Samia NI, Marchant R, Rosen O, Ioannidis JPA, Tanner MA, et al. A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York state. *Eur J Epidemiol.* 2020;**35**(8):733-42. doi: [10.1007/s10654-020-00669-6](https://doi.org/10.1007/s10654-020-00669-6). [PubMed: [32780189](https://pubmed.ncbi.nlm.nih.gov/32780189/)]. [PubMed Central: [PMC7417851](https://pubmed.ncbi.nlm.nih.gov/PMC7417851/)].
8. Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. *PLoS One.* 2020;**15**(7). e0236618. doi: [10.1371/journal.pone.0236618](https://doi.org/10.1371/journal.pone.0236618). [PubMed: [32730358](https://pubmed.ncbi.nlm.nih.gov/32730358/)]. [PubMed Central: [PMC7392248](https://pubmed.ncbi.nlm.nih.gov/PMC7392248/)].
9. Agieb RS. Machine learning models for the prediction the necessity of resorting to icu of covid-19 patients. *Int. J. Adv. Trends Comput.* 2020;**9**(5):6980-4. doi: [10.30534/ijatcse/2020/15952020](https://doi.org/10.30534/ijatcse/2020/15952020).
10. Assaf D, Gutman Y, Neuman Y, Segal G, Amit S, Gefen-Halevi S, et al. Utilization of machine-learning models to accurately predict the risk for critical COVID-19. *Intern Emerg Med.* 2020;**15**(8):1435-43. doi: [10.1007/s11739-020-02475-0](https://doi.org/10.1007/s11739-020-02475-0). [PubMed: [32812204](https://pubmed.ncbi.nlm.nih.gov/32812204/)]. [PubMed Central: [PMC7433773](https://pubmed.ncbi.nlm.nih.gov/PMC7433773/)].
11. Ryan L, Lam C, Mataraso S, Allen A, Green-Saxena A, Pellegrini E, 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**:207-16. doi: [10.1016/j.amsu.2020.09.044](https://doi.org/10.1016/j.amsu.2020.09.044). [PubMed: [33042536](https://pubmed.ncbi.nlm.nih.gov/33042536/)]. [PubMed Central: [PMC7532803](https://pubmed.ncbi.nlm.nih.gov/PMC7532803/)].
12. Allenbach Y, Saadoun D, Maalouf G, Vieira M, Hellio A, Bodaert J, et al. Development of a multivariate prediction model of intensive care unit transfer or death: A French prospective cohort study of hospitalized COVID-19 patients. *PLoS One.* 2020;**15**(10). e0240711. doi: [10.1371/journal.pone.0240711](https://doi.org/10.1371/journal.pone.0240711). [PubMed: [33075088](https://pubmed.ncbi.nlm.nih.gov/33075088/)]. [PubMed Central: [PMC7571674](https://pubmed.ncbi.nlm.nih.gov/PMC7571674/)].
13. Pan P, Li Y, Xiao Y, Han B, Su L, Su M, et al. Prognostic Assessment of COVID-19 in the Intensive Care Unit by Machine Learning Methods: Model Development and Validation. *J Med Internet Res.* 2020;**22**(11). e23128. doi: [10.2196/23128](https://doi.org/10.2196/23128). [PubMed: [33035175](https://pubmed.ncbi.nlm.nih.gov/33035175/)]. [PubMed Central: [PMC7661105](https://pubmed.ncbi.nlm.nih.gov/PMC7661105/)].
14. Gao Y, Cai GY, Fang W, Li HY, Wang SY, Chen L, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun.* 2020;**11**(1):5033. doi: [10.1038/s41467-020-18684-2](https://doi.org/10.1038/s41467-020-18684-2). [PubMed: [33024092](https://pubmed.ncbi.nlm.nih.gov/33024092/)]. [PubMed Central: [PMC7538910](https://pubmed.ncbi.nlm.nih.gov/PMC7538910/)].
15. Hernandez-Suarez DF, Ranka S, Kim Y, Latib A, Wiley J, Lopez-Candales A, et al. Machine-Learning-Based In-Hospital Mortality Prediction for Transcatheter Mitral Valve Repair in the United States. *Cardiovasc Revasc Med.* 2021;**22**:22-8. doi: [10.1016/j.carrev.2020.06.017](https://doi.org/10.1016/j.carrev.2020.06.017). [PubMed: [32591310](https://pubmed.ncbi.nlm.nih.gov/32591310/)]. [PubMed Central: [PMC7736498](https://pubmed.ncbi.nlm.nih.gov/PMC7736498/)].
16. Parchure P, Joshi H, Dharmarajan K, Freeman R, Reich DL, Mazumdar M, et al. Development and validation of a machine learning-based prediction model for near-term in-hospital mortality among patients with COVID-19. *BMJ Support Palliat Care.* 2020. doi: [10.1136/bmjspcare-2020-002602](https://doi.org/10.1136/bmjspcare-2020-002602). [PubMed: [32963059](https://pubmed.ncbi.nlm.nih.gov/32963059/)]. [PubMed Central: [PMC8049537](https://pubmed.ncbi.nlm.nih.gov/PMC8049537/)].
17. Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, et al. Federated Learning of Electronic Health Records to Improve Mortality Prediction in Hospitalized Patients With COVID-19: Machine Learning Approach. *JMIR Med Inform.* 2021;**9**(1). e24207. doi: [10.2196/24207](https://doi.org/10.2196/24207). [PubMed: [33400679](https://pubmed.ncbi.nlm.nih.gov/33400679/)]. [PubMed Central: [PMC7842859](https://pubmed.ncbi.nlm.nih.gov/PMC7842859/)].
18. Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health.* 2020;**2**(10):e516-25. doi: [10.1016/S2589-7500\(20\)30217-X](https://doi.org/10.1016/S2589-7500(20)30217-X). [PubMed: [32984797](https://pubmed.ncbi.nlm.nih.gov/32984797/)]. [PubMed Central: [PMC7508513](https://pubmed.ncbi.nlm.nih.gov/PMC7508513/)].
19. Yan L, Zhang H-T, Goncalves J, Xiao Y, Wang M, Guo Y, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* 2020;**2**(5):283-8. doi: [10.1038/s42256-020-0180-7](https://doi.org/10.1038/s42256-020-0180-7).
20. Karegowda AG, Manjunath AS, Jayaram MA. Comparative study of attribute selection using gain ratio and correlation based feature selection. *International Journal of Information Technology and Knowledge Management.* 2010;**2**(2):271-7.
21. Leclerc T, Donat N, Donat A, Pasquier P, Libert N, Schaeffer E, et al. Prioritisation of ICU treatments for critically ill patients in a COVID-19 pandemic with scarce resources. *Anaesth Crit Care Pain Med.* 2020;**39**(3):333-9. doi: [10.1016/j.accpm.2020.05.008](https://doi.org/10.1016/j.accpm.2020.05.008). [PubMed: [32426441](https://pubmed.ncbi.nlm.nih.gov/32426441/)]. [PubMed Central: [PMC7230138](https://pubmed.ncbi.nlm.nih.gov/PMC7230138/)].
22. Tambone V, Boudreau D, Ciccozzi M, Sanders K, Campanozzi LL, Wathuta J, et al. Ethical Criteria for the Admission and Management of Patients in the ICU Under Conditions of Limited Medical Resources: A Shared International Proposal in View of the COVID-19 Pandemic. *Front Public Health.* 2020;**8**:284. doi: [10.3389/fpubh.2020.00284](https://doi.org/10.3389/fpubh.2020.00284). [PubMed: [32612972](https://pubmed.ncbi.nlm.nih.gov/32612972/)]. [PubMed Central: [PMC7308475](https://pubmed.ncbi.nlm.nih.gov/PMC7308475/)].
23. Sun L, Mo Z, Yan F, Xia L, Shan F, Ding Z, et al. Adaptive Feature Selection Guided Deep Forest for COVID-19 Classification With Chest CT. *IEEE J Biomed Health Inform.* 2020;**24**(10):2798-805. doi: [10.1109/JBHI.2020.3019505](https://doi.org/10.1109/JBHI.2020.3019505). [PubMed: [32845849](https://pubmed.ncbi.nlm.nih.gov/32845849/)].
24. Li X, Ge P, Zhu J, Li H, Graham J, Singer A, et al. Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables. *PeerJ.* 2020;**8**. e10337. doi: [10.7717/peerj.10337](https://doi.org/10.7717/peerj.10337). [PubMed: [33194455](https://pubmed.ncbi.nlm.nih.gov/33194455/)]. [PubMed Central: [PMC7651477](https://pubmed.ncbi.nlm.nih.gov/PMC7651477/)].
25. Zhou Y, He Y, Yang H, Yu H, Wang T, Chen Z, et al. Exploiting an early warning Nomogram for predicting the risk of ICU admission in patients with COVID-19: a multi-center study in China. *Scand J Trauma Resusc Emerg Med.* 2020;**28**(1):106. doi: [10.1186/s13049-020-00795-w](https://doi.org/10.1186/s13049-020-00795-w). [PubMed: [33109234](https://pubmed.ncbi.nlm.nih.gov/33109234/)]. [PubMed Central: [PMC7590555](https://pubmed.ncbi.nlm.nih.gov/PMC7590555/)].
26. Zhang Y, Xin Y, Li Q, Ma J, Li S, Lv X, et al. Empirical study of seven data mining algorithms on different characteristics of datasets for biomedical classification applications. *Biomed Eng Online.* 2017;**16**(1):125. doi: [10.1186/s12938-017-0416-x](https://doi.org/10.1186/s12938-017-0416-x). [PubMed: [29096638](https://pubmed.ncbi.nlm.nih.gov/29096638/)]. [PubMed Central: [PMC5668968](https://pubmed.ncbi.nlm.nih.gov/PMC5668968/)].

Table 2. Descriptive Statistics of the Study Variables After Preprocessing<sup>a</sup>

| Variables                        | Qualitative | Values  |
|----------------------------------|-------------|---------|
| <b>Blood type</b>                |             |         |
| A-, A+                           |             | 17, 552 |
| B-, B+                           |             | 13, 126 |
| O-, O+                           |             | 29, 421 |
| AB-, AB+                         |             | 6, 61   |
| <b>Gender</b>                    |             |         |
| Male                             |             | 664     |
| Female                           |             | 561     |
| <b>Cough</b>                     |             |         |
| Yes                              |             | 958     |
| No                               |             | 267     |
| <b>Contusion</b>                 |             |         |
| Yes                              |             | 409     |
| No                               |             | 816     |
| <b>Nausea</b>                    |             |         |
| Yes                              |             | 401     |
| No                               |             | 824     |
| <b>Vomit</b>                     |             |         |
| Yes                              |             | 346     |
| No                               |             | 879     |
| <b>Headache</b>                  |             |         |
| Yes                              |             | 312     |
| No                               |             | 913     |
| <b>Gastrointestinal symptoms</b> |             |         |
| Yes                              |             | 252     |
| No                               |             | 973     |
| <b>Muscular pain</b>             |             |         |
| Yes                              |             | 623     |
| No                               |             | 602     |
| <b>Chill</b>                     |             |         |
| Yes                              |             | 591     |
| No                               |             | 634     |
| <b>Fever</b>                     |             |         |
| Yes                              |             | 628     |
| No                               |             | 597     |
| <b>Pneumonia</b>                 |             |         |
| Yes                              |             | 1044    |
| No                               |             | 181     |

|                                  |      |
|----------------------------------|------|
| <b>Oxygen therapy</b>            |      |
| Yes                              | 1053 |
| No                               | 172  |
| <b>Dyspnea</b>                   |      |
| Yes                              | 1078 |
| No                               | 147  |
| <b>Loss of taste</b>             |      |
| Yes                              | 272  |
| No                               | 953  |
| <b>Loss of smell</b>             |      |
| Yes                              | 305  |
| No                               | 920  |
| <b>Runny nose</b>                |      |
| Yes                              | 437  |
| No                               | 788  |
| <b>Sore throat</b>               |      |
| Yes                              | 444  |
| No                               | 781  |
| <b>Other underlying diseases</b> |      |
| Yes                              | 735  |
| No                               | 490  |
| <b>Cardiac disease</b>           |      |
| Yes                              | 306  |
| No                               | 919  |
| <b>Hypertension</b>              |      |
| Yes                              | 395  |
| No                               | 830  |
| <b>Diabetes</b>                  |      |
| Yes                              | 268  |
| No                               | 957  |
| <b>Smoking</b>                   |      |
| Yes                              | 41   |
| No                               | 1184 |
| <b>Alcohol addiction</b>         |      |
| Yes                              | 11   |
| No                               | 1214 |
| <b>C-Reactive protein</b>        |      |
| Positive                         | 1063 |
| Negative                         | 162  |
| <b>Hypersensitive troponin</b>   |      |
| Positive                         | 58   |
| Negative                         | 1167 |
| <b>ICU admission (outcome)</b>   |      |

|  |                                       |
|--|---------------------------------------|
| Yes  | 1055                                  |
| No   | 170                                   |
| <b>Quantitative</b>                          |                                       |
| <b>Age (y)</b>                               | 57.25 ± 17.8 (18 - 100)               |
| <b>Height</b>                                | 168.53 ± 8.5 (92 - 195)               |
| <b>Weight</b>                                | 75.20 ± 13 (6.5 - 163)                |
| <b>Creatinine</b>                            | 1.39 ± 1.4 (0.1 - 17.9)               |
| <b>Red cell count</b>                        | 4.56 ± 0.9 (1.38 - 13.1)              |
| <b>White cell count</b>                      | 8182.34 ± 4897.4 (1300 - 63000)       |
| <b>Hematocrit</b>                            | 39.20 ± 6.7 (3.6 - 73.9)              |
| <b>Hemoglobin</b>                            | 13.21 ± 2.4 (3.7 - 46)                |
| <b>Platelet count</b>                        | 215493.66 ± 88380.1 (108000 - 691000) |
| <b>Absolute lymphocyte count</b>             | 23.74 ± 11.8 (2 - 95)                 |
| <b>Absolute neutrophil count</b>             | 74.52 ± 12.3 (8 - 98)                 |
| <b>Calcium</b>                               | 9.68 ± 0.8 (0.9 - 14.1)               |
| <b>Phosphorus</b>                            | 3.5 ± 0.5 (2 - 12.4)                  |
| <b>Magnesium</b>                             | 2.16 ± 0.6 (1.14 - 19.1)              |
| <b>Sodium</b>                                | 137.94 ± 5.3 (37 - 157)               |
| <b>Potassium</b>                             | 3.98 ± 0.7 (2.5 - 14.2)               |
| <b>Blood urea nitrogen</b>                   | 42.52 ± 31.7 (0.5 - 251)              |
| <b>Total bilirubin</b>                       | 0.72 ± 0.7 (0.01 - 10)                |
| <b>Aspartate aminotransferase</b>            | 44.45 ± 53.5 (3.8 - 924)              |
| <b>Alanine aminotransferase</b>              | 38.29 ± 41.6 (2 - 672)                |
| <b>Albumin</b>                               | 4.02 ± 0.5 (0.2 - 8.9)                |
| <b>Glucose</b>                               | 136.09 ± 74.2 (18 - 994)              |
| <b>Lactate dehydrogenase</b>                 | 555.68 ± 339 (4.6 - 6973)             |
| <b>Activated partial thromboplastin time</b> | 28.56 ± 11.4 (1 - 120)                |
| <b>Prothrombin time</b>                      | 12.82 ± 1.9 (0.9 - 46.8)              |
| <b>Alkaline phosphatase</b>                  | 213.12 ± 139.2 (9.6 - 2846)           |
| <b>Erythrocyte sedimentation rate</b>        | 40.65 ± 28.8 (2 - 258)                |

<sup>a</sup> Values are expressed as No. (%) or mean ± SD (range).