Prediction of COVID-19 Mortality in the Intensive Care Unit Using Machine Learning

Aikaterini SAKAGIANNI[a], Christina KOUFOPOULOU[b], Vassilios VERYKIOS[c],
Evangelos LOUPELIS[d], Dimitrios KALLES[c] and Georgios FERETZAKIS[c,e,1]

[a] *Sismanogleio General Hospital, Intensive Care Unit, Marousi, Greece*
[b] *Aretaieio Hospital, National and Kapodistrian University of Athens, Anesthesiology Department, Athens, Greece*
[c] *School of Science and Technology, Hellenic Open University, Patras, Greece*
[d] *Sismanogleio General Hospital, IT department, Marousi, Greece*
[e] *Sismanogleio General Hospital, Department of Quality Control, Research and Continuing Education, Marousi, Greece*

Abstract. Since its emergence, the COVID-19 pandemic still poses a major global health threat. In this setting, a number of useful machine learning applications have been explored to assist clinical decision-making, predict the severity of disease and admission to the intensive care unit, and also to estimate future demand for hospital beds, equipment, and staff. The present study examined demographic data, hematological and biochemical markers routinely measured in Covid-19 patients admitted to the intensive care unit (ICU) of a public tertiary hospital, in relation to the ICU outcome, during the second and third Covid-19 waves, from October 2020 until February 2022. In this dataset, we applied eight well-known classifiers of the caret package for machine learning of the R programming language, to evaluate their performance in forecasting ICU mortality. The best performance regarding area under the receiver operating characteristic curve (AUC-ROC) was observed with Random Forest (0.82), while k-nearest neighbors (k-NN) were the lowest performing machine learning algorithm (AUC-ROC: 0.59). However, in terms of sensitivity, XGB outperformed the other classifiers (max Sens: 0.7). The six most important predictors of mortality in the Random Forest model were serum urea, age, hemoglobin, C-reactive protein, platelets, and lymphocyte count.

Keywords. Artificial intelligence; machine learning; COVID-19; SARS-CoV-2; ICU—intensive care unit.

---

[1] Corresponding Author: Georgios FERETZAKIS, Ph.D., School of Science and Technology, Hellenic Open University, Patra 263 35, Greece, E-mail: georgios.feretzakis@ac.eap.gr.

1. Introduction and Background

The 2019 coronavirus disease (COVID-19) pandemic has posed an ongoing threat to global health since its emergence. As of 9 December 2022, 643,875,406 confirmed cases of COVID-19, including 6,630,082 deaths, have been reported to WHO [1]. One of the biggest challenges facing healthcare systems around the world during the Covid-19 pandemic is the surge in patient numbers and concomitant limited medical resources. Several recent studies suggest that machine learning-based predictive models of disease severity and outcome could meet these requirements in the best possible way permitting efficient allocation of resources in the intensive care unit [2-7]. However, no sufficiently validated prognostic models are currently in wide clinical use.

In this study, we aim to compare the performance of eight different machine learning methods for predicting the mortality of critically ill COVID-19 patients admitted to the Intensive Care Unit (ICU), using a simple dataset containing demographic information and results of routine laboratory tests. We also aim to identify the most important variables associated with ICU mortality using machine learning methods.

2. Methods and Materials

A retrospective observational study was performed in the Intensive Care Unit (ICU) of a tertiary public hospital during the second and third waves of the COVID-19 pandemic. The study was approved by the Institutional Review Board (IRB) of Sismanogleio General Hospital (no. 4333/2022). For this analysis, we used the caret package [8] in R [9], a well-known framework for building ML models. The caret package provides functions to support model training for complex classification and regression tasks. During a 17-month period from October 2020 until February 2022, all the patients with PCR-confirmed COVID-19 pneumonia that were admitted to the ICU were included in the study. The primary outcome was ICU mortality, and the secondary outcome was the identification of predictive variables for mortality.

We investigated patient demographic data, blood biomarkers routinely measured in Covid-19 patients admitted to the ICU, bloodstream infection occurrence, and ICU outcome. Analytically, the data set includes the following variables: age, gender, Procalcitonin (PCT), Creatine Kinase (CPK), Troponin, Creatinine (CREA), C-Reactive Protein (CRP), Lactate Dehydrogenase (LDH), serum levels of Urea (UREA), Ferritin (Ferr), serum glutamate-pyruvate transaminase (SGOT), serum glutamate-oxaloacetate transaminase (SGPT), total bilirubin (TBIL), activated partial thromboplastin time (aPTT), D-Dimer, International Normalized Ratio (INR), hemoglobin (HGB), white blood cells count (WBC), lymphocyte count (LYM%), neutrophil count (NEUT%), platelets (PLT), the presence of bloodstream infection (microbial or/and fungal, none) and the binomial ICU outcome (survival or death).

2.1. Data preprocessing

We used the laboratory values of the two first days of the ICU stay. Then after taking the first 7 observations in chronological order from each patient, we take the average in each variable and with this value, we replace the missing values and after that, we keep only the first two lines from each patient.

2.2. Model Building

The classification problem that we had to deal, with consisted of two classes of patients; the patients who survived and were discharged alive from the ICU and the patients who died during their ICU stay. We applied and evaluated the following classification models: linear discriminant analysis (method: lda), Recursive Partitioning and Regression Trees (method: rpart), k–nearest neighbor (method: knn), support vector machines (method: svmRadial), random forests (method: rf), eXtreme Gradient Boosting (method="xgbLinear"), AdaBoost Classification Trees (method="ada"), and Stochastic Gradient Boosting (method="gbm") to try to predict hospitalization. The caret package uses the train() function to build any predictive model. We trained our models through a 10-fold cross-validation (CV) procedure to avoid overfitting in our analysis.

2.3. Model Evaluation

The performance metrics of the ML techniques that were evaluated in this study were the Sensitivity, Specificity, and Area Under the Receiver Operating Characteristic Curve AUC ROC from prediction scores.

3. Results

Among 373 patients (140 female, 233 male) with COVID-19 pneumonia that were included in the study, 102 (27.34%) died in the ICU. Patients who died were older (mean [SD] age 70 yr [11.6 yr] vs 64.4 yr [13.5 yr]) and more likely to be female (28.6% vs 26.6%). The descriptive statistics for Age, Gender, and Outcome are summarized in Table 1. The best performance considering the area under the receiver operating characteristic curve (AUC-ROC) was observed with Random Forest (0.82), followed by XGB (0.80), while k-nearest neighbors (k-NN) were the lowest performing machine learning algorithm (0.59). In terms of sensitivity, XGB outperformed the other classifiers, with maximum sensitivity reaching 0.7. The corresponding AUC-ROC curves of the eight classifiers in the specific data set are shown in Figure 1. The relative importance of the examined variables with Random Forest sorted in descending order is shown in Figure 2. The six highest-ranked variables were Urea, Age, Hgb, CRP, PLT, and lymphocyte count.

Table 1. Descriptive statistics for Age, Gender, and Outcome

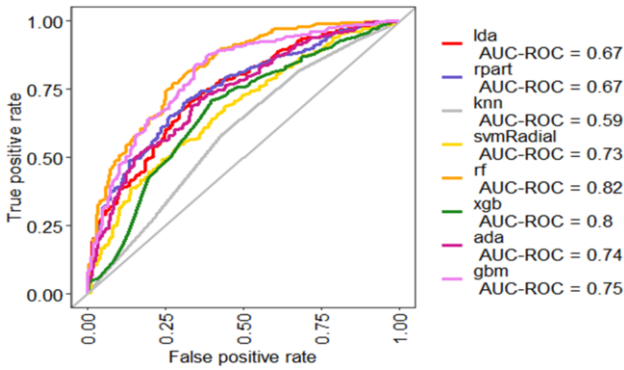| Age (years)		Gender (%)		Outcome (%)	
Median	68	Male	233 (62.46)	Survived	371 (72.65)
Mean (SD)	65.9 (13.21)	Female	140 (37.53)	Died	102 (27.34)
IQR	58-75	Total	373	Total	373

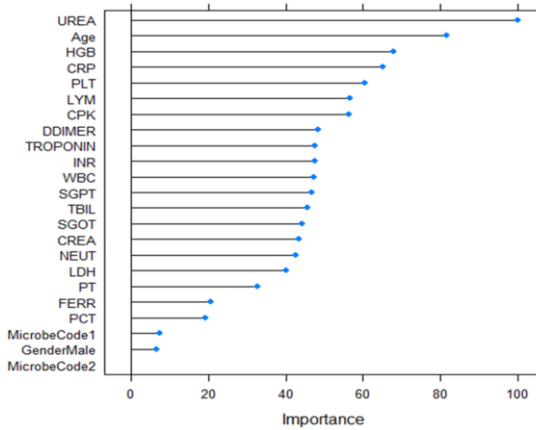**Figure 1.** Corresponding AUC-ROC curves of the eight classifiers.

Figure 2. Variable Importance with Random Forest

4. Discussion

Machine Learning approaches are increasingly used to support medical diagnosis or treatment [10, 11]. In this study, we evaluated eight classifiers to predict ICU mortality in critically ill COVID-19 patients. A random forest model outperformed other models in terms of AUC ROC among the various classification methods that were tested using a 10-fold CV. Renal function worsening and older age seem to be among the strongest predictors of unfavorable outcomes in compliance with similar studies [2,4,5,7,12]. Recent literature suggests that predictive modeling can be enhanced by using ML approaches, since they enable the detection of complex relationships between the outcomes of interest and the covariates, especially when trained in the individual data of a hospital, overcoming the limitations of traditional methods [2,3,6,13]. It is difficult to compare the performance of different predictive models, because each one was developed on patients with different characteristics, with different sets of attributes, and with different model development techniques [7].

There are several limitations to this study. First, it is a single-center study with data collected from a single ICU, restricting the generalizability and external validity of the

results. Second, the dataset is limited to demographic characteristics and routine laboratory results retrieved from the hospital information system (HIS) and does not contain clinical patient data, which would undoubtedly add significant accuracy to the predictive ability of the machine learning models. Third, a serious drawback of the proposed ML model (RF), is its limited sensitivity (max 0.5), which is apparently due to the lack of clinical data from the dataset. We anticipate that our future studies besides laboratory and demographic data will additionally include clinical data from patients over longer time periods.

5. Conclusion

We evaluated simple demographics, routine laboratory tests, and the presence of bloodstream infection during the first two days after admission to the ICU, for the prediction of ICU mortality. The best performance regarding AUC-ROC was observed with Random Forest (0.82), with the highest predictive values being urea, age, hemoglobin, CRP, platelet, and lymphocyte count.

References

[1]   World Health Organization. *WHO Coronavirus (COVID-19) Dashboard.* https://covid19.who.int/" https://covid19.who.int/ [Accessed 4th March 2023].
[2]   Magunia H, Lederer S, Verbuecheln R, Gilot BJ, Koeppen M, Haeberle HA, et al. Machine learning identifies ICU outcome predictors in a multicenter COVID-19 cohort. *Crit Care.* 2021; 25(1):295. doi: 10.1186/s13054-021-03720-4.
[3]   Laino ME, Generali E, Tommasini T, Angelotti G, Aghemo A, Desai A, et al. An individualized algorithm to predict mortality in COVID-19 pneumonia: a machine learning based study. *Archives of Medical Science.* 2022;18(3):587-595. doi: 10.5114/aoms/144980.
[4]   Jamshidi E, Asgary A, Tavakoli N, Zali A, Setareh S, Esmaily H, et al. Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. *Front Digit Health.* 2022; 3:681608. doi: 10.3389/fdgth.2021.681608.
[5]   Ustebay S, Sarmis A, Kaya GK, Sujan M. A comparison of machine learning algorithms in predicting COVID-19 prognostics. *Intern Emerg Med.* 2023;18(1):229-239. doi:10.1007/s11739-022-03101-x.
[6]   Elhazmi A, Al-Omari A, Sallam H, Mufti HN, Rabie AA, Alshahrani M, et al. Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. *J Infect Public Health.* 2022;15(7):826-834. doi: 10.1016/j.jiph.2022.06.008.
[7]   Lorenzoni G, Sella N, Boscolo A, Azzolina D, Bartolotta P, Pasin L, et al. COVID-19 ICU mortality prediction: a machine learning approach using SuperLearner algorithm. *J Anesth Analg Crit Care.* 2021; 1:3. doi:10.1186/s44158-021-00002-x.
[8]   Kuhn M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software.* 2008; 28(5):1–26. doi:10.18637/jss. v028.i05
[9]   R Core Team (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/. [Accessed 4th March 2023].
[10] Sakagianni A, Feretzakis G, Kalles D, Koufopoulou C, Kaldis V. Setting up an Easy-to-Use Machine Learning Pipeline for Medical Decision Support: A Case Study for COVID-19 Diagnosis Based on Deep Learning with CT scans. *Stud Health Technol Inform.* 2020; 272:13-16. doi: 10.3233/SHTI200481.
[11] Feretzakis G, Loupelis E, Sakagianni A, Kalles D, Martsoukou M, Lada M, et al. Using Machine Learning Techniques to Aid Empirical Antibiotic Therapy Decisions in the Intensive Care Unit of a General Hospital in Greece. *Antibiotics.* 2020;9(2):50. doi: 10.3390/antibiotics9020050.
[12] Afrash MR, Shanbehzadeh M, Kazemi-Arpanahi H. Predicting Risk of Mortality in COVID-19 Hospitalized Patients using Hybrid Machine Learning Algorithms. *J Biomed Phys Eng.* 2022;12(6):611-626. doi: 10.31661/jbpe.v0i0.2105-1334.
[13] Choi MH, Kim D, Choi EJ, Jung YJ, Choi YJ, Cho JH, et al. Mortality prediction of patients in intensive care units using machine learning algorithms based on electronic health records. *Sci Rep.* 2022;12(1):7180. doi: 10.1038/s41598-022-11226-4.