Received: 3 June 2023 | Revised: 7 February 2024 | Accepted: 18 April 2024

DOI: 10.1002/hsr2.2104

ORIGINAL RESEARCH

Health Science Reports Open Access WILEY

Application of boosted trees to the prognosis prediction of COVID-19

Sajjad Molaei[1] | Hadi Moazen[2] | Hamid R. Niazkar[3] | Masoud Sabaei[1] | Masoumeh G. Johari[3] | Abbas Rezaianzadeh[4]

[1]Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran

[2]Department of Computer Science and Software Engineering, Universite Laval, Quebec, Quebec, Canada

[3]Breast Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

[4]Colorectal Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

Correspondence

Hamid R. Niazkar, Breast Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.
Email: hrn185@hotmail.com

Abstract

Background and Aims: The precise prediction of COVID-19 prognosis remains a clinical challenge. In this regard, early identification of severe cases facilitates the triage and management of COVID-19 cases. The present paper aims to explore the prognosis of COVID-19 patients based on routine laboratory tests taken when patients are admitted.

Methods: A data set including 1455 COVID-19 patients (727 male, 728 female) and their routine laboratory tests conducted upon hospital admission, age, Intensive Care Unit (ICU) admission, and outcome were gathered. The data set was randomly split into the train (75% of the data) and test data set (25% of the data). The explainable boosting machine (EBM) and extreme gradient boosting (XGBoost) were used for predicting the mortality and ICU admission of COVID-19 cases. Also, feature importance was extracted using EBM and XGBoost.

Results: The EBM and XGBoost achieved 86.38% and 88.56% accuracy in the test data set, respectively. In addition, EBM and XGBoost predicted the ICU admission with an accuracy of 89.37%, and 79.29% in the test data set for COVID-19 patients, respectively. Also, obtained models indicated that aspartate transaminase (AST), lymphocyte, blood urea nitrogen (BUN), and age are the most significant predictors of COVID-19 mortality. Furthermore, the lymphocyte count, AST, and BUN level were the most significant ICU admission predictors of COVID-19 patients.

Conclusions: The current study indicated that both EBM and XGBoost could predict the ICU admission and mortality of COVID-19 cases based on routine hematological and clinical chemistry evaluation at the time of admission. Also, based on the results, AST, lymphocyte count, and BUN levels could be used as early predictors of COVID-19 prognosis.

KEYWORDS

COVID-19, explainable boosting machine, machine learning, prediction, prognosis, XGBoost

1 | INTRODUCTION

The coronavirus disease 2019 (COVID-19) has been a worldwide health concern in recent years. According to the statistics, as of September 2022, more than 600 million individuals have been infected with COVID-19, leading to more than 6 million deaths.[1] Most COVID-19 patients are asymptomatic or may experience a mild disease, while some require hospital care and intensive care unit admission.[2] Accordingly, several studies have been carried out to discover the patients with the highest risk of experiencing severe COVID-19 symptoms.[3,4] Predicting the outcome of those with COVID-19 helps physicians to provide timely management for those at risk of developing COVID-19 complications. Also, it eases the triage of infected individuals by allocating limited medical resources such as mechanical ventilation units and hospital beds.[5,6] In a bid to estimate the prognosis of COVID-19 cases, various artificial intelligence models have been applied. However, despite various studies on the prognosis prediction of COVID-19, the quest for identifying the severity indicators of COVID-19 is still ongoing.[7,8]

Various studies have applied machine learning methods to estimate COVID-19 prognosis.[9] For instance, An et al.[10] developed various machine learning models such as least absolute shrinkage and selection operator, random forest (RF), and linear support vector machine (SVM) to estimate outcome of COVID-19 cases considering the sociodemographic and past medical history (PMH).[10] They indicated that age and PMH, such as diabetes, cancer, and hypertension, are among the main predictors of COVID-19 prognosis.[10] Similarly, the outcome prediction of COVID-19 cases was carried out by Kocadaglia et al.[11] by applying hybrid machine learning methods using the symptoms and laboratory tests as predictors.[11] They demonstrated that age, dyspnea, Spo2, respiratory rate, neutrophil count, and c-reactive protein (CRP) have the highest importance in the prognosis prediction of COVID-19.[11] Moreover, Thimoteo et al.[12] applied logistic regression, explainable boosting machine (EBM), RF, and SVM for COVID-19 identification using laboratory tests and pathogen variables.[12] They indicated that the RF and EBM achieved the best results. Also, they observed that leukocytes were the most significant factor in diagnosing COVID-19.[12]

Furthermore, the current work tries to investigate COVID-19 prognosis prediction using EBM and extreme gradient boosting (XGBoost) based on routine laboratory tests conducted upon hospital admission. In addition, since the selected methods can predict and identify the significance of selected futures, variables with the highest significance for the prediction of COVID-19 prognosis will be identified.[13] The discriminative power of EBM, along with its inherent explainability, makes it an optimal choice for healthcare applications, and so it was selected for the current study. Also, XGBoost is known for providing an efficient variant gradient boosting algorithm capable of achieving satisfactory results for prediction purposes.[14]

2 | MATERIALS AND METHODS

2.1 | Study patients

The current retrospective study was carried out on COVID-19 patients hospitalized in Allameh Behlool Gonabadi during 2020–2021, after approval of the medical ethics committee of the Gonabad University of Medical Sciences (ethic code: IR.GMU.REC.1400.060).

2.2 | The inclusion and exclusion criteria

The inclusion criteria were having a confirmed nasopharyngeal RT-PCR test for COVID-19 and being admitted to either the infectious disease or internal medicine wards. On the other hand, the pregnant women, infants, neonates, and children with COVID-19, as well as those who left against medical advice, were transferred or referred to other hospitals, or had missing data were excluded.

2.3 | Data extraction

To gather the required data, a list of COVID-19 cases that were admitted to the hospital was compiled using the information provided to the hospital's registration system. This information included the patients' age, sex, and outcome (whether they were hospitalized in the ICU, died, or recovered). From this list, a total of 2660 cases that met the inclusion criteria were selected for the experiments. After excluding pregnant women (45 cases), infants (27 cases), neonates (two cases), and children with COVID-19 (212 cases), as well as those who left against medical advice (24 cases) or were transferred or referred to other hospitals (eight cases), a final count of 2342 patients remained. Various laboratory tests conducted upon admission, such as a complete blood count, coagulation indices such as partial thromboplastin time, prothrombin time (PT), inflammatory markers such as CRP, erythrocyte sedimentation rate (ESR), and biochemical factors including creatinine, blood urea nitrogen (BUN), aspartate aminotransferase (AST), alanine aminotransferase (ALT), and alkaline phosphatase (ALP), were retrieved from the hospital's electronic registration system using the national code of each patient. These results were then recorded in a checklist. After excluding patients with missing data (887 cases), a total of 1455 patients (727 male, 728 female) were included in the study.

2.4 | Data randomization

The data were randomly categorized into the train (75% of the data) and the test data set (25% of the data). The random selection method used in this study consists of four steps. In the first step, the data were anonymized by assigning an arbitrary ID number to each data point. In the second step, the data points were randomly permutated.

In the third step, data points corresponding to the first 25% of the permutated set were extracted and called test data.[15] In the fourth step, the range of each variable in the test data was compared with the range of the corresponding variable in the training data. If the range of all variables in the training data embraces the corresponding variables in the test data, the data are properly categorized. If not, the second to fourth steps should be repeated so that each parameter in the training data has a wider range than the corresponding variables in the test data.

2.5 | Brief introduction to EBM

The EBM was presented as part of InterpretML by Microsoft,[16] an interpretable method that can match the performance of more complex models such as XGBoost, is a generalized additive model (GAM) as presented in the below equation.

$$g(\mathrm{E}[y]) = \beta_0 + \sum f_j(x_j), \tag{1}$$

where x_j are the features in our data set and $f_j(.)$ are known as the link shape functions. The interpretability and simplicity of EBM, and GAMs in general, stems from the fact that the shape functions operate on single features, and any interaction terms between different features are ignored. This enables us to visualize and understand the effect each feature has on a decision made by the model on a single data point and can be further extended to explore the behavior of the model over the entire input space.

EBM can achieve this level of performance and interpretability by utilizing boosting. The backbone of EBM is a shallow decision tree with a single input feature. When training an EBM model, all features are used iteratively, and a decision tree is fitted using a single feature to the current residuals. The final model is the combination of all learned decision trees. The interpretability of EBM makes it a great choice for medical applications of AI as it allows the practitioners to understand the behavior of the model and builds trust in the usage of the model.

2.6 | Brief introduction to XGBoost

XGBoost is an implementation of the gradient boosting framework using decision trees. In gradient boosting, instead of optimizing the model parameters, new weak models are fitted to the residuals and combined with the previous weak models to achieve learning at each training iteration.

An XGBoost model is an ensemble of weak learners implemented using decision trees, and the model combines the outputs of these trees to generate the final output. Unlike EBM, XGBoost uses trees with more than a single input feature, which makes it more complex and much less interpretable but contributes to its performance. An XGBoost model composed of K decision trees is defined as in the below equation.

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \tag{2}$$

where $\hat{y}_i$ is the output of the XGBoost model, $f_k(.)$ are the decision trees, and x_i is an input sample. Please note the difference between the formulation of XGBoost and EBM that where in EBM each $f_j(.)$ corresponded to a single feature j of the input, decision trees in XGBoost take the entire samples as input.

3 | RESULTS

We take advantage of EBM and XGBoost for both tasks, prediction of death or recovery and prediction of ICU admission. For each task, both models were trained on the same training set, randomly selected from the collected data set. The input features of each model were the routine laboratory tests conducted upon hospital admission. The features were normalized before being fed into the models and the same normalization was used for training and test sets. One model was trained per task, and hence a total of four models were trained. For each model, a randomized grid search was performed to obtain the model parameters that resulted in the best performance on the training. The following sections discuss the performance of the models in more detail.

3.1 | Prediction of death or recovery

For the prediction of death or recovery, EBM achieved an area under the curve (AUC) of 0.9357 on the training data set while achieving an AUC of 0.8833 on the test data set. In addition, for forecasting death or recovery, XGBoost prediction models obtained an AUC of 0.9856 on the training data set and an AUC of 0.8760 on the test data set. Also, the proposed EBM model exhibited an accuracy of 87.65% on the training data set and an accuracy of 86.38% on the test set. In addition, the XGBoost prediction model achieved an accuracy of 94.28% on the training data set and an accuracy of 88.56% on the test data set. Figure 1 presents the receiver operating characteristic (ROC) curve and confusion matrices of the obtained models on the training and test datasets.

EBM exhibited a higher AUC in predicting COVID-19 mortality, while the XGBoost model outperformed the EBM in the accurate prediction of patients' outcomes.

The feature importance of input variables is presented in Figure S1. It should be noted that a higher F score indicates a higher significance in the prediction of the outcome. According to the EBM prediction model, among studied variables, the age, AST, BUN, lymphocyte count, and Cr features have the highest importance in the prediction of COVID-19 mortality. However, the XGBoost model indicated that the AST, lymphocyte count, RBC, BUN, and age have the highest significance in predicting COVID-19 mortality, respectively. Overall, AST, lymphocyte count, BUN, and age are the strongest predictors of COVID-19 mortality.

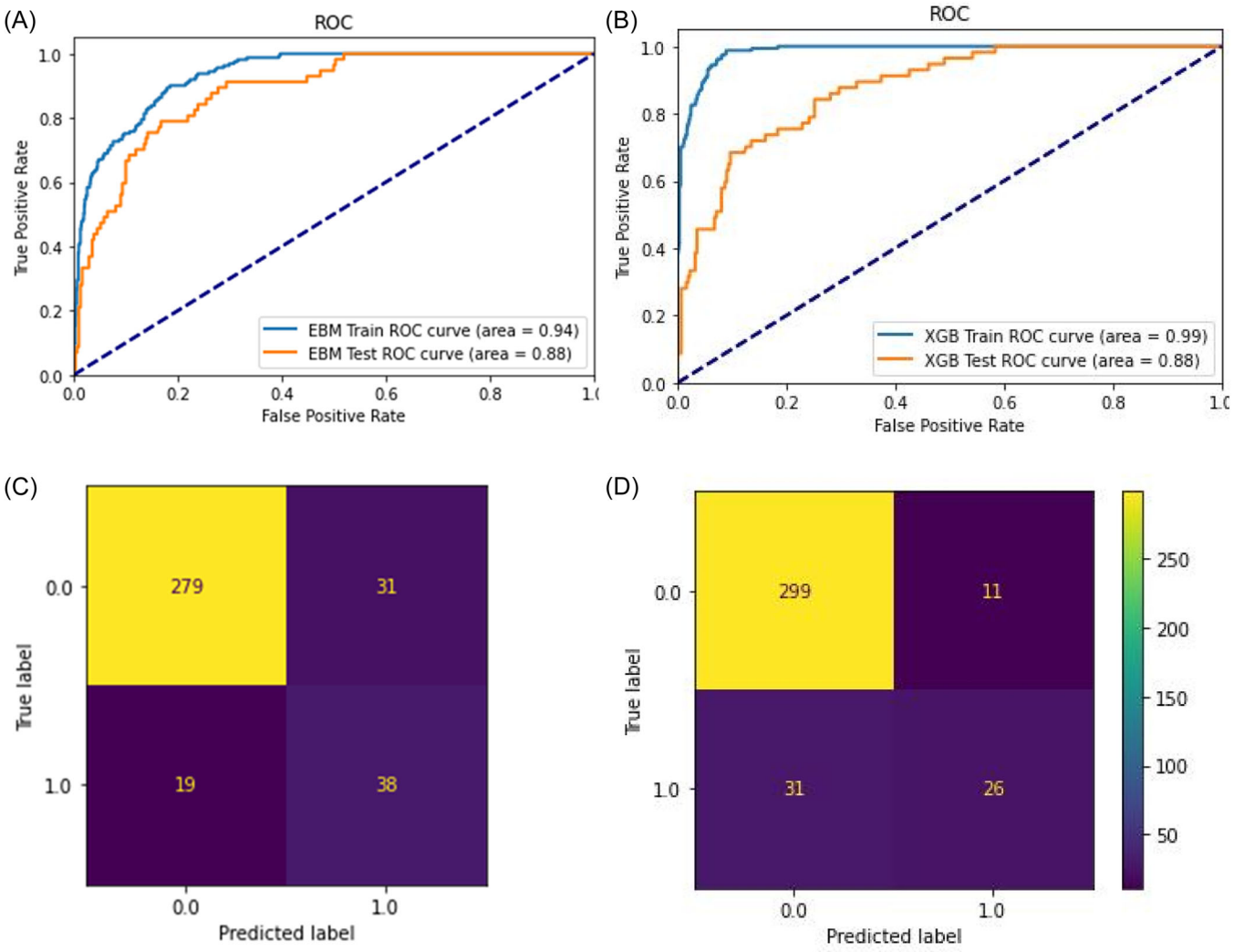

FIGURE 1 The ROC of mortality prediction models in the training and test datasets using explainable boosting machine (EBM) (A) and extreme gradient boosting (XGBoost) (B), the confusion matrix of COVID-19 mortality prediction models using EBM (C) and XGBoost (D). ROC, receiver operating characteristic.

In a bid to determine the critical threshold, the trend of contributing variables with the highest importance in the EBM prediction models was obtained and illustrated in Figure 2.

As presented in Figure 2, individuals older than 70 years old have a significantly higher probability of COVID-19 mortality. While for infected individuals younger than 60 years old, age may play a positive role in the COVID-19 mortality prediction. Furthermore, an AST level higher than 40 (U/L), a BUN level higher than 20 (mg/dL), Cr higher than 1.3 (mg/dL), and absolute lymphocyte count lower than 750 indicate a higher mortality rate in COVID-19 patients.

3.2 | Prediction of ICU admission

For ICU admission prediction, the EBM prediction model obtained an AUC of 0.9343 on the training set and an AUC of 0.8236 on the test set. Also, the XGBoost prediction model achieved an AUC of 0.9854 on the training set and an AUC of 0.8409 on the test data set. The EBM predicted the ICU admission with an accuracy of 95.00% and 89.37% on the training and test datasets, respectively. Also, XGBoost obtained an accuracy of 85.56% and 79.29% in the prediction of COVID-19 ICU admission. Figure 3 presents the ROC and confusion matrices of developed models for the training and test datasets.

XGBoost achieved a higher AUC than EBM in predicting ICU admission. Also, the EBM prediction model performed more accurately in estimating ICU admission.

In addition, the feature importance of the studied variables is depicted in Figure S2. According to Figure S2, the EBM estimation model indicated that the lymphocyte count, AST, BUN, PT, and age level are the most significant predictors of ICU admission in COVID-19. Also, in the XGBoost prediction model, lymphocyte count, BUN, AST, RBC, and neutrophil count were the most significant variables in predicting COVID-19 ICU admission. Overall, lymphocyte count, AST, and BUN

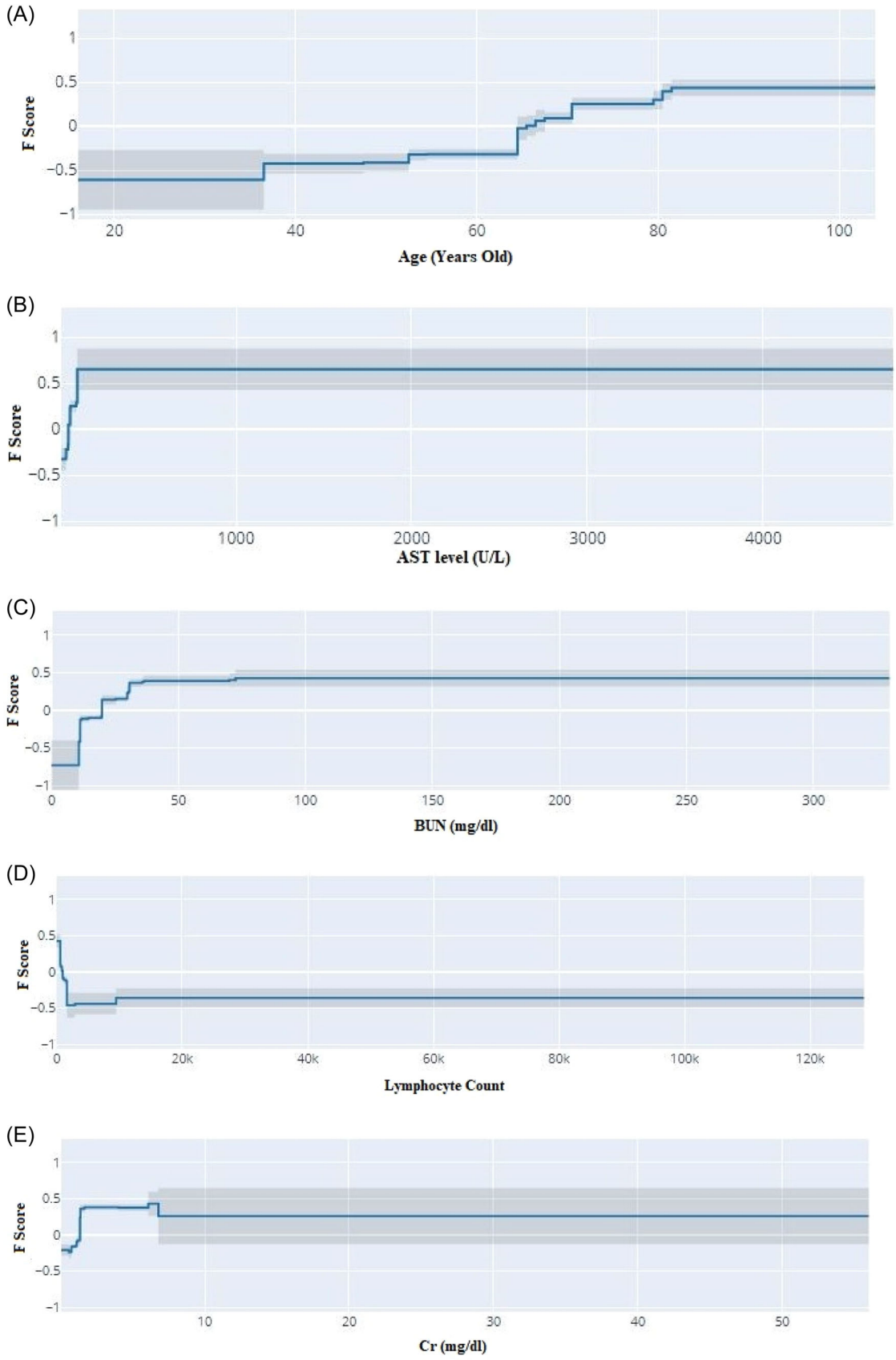

FIGURE 2 The trend of age (A), aspartate transaminase (AST) (B), blood urea nitrogen (BUN) (C), lymphocyte count (D), and Cr (E) in the prediction of COVID-19 mortality using explainable boosting machine (EBM).

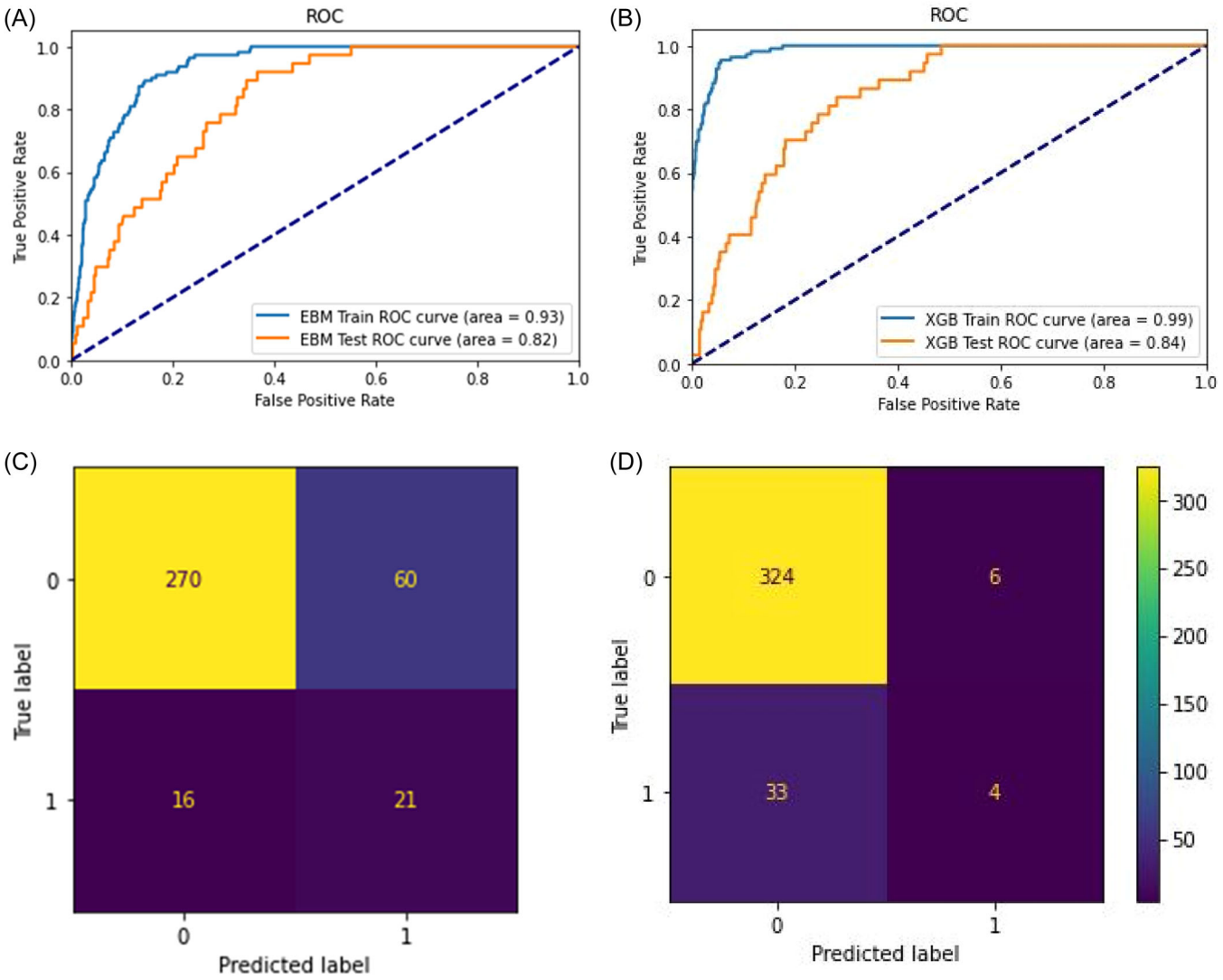

FIGURE 3 The ROC of intensive care unit (ICU) admission prediction models on the training and test datasets using explainable boosting machine (EBM) (A) and extreme gradient boosting (XGBoost) (B), the confusion matrices of COVID-19 ICU admission prediction models using EBM (C) and XGBoost (D). ROC, receiver operating characteristic.

level were common in both models, which may have the highest importance in predicting ICU admission.

Figure 4 presents the trend of contributing variables with the highest importance in the EBM prediction models.

According to Figure 4, lymphopenia (lower than 815), AST higher than 130 (U/L), BUN higher than 30 (mg/dL), PT higher than 15 (seconds), and individuals older than 70 years old who contract COVID-19 have a higher probability of being admitted to ICU.

4 | DISCUSSION

The prompt spread of COVID-19 brought about a worldwide healthcare challenge. One of the major concerns during the pandemic is the screening of infected patients.[17,18] This helps decision-makers allocate limited medical facilities such as hospital beds and mechanical ventilators. Also, identifying those susceptible to severe COVID-19 symptoms aids physicians in providing timely necessary management.[19] For this purpose, the current study investigated the prognosis of COVID-19 patients using EBM and XGBoost. In the current study, an attempt was made to construct a triage framework by exclusively utilizing the common hematological and clinical chemistry tests conducted upon hospital admission. Consequently, the proposed prognostic models can be effectively employed for triaging COVID-19 cases in healthcare facilities with limited resources. Undoubtedly, incorporating the clinical changes during the hospitalization along with more sophisticated laboratory examinations holds the potential to yield more precise prognostic models for predicting the outcome of COVID-19.

The present study revealed that both EBM and XGBoost models can predict the ICU admission and mortality of COVID-19 patients. According to the obtained results, XGBoost achieved higher accuracy

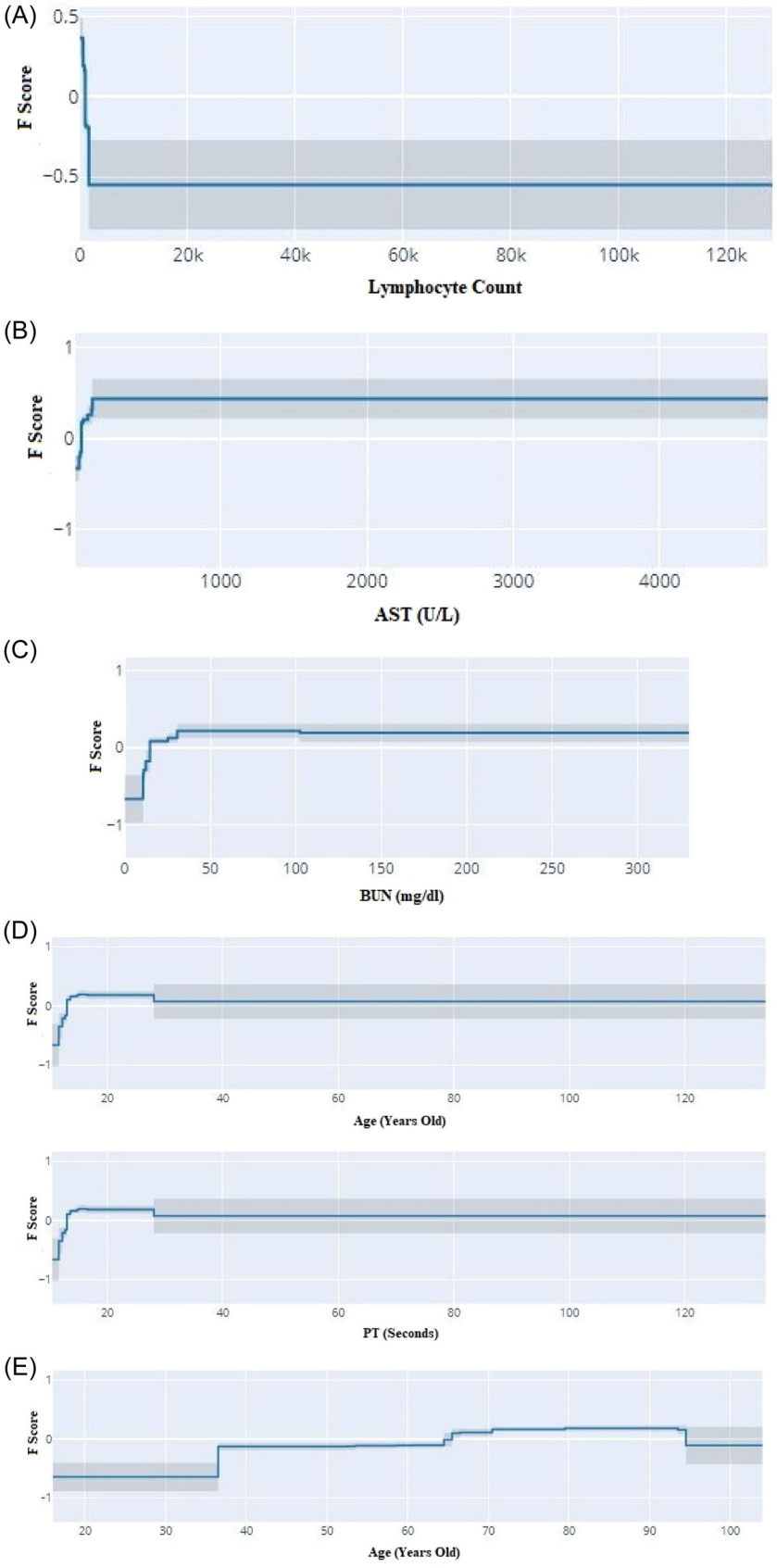

FIGURE 4 The trend lymphocyte count (A), aspartate transaminase (AST) (B), blood urea nitrogen (BUN) (C), prothrombin time (PT) (D), and age (E) in the prediction of COVID-19 intensive care unit (ICU) admission using explainable boosting machine (EBM).

in predicting mortality in COVID-19 cases. Conversely, the EBM model outperformed XGBoost in predicting ICU admission in COVID-19 cases. Also, feature importance analysis revealed that AST, lymphocyte count, BUN, and age are the most significant predictors of COVID-19 mortality. In addition, the lymphocyte count, AST, and BUN values had the most significant impact on the risk of ICU admission. Considering the importance of AST, lymphocyte count, and BUN in predicting mortality and ICU admission, it seems that COVID-19-induced or pre-existing liver and kidney injuries may contribute to a major part of COVID-19 mortality. Similar to the present study, previous clinical investigations demonstrated that elevated liver enzyme levels might be common during COVID-19.[20] Studies showed that COVID-19 cases with acute liver injury are susceptible to developing severe COVID-19.[21] Also, acute kidney injury was observed in one-third of COVID-19 patients, and it was reported to be an indicator of COVID-19 severity.[22,23] Additionally, similar to our findings, a meta-analysis by Huang and Pranata[24] indicated that lymphopenia at the time of admission was correlated with a poor prognosis in COVID-19 cases.[24]

Additionally, this work provided a threshold in the trends of significant variables in predicting COVID-19 prognosis. According to EBM's results, lymphocyte levels lower than 750, AST higher than 40 (U/L), and BUN higher than 20 (mg/dL) are strongly associated with COVID-19 mortality. Also, a lymphocyte level lower than 815, AST higher than 130 (U/L), and BUN higher than 30 (mg/dL) are the major predictors of ICU admission in COVID-19 patients. Needless to say, these thresholds may help healthcare providers to identify susceptible patients at risk of COVID-19 mortality.

The current work has several limitations. In the present study, only routine hematological and clinical chemistry tests at the time of admission and age were included as predictors. However, other variables, including symptoms, Spo2, past medical histories, and smoking habits, may unarguably contribute to COVID-19 mortality. In this regard, further studies investigating such variables are encouraged. Secondly, since most COVID-19 patients recover without ICU admission, its datasets are usually imbalanced, which may impact the obtained accuracy. In addition, the current study is a single-center study (Allameh Behlool Gonabadi Hospital), and it may be subjected to an institutional bias. Moreover, vaccination, possible mutations, and new variants of SARS-CoV-2 may bias the findings of the current study. Lastly, the present study excluded pregnant women, children, and neonates with COVID-19. On this subject, similar studies could be conducted in these excluded populations.

5 | CONCLUSION

The present study revealed that both EBM and XGBoost can predict the mortality and ICU admission of COVID-19 cases using routine hematological and clinical chemistry evaluation conducted upon hospital admission. Furthermore, the obtained models indicated that AST, lymphocyte count, BUN, and age have the highest impact in predicting COVID-19 mortality. In addition, the lymphocyte count, AST, and BUN values had the highest impact on the risk of ICU admission. Therefore, evaluating AST, lymphocyte count, and BUN levels is suggested for an early prognosis prediction of COVID-19 patients.

AUTHOR CONTRIBUTIONS

Sajjad Molaei: Conceptualization; formal analysis; methodology. **Hadi Moazen**: Formal analysis; methodology. **Hamid R. Niazkar**: Data curation; methodology; project administration; writing—original draft. **Masoud Sabaei**: Formal analysis; investigation. **Masoumeh G. Johari**: Conceptualization; methodology; writing—review and editing. **Abbas Rezaianzadeh**: Project administration; writing—review and editing.

ACKNOWLEDGMENTS

All authors have read and approved the final version of the manuscript. Hamid R. Niazkar had full access to all the data in this study and takes complete responsibility for the integrity of the data and the accuracy of the data analysis.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Due to ethical, and legal concerns, the supporting data could not be publicly available. However, the data that support the findings of this study are available from the corresponding author upon reasonable request.

TRANSPARENCY STATEMENT

The lead author Hamid R. Niazkar affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

ORCID

Sajjad Molaei http://orcid.org/0000-0001-5569-5178
Hadi Moazen http://orcid.org/0000-0003-3832-8856
Hamid R. Niazkar http://orcid.org/0000-0002-6143-9979
Masoud Sabaei http://orcid.org/0000-0002-3294-3113
Masoumeh G. Johari http://orcid.org/0000-0003-3486-7182
Abbas Rezaianzadeh http://orcid.org/0000-0002-0067-0659

REFERENCES

1. *WHO Coronavirus Disease (COVID-2019) Situation Reports*. 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/
2. Musazadeh V, Karimi A, Bagheri N, et al. The favorable impacts of silibinin polyphenols as adjunctive therapy in reducing the complications of COVID-19: a review of research evidence and underlying mechanisms. *Biomed Pharmacother Biomed Pharmacother*. 2022;154:113593.
3. Figliozzi S, Masci PG, Ahmadi N, et al. Predictors of adverse prognosis in COVID-19: a systematic review and meta-analysis. *Eur J Clin Invest*. 2020;50(10):e13362.
4. Sperrin M, Grant SW, Peek N. Prediction models for diagnosis and prognosis in COVID-19. *Br Med J Publish Group*. 2020;369:m1464.

5. Niazkar M, Niazkar HR. COVID-19 outbreak: application of multi-gene genetic programming to country-based prediction models. *Electr J General Med*. 2020;17(5):em247.
6. Feinstein MM, Niforatos JD, Hyun I, et al. Considerations for ventilator triage during the COVID-19 pandemic. *Lancet Respirat Med*. 2020;8:e53.
7. Karimi Shahri M, Niazkar HR, Rad F. COVID-19 and hematology findings based on the current evidences: a puzzle with many missing pieces. *Int J Lab Hematol*. 2021;43(2):160-168.
8. Tahtasakal CA, Oncul A, Sevgi DY, et al. Could we predict the prognosis of the COVID-19 disease? *J Med Virol*. 2021;93(4):2420-2430.
9. Jamshidi MB, Roshani S, Talla J, et al. A review of the potential of artificial intelligence approaches to forecasting COVID-19 spreading. *AI*. 2022;3(2):493-511.
10. An C, Lim H, Kim D-W, Chang JH, Choi YJ, Kim SW. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci Rep*. 2020;10(1):18716.
11. Kocadagli O, Baygul A, Gokmen N, Incir S, Aktan C. Clinical prognosis evaluation of COVID-19 patients: an interpretable hybrid machine learning approach. *Curr Res Transl Med*. 2022;70(1):103319.
12. Thimoteo LM, Vellasco MM, Amaral J, Figueiredo K, Yokoyama CL, Marques E. Explainable artificial intelligence for COVID-19 diagnosis through blood test variables. *J Control Automat Electr Syst*. 2022;33(2):625-644.
13. Sarica A, Quattrone A, Quattrone A, editors. Explainable boosting machine for predicting Alzheimer's Disease from MRI hippocampal subfields. International Conference on Brain Informatics; 2021: Springer.
14. Ogunleye A, Wang Q-G. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinf*. 2020;17(6):2131-2140.
15. Niazkar HR, Moshari J, Khajavi A, Ghorbani M, Niazkar M, Negari A. Application of multi-gene genetic programming to the prognosis prediction of COVID-19 using routine hematological variables. *Sci Rep*. 2024;14(1):2043.
16. Nori H, Jenkins S, Koch P, Caruana R. Interpretml: a unified framework for machine learning interpretability. arXiv preprint arXiv:190909223. 2019.
17. Palayew A, Norgaard O, Safreed-Harmon K, Andersen TH, Rasmussen LN, Lazarus JV. Pandemic publishing poses a new COVID-19 challenge. *Nat Hum Behav*. 2020;4(7):666-669.
18. Hautz WE, Exadaktylos A, Sauter TC. Online forward triage during the COVID-19 outbreak. *Emerg Med J*. 2021;38(2):106-108.
19. Liu Y, Wang Z, Ren J, et al. A COVID-19 risk assessment decision support system for general practitioners: design and development study. *J Med Internet Res*. 2020;22(6):e19786.
20. Wijarnpreecha K, Ungprasert P, Panjawatanan P, et al. COVID-19 and liver injury: a meta-analysis. *Eur J Gastroenterol Hepatol*. 2021;33(7):990-995.
21. Phipps MM, Barraza LH, LaSota ED, et al. Acute liver injury in COVID-19: prevalence and association with clinical outcomes in a large US cohort. *Hepatology*. 2020;72(3):807-817.
22. Yang X, Jin Y, Li R, Zhang Z, Sun R, Chen D. Prevalence and impact of acute renal impairment on COVID-19: a systematic review and meta-analysis. *Crit Care*. 2020;24(1):356.
23. Sharma A, Jaiswal P, Kerakhan Y, et al. Liver disease and outcomes among COVID-19 hospitalized patients–a systematic review and meta-analysis. *Ann Hepatol*. 2021;21:100273.
24. Huang I, Pranata R. Lymphopenia in severe coronavirus disease-2019 (COVID-19): systematic review and meta-analysis. *J Intensive Care*. 2020;8(1):36.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Molaei S, Moazen H, Niazkar HR, Sabaei M, Johari MG, Rezaianzadeh A. Application of boosted trees to the prognosis prediction of COVID-19. *Health Sci Rep*. 2024;7:e2104. doi:10.1002/hsr2.2104