

Can artificial intelligence predict COVID-19 mortality?

A.C. GENÇ¹, D. CEKIC¹, K. ISSEVER¹, F.T. GENÇ¹, A.B. GENÇ¹, A. TOÇOĞLU¹, Y. DURMAZ², H. ÖZKÖK², A. NALBANT¹, S. YAYLACI¹

¹Department of Internal Medicine, Faculty of Medicine, Sakarya University, Sakarya, Turkey

²Department of Intensive Care Unit, Faculty of Medicine, Sakarya University, Sakarya, Turkey

Abstract. – OBJECTIVE: COVID-19 infection rapidly spread across the globe and evolved into a pandemic. Artificial intelligence (AI) has been used to predict the spread of the virus and diagnose and treat the disease. In this study, we aimed to predict whether patients admitted to the intensive care unit (ICU) due to COVID-19 infection will result in mortality.

PATIENTS AND METHODS: Ninety parameters were used for each 589 ICU patient. The nine parameters with the highest effect on mortality were determined. Four hundred seventy-one patients were used to train the AI with these nine parameters. AI was tested with 118 patient data.

RESULTS: AI estimated mortality with 83% sensitivity, 84% specificity, 84% accuracy, and 0.81 F1 score. Therefore, the area under the curve was calculated as 0.91. The results indicate that mortality among COVID-19 patients admitted to the ICU can be predicted based on their laboratory parameters on the first day.

CONCLUSIONS: These findings underscore the potential benefits of utilizing AI in managing pandemics.

Key Words:

Artificial intelligence, COVID-19, mortality, Intensive care unit.

Introduction

COVID-19 infection, initially detected in China, towards the end of 2019, rapidly spread across the globe and evolved into a pandemic¹. Droplets transmit this RNA virus from person to person, causing respiratory disease. This new virus has spread worldwide and turned into a pandemic. Reports indicate that the COVID-19 infection has reached many countries, surpassing 210 nations on a global scale².

Artificial intelligence (AI) technology has been used in almost all areas of pandemic management, such as the diagnosis, severity, course, treatment,

drug development when quarantine should be applied to whom, which environmental factors are influential in the spread of the virus, and early prediction of the number of cases³⁻⁵. To forecast daily trends in COVID-19 cases, a machine learning model was employed. This model utilized a dataset comprising the 30-day minimum, maximum, and average air temperature, humidity levels, wind speed, city population density, and the number of COVID-19 cases in Wuhan province, China⁶.

The timely implementation of quarantine measures for individuals who have been in contact with confirmed or suspected COVID-19 cases, including in larger communities or settlements, could effectively mitigate the spread of the virus. For this reason, it was suggested by a group of researchers⁷ to collect travel information, rather than symptoms, with a mobile phone-based application. Furthermore, it was argued that the spread of the virus would be reduced by evaluating the risks of COVID-19 for those using the application and suggesting the quarantine of the machine learning model to people at high risk. Researchers⁸ aimed to predict the mortality probability of COVID-19-diagnosed patients, even if they were not in the intensive care unit (ICU). They employed 7 variables and successfully predicted mortality with an approximate accuracy of 85%.

Ozturk et al⁹ developed a machine-learning model that predicts whether a chest X-ray indicates COVID-19 pneumonia. 4,352 computerized thorax tomographies were used to detect COVID-19 pneumonia with an AI model, which was successful over 90%¹⁰.

AI was able to determine the risk factors of diseases, when they will occur, which pathophysiological pathway was used, and the mortality¹¹.

Our aim was to assess the likelihood of mortality by analyzing age, gender, and laboratory findings obtained within the first day of intensive care unit admission. In this way, if it is possible to

predict the patients who will result in mortality, we can reduce the mortality rate with closer observation and more rigorous treatment methods.

Patients and Methods

Data Collection

The study included a cohort of 589 patients who were admitted to the ICU and tested positive for COVID-19 using the reverse transcription polymerase chain reaction (RT-PCR) method. The study dataset consisted of various variables, including age (in years), gender, presence of chronic diseases, regular drug usage, reason for hospital admission, microbiology findings, and biochemistry laboratory parameters recorded at the time of ICU admission.

RT-PCR Method

COVID-19 infection was detected by analyzing the nasopharyngeal swab sample with the RT-PCR method.

Artificial Intelligence Model

The machine learning model was developed based on the Categorical Boosting (CatBoost) algorithm to predict COVID-19 mortality. Gradient boosting methods have demonstrated remarkable effectiveness when applied to tasks involving tabular data, making them widely utilized in machine learning to achieve state-of-the-art results. The CatBoost algorithm is specifically robust in its processing of categorical features. This situation makes it especially aligned with the used dataset. The model is built to perform a binary classification of the output class, the patient's future health status, encoded as "0" for recovery and "1" for death.

Various implementations of gradient boosting algorithms were compared towards the selection of the model. In order to assess the average performance of the final model, a five-fold training was conducted using Extreme Gradient Boosting, CatBoost, and Light Gradient Boosting Machine. The performances of these models were comparable, with CatBoost demonstrating the best results among them.

The Train and Test Sets

Train and test sub-samples are randomly generated through a split dataset of 80 and 20 percent of the patients. The CatBoost model was trained on the training dataset to predict the output class

from the inputted data. In order to perform feature selection to determine the feature set of the final trained model, the model is first trained with the complete feature set of 90 features. Then, feature importance values, calculated by SHAP, are used to narrow down the final feature set of 9 features. SHapley Additive exPlanations (SHAP) is a framework for interpreting predictions that provides feature importance values based on the impact of each feature on the model.

Statistical Analysis

Descriptive data analyses were conducted to provide comprehensive information about the study population. Analytical and visual techniques were employed to assess the distribution pattern. In order to assess the precision of the AI model in predicting mortality, a receiver operating characteristic (ROC) curve analysis was conducted. This analysis determined the area under the curve (AUC) and provided valuable information on the model's predictive performance. Sensitivity and specificity percentages were also calculated and reported to assess the model's ability to identify positive and negative cases correctly. Additionally, positive and negative predictive values were calculated, and the findings were presented using a confusion mortality matrix, which further illustrated the model's predictive capabilities.

Results

The study included 589 patients who received treatment in the ICU. Among these patients, 294 experienced mortality, while 295 were discharged from the ICU. The average age of the patient population was 65.6 ± 15 . Among the patients, 358 were male, and 231 were female. Out of the 589 patients, 118 were used as the test dataset.

Out of the 90 variables in the dataset, we have identified the nine variables that exhibit the strongest association with mortality. The variables found to be the most successful in predicting mortality were estimated glomerular filtration rate (eGFR), age (year), lactate dehydrogenase U/L (LDH), procalcitonin ng/ml (PCT), high sensitive troponin I ng/L (hs troponin I), C-reactive protein mg/L (CRP), serum calcium mg/dL (Ca), lymphocyte K/uL, urea mg/dL (Figures 1,2).

The included parameters of the patients constituting the test set have the following mean and

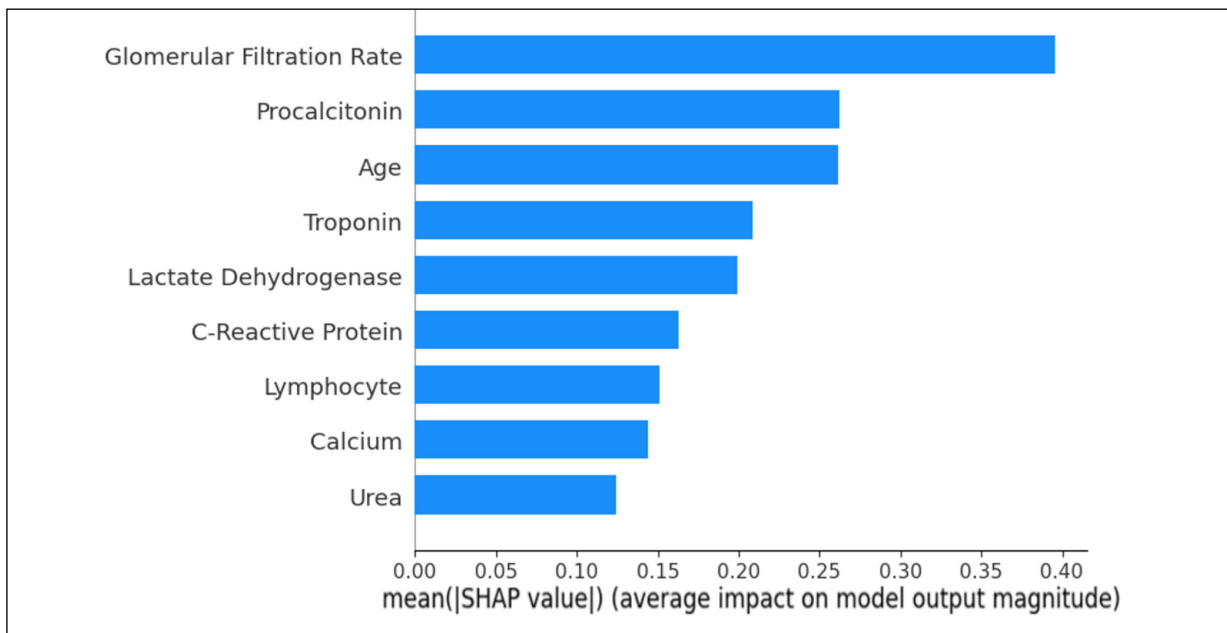


Figure 1. The SHAP feature importance chart displays the average absolute SHAP value for each feature.

standard deviation values in our study: Age 65.87 ± 14.12 years, eGFR 73.36 ± 29.88 , urea 70.31 ± 52.31 mg/dL, LDH 419.42 ± 199.50 U/L, PCT

2.39 ± 8.61 ng/ml, hs troponin I $311.29 \pm 1,473.58$ ng/L, CRP 103.76 ± 88.41 mg/L, Ca 8.29 ± 0.73 mg/dL, and lymphocyte count 0.79 ± 0.69 U/L.

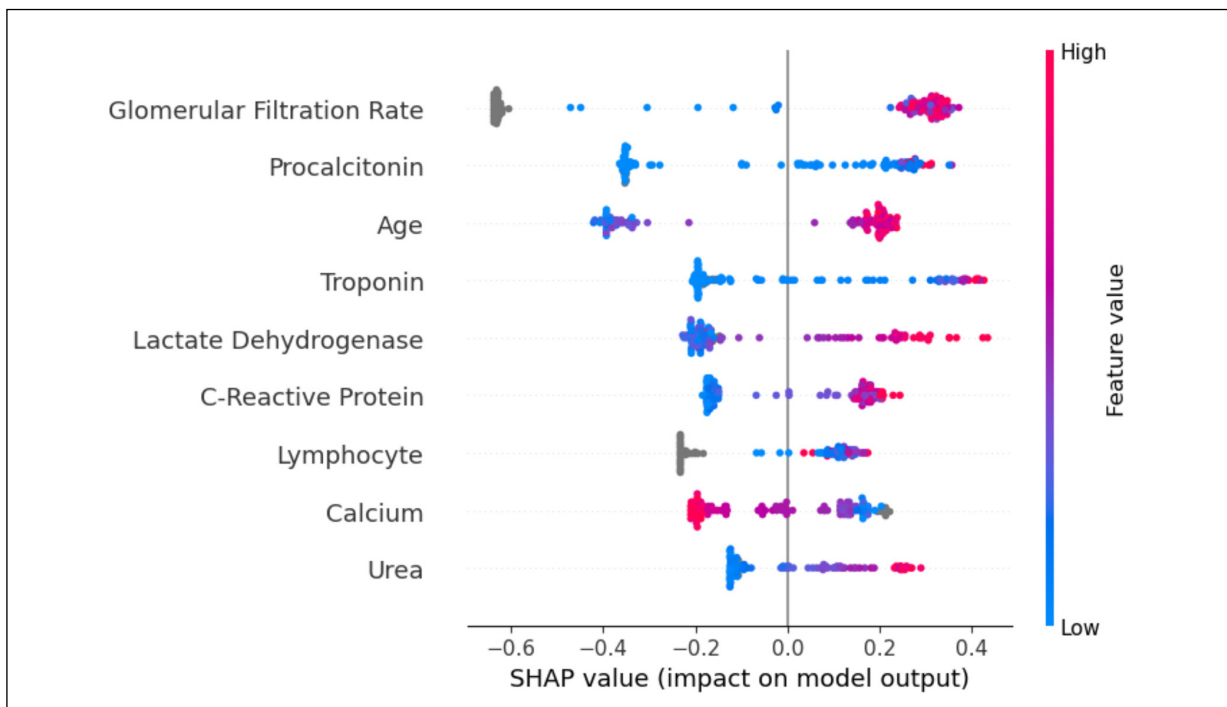


Figure 2. SHAP summary plot for a sample of top features. The x-axis indicates the impact on the model (no change to change from left to right). The color is the feature value.

Table I. Model performance of the final CatBoost model trained on the top 9 features.

CatBoost	AUC	Specificity (%)	Sensitivity (%)	Accuracy (%)	F1 score	Precision	Recall
Test performance of mortality	0.91	83	84	84	0.81	0.83	0.84

AUC: Area under the curve.

In the ROC analysis of the AI model tested with 118 patients, the AUC was determined to be 0.91. The model’s specificity was 83%, sensitivity 84%, accuracy 84%, F1 score 0.81, precision 0.83, and recall 0.84 (Table I, Figure 3).

Discussion

Numerous medical scores have been devised to estimate the probability of patients developing severe COVID-19 infection. The purpose of these scores is to evaluate a range of factors and clinical parameters in order to assess the severity of the disease and predict potential outcomes for patients¹². For numerous years, the Acute Physiology and Chronic Health Evaluation II (APACHE II) score has been widely used to calculate the mortality risk of patients admitted to the ICU. This scoring system offers a standardized approach to evaluate the severity of illness and predict the likelihood of mortality in critically ill patients¹³. While the APACHE II score is not specifically designed for evaluating COVID-19 patients, it has been widely used and proven to

provide a reliable estimation of ICU mortality in various critical illness scenarios. Therefore, although not specific to COVID-19, the APACHE II score can still offer valuable insights into the prognosis and potential outcomes of patients requiring intensive care management. The clinical examination findings of the physician are also required in calculating the APACHE II score. In addition to clinical evaluation data, measurable laboratory parameters are required¹⁴. In a study¹⁵ of 6,995 patients with COVID-19, 161 had a severe course. In predicting disease severity, a comparison was made between the APACHE II score and artificial intelligence (AI) techniques. The results indicated that AI outperformed the APACHE II score in accurately predicting the severity of critically ill patients.

The rapid escalation of COVID-19 cases, leading to a global pandemic, has necessitated the healthcare system to enhance its capacity to address the escalating demands. Health professionals have been compelled to identify cases that may progress to a severe clinical condition and augment their clinical skills accordingly. An AI model was developed with the aim of predicting the progression of acute respiratory distress syndrome (ARDS) in patients who initially presented with a mild clinical course upon their admission to the hospital. This model was trained using data from COVID-19 patients who initially presented with mild symptoms but later progressed to ARDS. Its purpose is to identify patients who are at risk of mortality and worsening of the disease. The presence of myalgia, slightly elevated alanine aminotransferase, and high hemoglobin levels were the most critical parameters predicting the development of ARDS. The AI model learned from the past clinical and laboratory information of patients with ARDS and was 70-80% successful in predicting severe cases¹⁶.

Clinical assessment data can be subjective. However, this data cannot be transferred to electronic media. For this reason, it will be more objective to predict mortality with the data of our study. Therefore, the AI model in this study was as successful as medical scores, without the

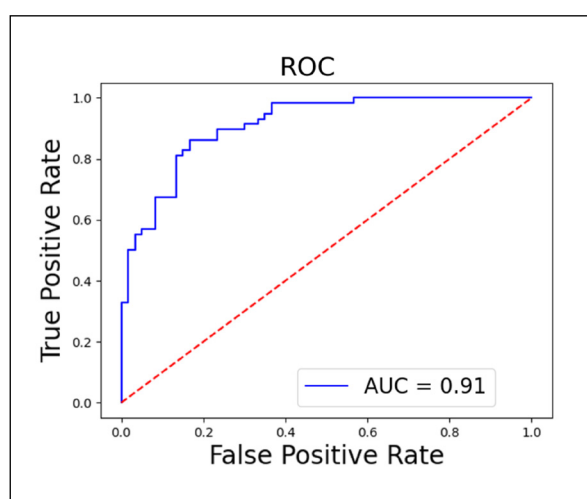


Figure 3. Receiver Operating Characteristic (ROC) curve for prediction of mortality of COVID-19 infection in intensive care unit admission.

need for personal information with age and eight laboratory data.

Limitations

There is no APACHE II score in this study's data, but many parameters are used in this score. The observational parameters in the APACHE II score are not available in our dataset. The retrospective design of the study introduces inherent constraints and limitations, which should be acknowledged as another notable limitation.

Conclusions

Newly identified viruses can quickly infect humans and become a global health problem. Unfortunately, the clinical knowledge and skills of healthcare professionals may make it difficult for them to overcome a newly identified disease. AI models can positively contribute to human life by helping to overcome this problem.

Conflict of Interest

The authors declare that they have no conflict of interests.

Acknowledgements

We respectfully commemorate all our patients and healthcare workers affected by COVID-19. We want to express our gratitude to Zeyneb Nahide Kaya for her valuable contributions to the field of artificial intelligence during the creation of this article.

Authors' Contribution

Conceptualization: ACG, SY, YD, AN, ABG; Literature Search: ACG, DC, KI, FTG, AT; Methodology: ACG, ABG, AT, YD; Data: ACG, FTG, YD, HO, DC, ABG; Formal analysis and investigation: ACG, DC, KI, F; Writing - review and editing: ACG, AN, SY, AT, YD, HO; Supervision: all authors read and approved the final manuscript.

Ethics Approval

The research study strictly followed the ethical guidelines stipulated in the Declaration of Helsinki. Prior to its initiation, the study obtained the necessary ethical approval from the Ethics Committee of Sakarya University Faculty of Medicine (Data: 12.05.2022, Protocol No.: E-71522473-050.01.04-194687-333).

Informed Consent

Informed consent was waived due to the retrospective nature of the study.

Funding

This research received no specific grant from any funding agency.

ORCID ID

Ahmed Cihad Genc: 0000-0002-7725-707X
Deniz Cekic: 0000-0002-7114-9334
Kubilay Issever: 0000-0002-1376-1488
Fevziye Turkoglu Genc: 0000-0002-1607-6355
Ahmed Bilal Genc: 0000-0002-1607-6355
Aysel Toçoğlu: 0000-0002-9078-6996
Yusuf Durmaz: 0009-0001-9320-8411
Hüseyin Özkök: 0000-0001-9948-4233
Ahmet Nalbant: 0000-0003-4756-3575
Selcuk Yaylaci: 0000-0002-6768-7973.

References

- 1) Singer DR. A new pandemic out of China: the Wuhan 2019-nCoV coronavirus syndrome. *Health Policy Technol* 2020; 9: 1-2.
- 2) Sharma A, Ahmad Farouk I, Lal SK. COVID-19: a review on the novel coronavirus disease evolution, transmission, detection, control and prevention. *Viruses* 2021; 13: 202.
- 3) Liu F, Zhang Q, Huang C, Shi C, Wang L, Shi N, Fang C, Shan F, Mei X, Shi J, Song F, Yang Z, Ding Z, Su X, Lu H, Zhu T, Zhang Z, Shi L, Shi Y. CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients. *Theranostics* 2020; 10: 5613-5622.
- 4) Mei X, Lee HC, Diao K, Huang M, Lin B, Liu C, Xie Z, Ma Y, Robson PM, Chung M, Bernheim A, Marni V, Calcagno C, Li K, Li S, Shan H, Lv J, Zhao T, Xia J, Long Q, Steinberger S, Jacobi A, Deyer T, Luksza M, Liu F, Little BP, Fayad ZA, Yang Y. Artificial intelligence-enabled rapid diagnosis of COVID-19 patients. *Nat Med* 2020; 26: 1224-1228.
- 5) Stebbing J, Krishnan V, de Bono S, Ottaviani S, Casalini G, Richardson PJ, Monteil V, Lauschke VM, Mirazimi A, Youhanna S, Tan YJ, Baldanti F, Sarasini A, Terres JAR, Nickoloff BJ, Higgs RE, Rocha G, Byers NL, Schlichting DE, Nirula A, Cardoso A, Corbellino M; Sacco Baricitinib Study Group. Mechanism of baricitinib supports artificial intelligence-predicted testing in COVID-19 patients. *EMBO Mol Med* 2020; 12: e12697.
- 6) Pirouz B, Shaffiee Haghshenas S, Shaffiee Haghshenas S, Piro P. Investigating a serious challenge in the sustainable development process: analysis of confirmed cases of COVID-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. *Sustainability* 2020; 12: 2427.
- 7) Srinivasa Rao AS, Vazquez JA. Identification of COVID-19 can be quicker through artificial intel-

- ligence framework using a mobile phone-based survey when cities and towns are under quarantine. *Infect Control Hosp Epidemiol* 2020; 41: 826-830.
- 8) Al-Najjar H, Al-Rousan N. A classifier prediction model to predict the status of Coronavirus COVID-19 patients in South Korea. *Eur Rev Med Pharmacol Sci* 2020; 24: 3400-3403.
 - 9) Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 2020; 121: 103792.
 - 10) Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology* 2020; 296: E65-E71.
 - 11) Gautier T, Ziegler LB, Gerber MS, Campos-Náñez E, Patek SD. Artificial intelligence and diabetes technology: A review. *Metabolism* 2021; 124: 154872.
 - 12) Sungono V, Hariyanto H, Soesilo TEB, Adisasmita AC, Syarif S, Lukito AA, Widysanto A, Puspitasari V, Tampubolon OE, Sutrisna B, Sudaryo MK. Cohort study of the APACHE II score and mortality for different types of intensive care unit patients. *Postgrad Med J* 2022; 98: 914-918.
 - 13) Godinjak A, Iglica A, Rama A, Tančica I, Jusufović S, Ajanović A, Kukuljac A. Predictive value of SAPS II and APACHE II scoring systems for patient outcome in medical intensive care unit. *Acta Med Acad* 2016; 45: 97-103.
 - 14) Beigmohammadi MT, Amoozadeh L, Rezaei Motlagh F, Rahimi M, Maghsoudloo M, Jafarnejad B, Eslami B, Salehi MR, Zendehtdel K. Mortality Predictive Value of APACHE II and SOFA Scores in COVID-19 Patients in the Intensive Care Unit. *Can Respir J* 2022; 2022: 5129314.
 - 15) Assaf D, Gutman Y 'ara, Neuman Y, Segal G, Amit S, Gefen-Halevi S, Shilo N, Epstein A, Mor-Cohen R, Biber A, Rahav G, Levy I, Tirosh A . Utilization of machine-learning models to accurately predict the risk for critical COVID-19. *Intern Emerg Med* 2020; 15: 1435-1443.
 - 16) Jiang X, Coffee M, Bari A, Wang J, Jiang X, Huang J, Shi J, Dai J, Cai J, Zhang T, Wu Z, He G, Huang Y. Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. *Comput Mater Contin* 2020; 62: 537-551.