



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Machine learning model for predicting the length of stay in the intensive care unit for Covid-19 patients in the eastern province of Saudi Arabia

Dina A. Alabbad^{a,*}, Abdullah M. Almuhaideb^b, Shikah J. Alsunaidi^c, Kawther S. Alqudaihi^c, Fatimah A. Alamoudi^c, Maha K. Alhobaishi^a, Naimah A. Alaqeel^a, Mohammed S. Alshahrani^d

^a Department of Computer Engineering, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, 31441, Saudi Arabia

^b Department of Networks and Communications, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, 31441, Saudi Arabia

^c Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, 31441, Saudi Arabia

^d Department of Emergency Medicine, College of Medicine, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, 31441, Saudi Arabia

ARTICLE INFO

Keywords:

Coronavirus disease 2019 (COVID-19)
Intensive care unit (ICU)
Length of stay (LoS)
Predation
Resource management
Machine learning (ML)

ABSTRACT

The COVID-19 virus has spread rapidly throughout the world. Managing resources is one of the biggest challenges that healthcare providers around the world face during the pandemic. Allocating the Intensive Care Unit (ICU) beds' capacity is important since COVID-19 is a respiratory disease and some patients need to be admitted to the hospital with an urgent need for oxygen support, ventilation, and/or intensive medical care. In the battle against COVID-19, many governments utilized technology, especially Artificial Intelligence (AI), to contain the pandemic and limit its hazardous effects. In this paper, Machine Learning models (ML) were developed to help in detecting the COVID-19 patients' need for the ICU and the estimated duration of their stay. Four ML algorithms were utilized: Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), and Ensemble models were trained and validated on a dataset of 895 COVID-19 patients admitted to King Fahad University hospital in the eastern province of Saudi Arabia. The conducted experiments show that the Length of Stay (LoS) in the ICU can be predicted with the highest accuracy by applying the RF model for prediction, as the achieved accuracy was 94.16%. In terms of the contributor factors to the length of stay in the ICU, correlation results showed that age, C-Reactive Protein (CRP), nasal oxygen support days are the top related factors. By searching the literature, there is no published work that used the Saudi Arabia dataset to predict the need for ICU with the number of days needed. This contribution is hoped to pave the path for hospitals and healthcare providers to manage their resources more efficiently and to help in saving lives.

1. Introduction

The quick spread of the Coronavirus Disease (COVID-19) worldwide has threatened most healthcare systems. Generally, the rapid increase in the number of infected patients has raised the demand for Intensive Care Unit (ICU) beds [1]. The shortage in hospital resources and bed capacity is one of the most critical factors that have an impact on increasing death rates of COVID-19 [2]. A study based on a sample of COVID-19 patients collected from 88 US Department of Veterans Affairs hospitals indicated that the risk of death for COVID-19 patients in ICU increased with the

increase in the demand for the ICU compared to the period when the demand rate was approximately 25% lower [3].

Various measures were taken to address the lack of medical resources issue, including following specific guidelines for prioritizing patients and selecting those deserving of admission to the ICU for needed care [4, 5]. Although these measures play a significant role in managing medical resources, they, in turn, may put the lives of COVID-19 patients at risk, as happened in the United Kingdom when several patients adhered to home quarantine, which led to their death, and their condition was not discovered for a period up to two weeks [6].

* Corresponding author.

E-mail addresses: daalabbad@iau.edu.sa (D.A. Alabbad), amAlmuhaideb@iau.edu.sa (A.M. Almuhaideb), shikah.sunaidi@gmail.com (S.J. Alsunaidi), 2190500127@iau.edu.sa (K.S. Alqudaihi), 2170007764@iau.edu.sa (F.A. Alamoudi), mahaakhh@gmail.com (M.K. Alhobaishi), naimahalaqeel@gmail.com (N.A. Alaqeel), msshahrani@iau.edu.sa (M.S. Alshahrani).

<https://doi.org/10.1016/j.imu.2022.100937>

Received 21 September 2021; Received in revised form 31 March 2022; Accepted 31 March 2022

Available online 14 April 2022

2352-9148/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

During the period of such crises, it is necessary to estimate the future need for medical resources such as ICU beds and nasal oxygen support. For such purposes, Artificial Intelligence (AI), namely Machine Learning (ML), has made a significant contribution during the COVID-19 pandemic [7].

Several studies in the literature focused on predicting the number of patients who need admission to the ICU [8–13]. Moreover, identifying the factors that increase the likelihood of ICU admission are also important [8,14–17]. Several factors play role in determining the need for ICU admission for COVID-19 patients, e.g. air quality in the geographical region, which could indicate the rate and severity of the infection in specific countries [18]. Age also is an indicator, as Estiri et al. stated that patients at age of 51 and above are more likely to be hospitalized or face mortality risk [8]. Izquierdo et al. found that the most common clinical characteristics determining admission to ICU were fever, age, and tachypnea with/without respiratory crackles [17]. Aktar et al. illustrate the strong relationship between abnormal blood parameters and the severity level of the disease in COVID-19 patients [16]. Subudhi et al. found that the amount of LDH, CRP, and O₂ saturation had significant influence in making the ICU admission decision while eGFR < 60 ml/min/1.73m², and the percentages of lymphocytes and neutrophils were useful in predicting mortality [13]. Moreover, Heldt et al. found that selected laboratory findings such as creatinine, blood lactate levels, and clinical indicators of patient oxygenation were most predictive of endpoints for COVID-19 patients. Furthermore, patient age and measures of oxygen status during the ED stay can help in the initial prediction of poor patient outcomes [9].

By searching the literature, there are no published studies that targeted developing machine learning models to estimate the needed number of days in the ICU for COVID-19 patients in Saudi Arabia which is the focus of this study. Also, in this work, the most relevant features related to the ICU admission for COVID-19 patients in Saudi Arabia are being determined.

The research conducted in this paper aims at:

1. Reviewing the published works that aimed at predicting the need of COVID-19 patients for ICU admission and the expected number of days to be spent in the unit worldwide.
2. Developing machine learning models to predict and estimate the number of days COVID-19 patients in Saudi Arabia may spend in the ICU to assist health sectors to better manage their resources and determine their readiness to receive new patients and provide the needed care for them.
3. Identifying the most relevant features that indicate the patient's need for ICU and the expected length of stay in the unit.

The rest of this paper is organized as follows: Section 2 presents the related works, section 3 introduces the methodology, section 4 explains the conducted experiments, and the experimental results are discussed in Section 7. Finally, section 8 concludes the paper and presents future research directions.

2. Related works

The spread of COVID-19 had a significant impact on the healthcare systems around the globe. In particular, on the availability of beds in the intensive care unit. This pandemic increased the research community's interest in finding practical solutions to mitigate the effects of this problem, such as prediction models for ICU admissions, length of stay in ICU, and discharge date. To conduct our research, over 41 papers were reviewed. The main focus was on the latest published works in high reputed journals to analyze the effectiveness of ML models and algorithms in solving the problems of resource shortage.

Many studies have tackled the issue of predicting the need for different hospital resources for COVID-19 patients, and some of them have utilized machine learning for that goal. Works that identify the

related features for ICU admission are also reviewed. The below discussion reviews the related works and their used methods. It also provides a summary table of the related works in terms of aim, method, dataset, features, and results.

Epstein and Dexter [19] have designed an analytical model to predict the future need for beds and ventilators during the COVID-19 pandemic for a specific hospital by analyzing its internal data. The model used COVID-19 data of admitted patients to estimate the number of days they may need a ventilator. The authors found that there is no relationship between gender and duration of hospitalization, or between age and the need for ventilators. The performance of the model was good as the mean absolute error of the daily prediction was small < 1.25 patients/day for the census, and < 0.5 ventilators/day for ventilators. However, the results of the model could not be generalized, as the authors recommended resetting the input parameters for each hospital to achieve more accurate results.

In contrast with the previous study, López-Cheda et al. in [20] found that gender and age affect the stay duration in the ICU. They applied a non-parametric model to predict the LoS of COVID-19 patients in ICU and the time to discharge or death. They simulated the COVID-19 hospital demand using a Monte Carlo algorithm. The results found that the LoS in hospital is 11 days on average.

The work of Henzi et al., presented in [21], has also demonstrated that gender has an effect on the LoS in the ICU. A semi-parametric model to probabilistically predict the LoS of COVID-19 patients in ICU was applied. The model was trained using data from patients with acute respiratory distress syndrome (ARDS) and validated using data of COVID-19 patients. The results indicated that the LoS for females tended to be shorter than that for males.

Researchers are often interested in improving prediction scenarios that influence public decision guidance. To this end, Lapidus et al. in [22] conducted a study to assess the average LoS (ALoS) for COVID-19 patients in ICU by examining two methods of estimation, the Discharged Patient Estimation (DPE) and the Censored Patient Estimation (CPE). Although the true ICU_ALoS for the series was > 21 days, which is significantly higher than other reported estimates, they concluded that it is possible to rely on the ALoS in decision-making. They also recommended that censored patients should be included in the analysis along with discharged cases to reduce the bias rate.

Dan et al. built a machine learning model to predict ICU admission, LoS in the ICU, and mortality of COVID-19 patients. The model could predict events using clinical data collected within 1–15 days before actual admission to the ICU. They found that the length of stay in the ICU for elderly people with heart disease is high. Moreover, LoS in the ICU is affected by abnormal values of many factors, namely lymphocyte absolute value, erythrocyte count, total cholesterol, adenovirus IgM antibody, hypersensitive C-reactive protein, high sensitivity troponin I, and Q fever Rickettsia IgM antibody [23].

Identifying the factors and the clinical characteristics that help in predicting the admission of COVID-19 patients to the ICU is crucial too. Hong et al. have described COVID-19 clinical characteristics outside of Wuhan. They found that fever is not a permanent initial symptom of COVID-19 as only 70% of the study sample reported this symptom. They found also that age, weight, gender, and career are not affecting the length of hospital stay. They designed a multivariate model to predict the risk of a long hospital stay. Long periods of hospital stay increase medical costs and increase the level of risk. Early estimation helps in taking many decisions and allocating resources [24].

Moreover, the research conducted by Gunduz et al. in [10] found that the CHA₂DS₂-VASc (Congestive heart failure, hypertension, age, diabetes mellitus, stroke, vascular disease) and the M-CHA₂DS₂-VASc (modified CHA₂DS₂-VASc) scores can be used to predict the need of ICU admission, the LoS in the ICU, and mortality of COVID-19 patients.

Table 1 summarizes the related studies by showing their aim, the used model and algorithm, the dataset, features, and the obtained results. Table 2 shows the relationship between the LoS and some clinical

Table 1
Summary of works aimed at estimating the need and LoS in ICU for COVID-19 patients worldwide.

Ref	Aim	Model	Algorithms/ Methods	Dataset	Model inputs/Extracted features	Results
[10]	Predict ICU admission, LoS in the ICU, and mortality for COVID-19 patients	Multivariate logistic regression	LR	EHRs of 1668 COVID-19 patients at Merkezefendi State Hospital, Manisa, Turkey)	Clinical data	AUC: 0.89
[19]	Predict patient census and estimate ventilator needs for a specific hospital during the COVID-19 pandemic	Analytical model (Weibull distribution)	Linear and log-linear regression	EHRs from UHT and UIHC of COVID-19 patients	LoS in hospital, and duration of using the ventilator	MAE: <1.25 patients/day, <0.5 ventilators/day
[20]	Estimate LoS of hospitalized COVID-19 patients	Non-parametric model	–	EHRs of 10,454 confirmed COVID-19 cases in Galicia (Northwest Spain)	Age, and gender	–
[21]	Estimate the LoS of COVID-19 patients in ICU	Semiparametric distributional index model	Distributional regression model	EHRs of 2411 patients with ARDS for training, and EHRs of 557 COVID-19 patients for testing, the data from the Swiss Society of Intensive Care Medicine	Age, and gender	Accuracy predict patient discharged from the ICU in 20 days = 80%
[24]	Describe COVID-19 clinical characteristics outside of Wuhan and predict the risk of long LoS in hospital	Multivariate regression model	Statistical methods (t-test, Chi-square)	EHRs of 75 COVID-19 patients in Zhejiang Tertiary Care Hospital	Demographic data, comorbidities, laboratory results symptoms, and vital signs	AUC: 0.84
[23]	Predict ICU admission, LoS in the ICU, and mortality for COVID-19 patients	ML model	SVM	EHRs of 733 COVID-19 patients in Wuhan, China	Demographic, laboratory, and clinical data	Accuracy of: 1. Prediction of ICU admission (0.83, 0.84) 2. Prediction of ICU death (0.92, 0.98) MAE of the prediction of LoS in ICU (0.723) DPE and CPE estimates of ICU_ALoS (95% CI)
[22]	Estimate average LoS in the ICU for COVID-19 patients	Mathematical model	Two estimation methods: DPE and CPE	EHRs of COVID-19 patients entered the ICU of ZHWU	Age and gender	DPE and CPE estimates of ICU_ALoS (95% CI)

Note: Area Under the Curve (**AUC**), Acute Respiratory Distress Syndrome (**ARDS**), Electronic Healthcare Records (**EHR**), Discharged Patient Estimation (**DPE**), Censored Patient Estimation (**CPE**), Linear Regression (**LR**), Mean Absolute Error (**MAE**), Support Vector Machine (**SVM**), University of Iowa Hospitals and Clinics (**UIHC**), University of Miami UHealth Tower (**UHT**), Zhongnan Hospital of Wuhan University (**ZHWU**).

Table 2
Relationship between LoS and clinical features of patients.

Factor	Studies that confirmed relation	Studies that confirmed no relation
Age	[10,20,23]	[19,24]
Gender	[20,21]	[19,24]
Weight		[24]
Career		[24]
Heart disease	[10,23]	–
Hypertension	[10]	–
Diabetes mellitus	[10]	–
Stroke	[10]	–
Vascular disease	[10]	–
Lymphocyte absolute value	[23]	–
Erythrocyte count	[23]	–
Total cholesterol	[23]	–
Adenovirus IgM antibody	[23]	–
Hypersensitive C- reactive protein	[23]	–
High sensitivity troponin I	[23]	–
Q fever Rickettsia IgM antibody	[23]	–

features according to those studies.

3. Methodology

This section presents the approach to developing the prediction models to estimate the period of stay in the ICU for COVID-19 patients. It provides an overview of the approach, a brief about the machine learning algorithms utilized, and the dataset description and

preparation. The first step was to build the machine learning model. The second was to collect the dataset and prepare it for processing. Then, experiments were conducted to select the best-performing model. The evaluation of the developed models then took place. The final step is to utilize the developed model for the prediction purpose.

3.1. Overview of the proposed approach

This paper aims at predicting the likelihood of COVID-19 patients for ICU admission and the length of their stay using demographic and clinical data of positive COVID-19 patients obtained from King Fahad University Hospital in Dammam city in Saudi Arabia. The first step was to review different machine learning algorithms and select the most suitable ones to be utilized in developing the prediction models. Next, data pre-processing was applied to recover the missing data and solve the imbalanced data problem which is described in detail in section 3.3.2. After that, the selected machine learning algorithms were trained and validated on the selected dataset, once with and once without applying feature selection, and by trying different k-fold as demonstrated in section 4. The classification results of the developed models were compared in the third stage which is described in section 5 to identify the best performance model.

3.2. Machine learning algorithms

In this section, the machine learning classifiers used to build the prediction models are discussed. Since the literature addressed different classifiers as well-known to perform with high accuracy to solve the problem of developing prediction models for medical purposes, the first task was to select the top classifiers and test their performance for our proposed model. The selected classifiers are Random Forest RF, Gradient

Boosting GB, Extreme Gradient Boosting XGBoost, and Ensemble model. Based on [25], the most used ML algorithms for mortality, severity, and length of stay in ICU are Random Forest RF and XGBoost. In their work, the prediction model attained its best result by utilizing RF which helped to alarm the medical service providers in 6 h earlier that the patient must be admitted to the ICU.

In the research addressed in [26], Random Forest outperforms other classifiers in predicting COVID-19 affected cases, mortality, and cured cases in India.

Also, these algorithms have been applied for many prediction purposes in the medical field such as breast cancer, diabetes, and other diseases. For example, the work presented in [27] built a risk prediction model that utilized RF and XGBoost algorithms for weighted feature selection to diagnose type 2 diabetes. They stated that the best prediction accuracy was achieved by using RF. XGBoost classifier can be used to predict accurately the infection with breast cancer, and it can achieve the highest accuracy as discussed in [28]. Also in [29], GB and LR were utilized to predict the need for recurrent bleeding, therapeutic intervention, and severe bleeding. The work demonstrated that the GB algorithm is a robust classification that can handle large input sizes and fits with simple models to achieve higher accuracy.

The proposed models in this work were implemented using Python programming language which provides several tools for machine learning tasks.

Below is a brief description of the used classifiers to build the proposed models.

a. Random Forest (RF)

The random forest, as the name implies, is made up of a huge number of individual decision trees that work together as an ensemble. It works to enhance accuracy by relying on a group of decision models rather than a single learning model. The key distinction between this approach and traditional decision tree algorithms is that the root nodes have splitting nodes that are produced at random [30]. The trees protect each other from their flaws, which explains why they have such a strong effect. While some trees may estimate incorrect classification, many others will be correct, allowing the trees to progress in the proper direction. As a result, the predictions, and thus errors, generated by individual trees must have minimal correlations with each other for the random forest to perform well [31]. Furthermore, RF offers many advantages, such as the ability to be utilized for both classification and regression tasks, and it can process missing variables. Additionally, when additional decision trees are added to the forest, overfitting is less likely to occur [32].

b. Gradient Boosting (GB)

Gradient boosting is a type of boosting technique that is an ensemble mechanism for combining numerous simple models into a single composite model. The entire model becomes a stronger predictor when additional simple models are introduced [33]. Each model attempts to compensate for the flaws of its predecessor by selecting a random sample of data, fitting it with a model, and then training it consecutively. Each iteration combines the weak rules of each classifier to generate a single and strong prediction rule [34]. Gradient boosting is a technique that can be utilized for both regression and classification tasks [33]. It can train many models in a progressive, cumulative, and sequential manner. The input of GB is three requirements which are: loss function, prediction maker (weak learner) which is generally a decision tree, and an additive model which minimizes the loss function by adding the weak learners. The loss function describes how the dataset is modeled by the algorithm. Mainly, it is the gap between actual and projected values. It has a different function for each task. For example, for the classification task, the binary cross entropy loss can be used. On the other hand, the decision trees can be used as a weak learner to reduce the error generated from the previous models and return a strong model [35].

c. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting is a fast open-source version of the stochastic gradient boosting ensemble method which is donated as a scalable tree boosting system. As an ensemble machine learning algorithm, XGBoost is built on decision tree models. Trees are introduced to the ensemble one at a time and fit to correct the prediction mistakes caused by preceding models [36]. XGBoost has been deployed on a variety of challenges, and the results show that this algorithm produces state-of-the-art outcomes on a wide range of difficult problems [37]. It is intended to be both computationally efficient (i.e. fast to execute) and extremely effective. The most significant factor in XGBoost's success is its scalability in all situations. On a single machine, the system operates more than 10 times quicker than the existing popular solutions and scales to billions of samples in distributed or memory-limited scenarios.

d Ensemble Classifier

By integrating many models, ensemble learning aids in improving machine learning results. This method is expected to produce greater predictive performance than a single model because it employs a group of classifiers rather than a single one to classify unknown data. The ensemble's classifiers all anticipate the proper class of each unseen instance, and their predictions are subsequently aggregated using some kind of voting system [38]. It is only appropriate to employ this approach if the output of multiple classifiers disagrees. Combining a large number of identical classifiers provides no benefit. As a result, approaches for creating ensembles revolve around producing classifiers that disagree with their predictions [39]. Table 3 discusses the strengths and weaknesses of the selected prediction algorithms.

3.3. Data collection and preparation

This section describes the dataset used in the research and the process of data preparation and pre-processing.

3.3.1. Data description

The dataset was obtained from King Fahad University hospital in Dammam City at the eastern province of Saudi Arabia. Since the start of the pandemic, this hospital was one of the main centers to receive and treat COVID-19 patients from all cities in the eastern province. The dataset consists of the clinical and demographics data of 895 patients who attended to the hospital and tested positively for COVID-19. Each patient record contained 47 features collected at admission time and from the medical history of the patient, where each patient's record was updated daily. Those features included a wide range of clinical and demographics data that contains comorbidities, laboratory results symptoms, and vital signs such as age, gender, obesity, smoking, vitamin D def, fever, headache, liver disease, ferritin, LDH, AST, Trop, and other data. The data set also contained the number of days each patient spent in the ICU before being discharged or expired. The range of the days in the data was between 0 and 58 days. For classification and prediction purposes, these days were categorized in small intervals of 5 days and given one label of 9 classes (from 1 to 9) to represent the LoS days data. The Zero class was used to label COVID-19 patients who did not need to attend the ICU. Class 1 was used to represent patients who entered the ICU for less than 24 h. The remaining classes 2, 3, 4, 5, 6, 7, 8, and 9 were used to represent the following period of days respectively: 1–5, 6–10, 11–15, 16–20, 21–25, 26–30, and more than 30 days.

3.3.2. Data preparation and preprocessing

The data preparation and preprocessing involved two tasks: first, filling in the missing data, second, solving the issue of imbalanced data.

Since the dataset contained some missing values, the KNN imputation method was used to fill in these missing values. The KNN imputation algorithm replaces the missing value with a value obtained from a

Table 3
Strengths and weaknesses for the implemented ML algorithms RF, GB, XGB, and ensemble classifier.

Algorithm	Strength	Weakness
Random Forest (RF)	<ul style="list-style-type: none"> • Collection of decision trees that fit the data and cause high variation in classification • Data classification is based on the most votes. • Lower chance of variation in data training. • Good scale for big dataset. • Knows what is better fits in the classification [40]. 	<ul style="list-style-type: none"> • Very sensitive to training data which makes it error-prone. • Complex and computationally expensive • The base classifiers need to be defined • It prefers the parameters that take higher different values [40].
Gradient Boosting (GB)	<ul style="list-style-type: none"> • It improves the prediction performance [41]. • The algorithm builds relations by shortening the number of errors from old weak classifiers [42]. 	<ul style="list-style-type: none"> • Up-sampling of similar data does not show any impact in improving results [42].
Extreme Gradient Boosting (XGBoost)	<ul style="list-style-type: none"> • Designed to be used with large complex datasets and avoid model overfitting. • The method is scalable in all cases. • It can handle sparse data and also parallel and distributed computation which makes learning process faster and quicker [43]. • Always involves many classification and regression trees [44]. 	<ul style="list-style-type: none"> • Complex and computationally expensive [40].
Ensemble Classifier	<ul style="list-style-type: none"> • It is combined by weighted averaging or the voting of a collection of single classifiers. • The ensemble method combines multiple weak classifiers as a strong classifier. An empirical study shows that the price of building a base classifier is lower than the price of building a strong classifier. • It can maximize the information of the base learner and improve the overall ability of classification [45]. 	<ul style="list-style-type: none"> • The method robustness is affected by the quality of the dataset [45].

neighbor of this empty value. The parameter k defines the number of neighbors to be included in the voting process. Besides the capability of this algorithm of obtaining a closer value as much as possible and using it in filling the missing value. The KNN imputation algorithm can preserve the normal distribution of the data which is very important in the case of medical datasets [46]. Since the usage of different k values in the KNN imputation algorithm can result in different results, the model performance was the average of using four different k values: $k = 3$, $k = 5$, $k = 7$, $k = 10$, $k = 15$.

Similar to a lot of other datasets, this dataset is considered an imbalanced dataset. Fig. 1 shows the frequency of every class in the attribute “Days to discharge from the ICU”. As shown, the majority of the records belong to class 0 while class 6 and class 7 had only 30 and 12 records respectively. This may make the model biased to class 0. The imbalanced dataset is considered a problem as the model will not be able to predict the minority classes because the model will focus only on optimizing the accuracy without taking into consideration the overall distribution of each class. To solve this imbalanced dataset issue, Synthetic Minority Oversampling Technique (SMOTE) was used. This technique oversamples the minority class by creating “synthetic”

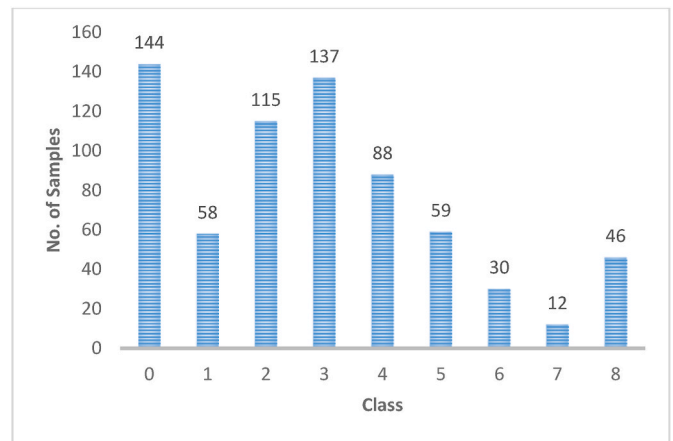


Fig. 1. Days to discharge from ICU class distribution before oversampling.

examples rather than already existing examples [47]. Fig. 2 demonstrates the frequency of every class in the attribute “Days to discharge from the ICU” after applying SMOTE technique in which all the classes now have the same number of records which is 144 records.

Also, the dataset attributes were evaluated using the entropy evaluation to understand how the impurity or the heterogeneity of the target class is computed. Table 4 below demonstrates the result of the entropy evaluation for each attribute.

4. Experiments

In order to achieve the goal of this research, a number of experiments were conducted to test the performance of the four developed prediction models mentioned earlier in section 3.2. The data were randomly divided into two folds, 80% for the training process, and the remaining 20% was for validation. All models were run using different K-fold values: 3, 5, 7, 10, and 15. The same dataset was used with each prediction model, and the performance was evaluated. Because the dataset is imbalanced, we could not depend on one performance measure. In this case, the performance measures included accuracy (the proportion of correctly classified test records), precision (positive predictive value), recall (negative predicted value), and f-score (the harmonic mean of precision and recall). Also, feature selection was applied once with each classifier using the Boruta algorithm to reduce the feature set each model needs to handle. The developed models and their performance are discussed in the below sections.

4.1. Model 1: random forest (RF)

The first prediction model was built using RF classifier. Table 5

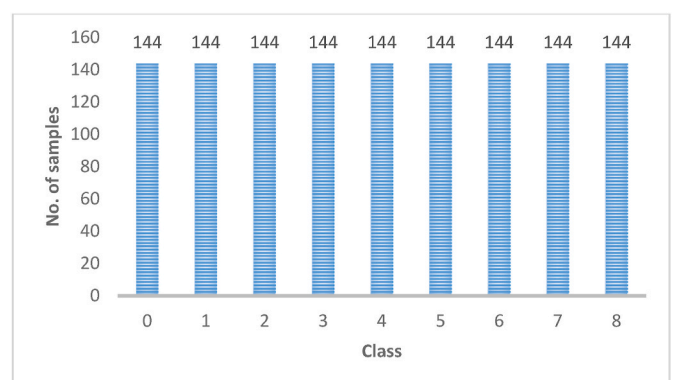


Fig. 2. Days to discharge from ICU class distribution after oversampling.

Table 4
Entropy values for features included in the study.

Attribute	Entropy Evaluation
Gender	0.64
Age	4.03
DM	0.71
HTN	0.68
Cardiac	0.40
Obesity	0.31
Smoking	0.2
Vitamin D def	0.04
Renal	0.23
Liver disease	0.12
Autoimmune disease	0.10
Fever	0.71
Dry Cough	0.71
Fatigue	0.45
Headache	0.35
Dyspnea	0.64
Flu Symptoms	0.36
diarrhea	0.53
RR	2.86
Blood oxygen saturation (SATS)	3.28
Mode of infection	0.80
1st sample	0.08
Chloroquine/ Hydroxychloroquine	0.72
Azithromycin/ Antibiotics	0.43
Steroids	0.63
Tocilizumab	0.56
Respiratory failure	0.17
Acute renal failure	0.10
Acute coronary syndrome	0.06
ARDS	0.06
GI complication	0.03
Nosocomial infection	0.03
Septic shock	0.06
Organ dysfunction or failure	0.13
lymph	5.35
Neut	5.56
LDH	5.64
ALT	4.34
AST	4.57
Trop	4.73
ferritin	5.93
D-dimer	5.47
CRP	5.90

Table 5
The results of predicting the number of days to discharge from ICU class for **Model 1**.

K-fold	Accuracy	Precision	Recall	F1-Score
3	94.16%	94.14%	94.16%	94.14%
5	93.55%	93.51%	93.55%	93.53%
7	87.23%	87.23%	87.23%	87.20%
10	86.07%	86.11%	86.07%	85.96%
15	92.38%	92.31%	92.38%	92.33%

demonstrates the results obtained by this model on the validation set for predicting the class of the number of days to discharge from ICU using different values of K-fold. The 3-fold achieved the highest accuracy performance compared to the other K-folds using the RF classifier. The

Table 6
The results of predicting the number of days to discharge from ICU class for **Model 1** with feature selection.

K-fold	Accuracy	Precision	Recall	F1-Score
3	93.30%	93.30%	93.30%	93.30%
5	89.45%	89.24%	89.45%	89.26%
7	87.23%	87.26%	87.23%	87.17%
10	86.30%	86.36%	86.30%	86.28%
15	79.59%	80.17%	79.59%	79.67%

results of RF classifier after the feature selection are shown in **Table 6**, which also shows that 3-fold provides the highest performance.

4.2. Model 2: gradient boosting (GB)

The second model was developed using GB classifier, **Table 7** and **Table 8** show the results of validating the model without and with applying feature selection. Noticed that, the highest accuracy achieved does not depend on the number of K-fold where it was 88.14% with 15-fold before the feature selection and 87.29% with 3-fold after the feature selection.

4.3. Model 3: Extreme Gradient Boosting (XGBoost)

The third model was built using the XGBoost classifier which achieved the highest accuracy with 3-fold with and without applying the feature selection which was 91.49% and 90.80% respectively. **Table 9**, and **Table 10** show the results obtained by using the third model with XGBoost classifier.

4.4. Model 4: Ensemble Classifier

The fourth model was developed using the Ensemble classifier where RF, GB, XGBoost, and Adaptive Boosting decisions techniques were combined. The obtained results with this model are shown in **Table 11** which achieved 93.13% accuracy without feature selection. **Table 12** shows the results of the model with applying feature selection which scored a slightly higher accuracy of 93.81%.

5. Experimental results and discussion

The goal of the presented research is to develop a machine learning model that estimates the required number of days for COVID-19 patients in the ICU to help healthcare providers to manage their resources and to plan for the expected COVID-19 affected patients. For this purpose, four prediction models were implemented using different ML classifiers with multiple tuning of the parameters. The results show that each model performed differently with the used dataset. By reviewing the obtained results in the previous section, it can be deduced that with the majority of the classifiers, using the full set of features improves the prediction accuracy compared to applying features selection. Only Model 4 (the ensemble) showed slight improvement with feature selection using Borota as the accuracy reached 93.81% compared to 93.13% before applying the feature selection. **Table 13** shows the results achieved by each model by applying feature selection. **Table 14** shows the highest obtained results of each classifier without the feature selection. Results were compared in order to identify the highest performance model.

Fig. 3 also shows the obtained accuracy of each model before and after feature selection.

By comparing the obtained results of all models, we can see that model 1, which applied the Random Forest, has achieved the highest validation accuracy of 94.16% to predict the number of days range to discharge from ICU by using 3-fold. It can be seen that RF provides the best result in all performance measures as it scored 94.16%, 94.14%, 94.16%, and 94.14% for accuracy, precision, recall, and F-score

Table 7
The results of predicting the number of days to discharge from ICU class for **Model 2**.

K-fold	Accuracy	Precision	Recall	F1-Score
3	86.21%	86.24%	86.21%	86.11%
5	85.16%	85.26%	85.16%	85.03%
7	86.33%	86.37%	86.33%	86.30%
10	83.33%	83.50%	83.33%	83.36%
15	88.14%	88.17%	88.14%	88.08%

Table 8

The results of predicting the number of days to discharge from ICU class for **Model 2** with feature selection.

K-fold	Accuracy	Precision	Recall	F1-Score
3	87.29%	87.09%	87.29%	87.02%
5	83.98%	83.80%	83.98%	83.80%
7	80.85%	80.93%	80.85%	80.62%
10	76.48%	76.42%	76.48%	76.25%
15	77.26%	77.99%	77.26%	77.48%

Table 9

Results of predicting the number of days to discharge from ICU class for **Model 3**.

K-fold	Accuracy	Precision	Recall	F1-Score
3	91.41%	91.49%	91.41%	91.42%
5	91.21%	91.10%	91.21%	91.10%
7	87.02%	87.07%	87.02%	86.97%
10	83.56%	83.67%	83.56%	83.49%
15	82.69%	83.27%	82.69%	82.89%

Table 10

The result of predicting the number of days to discharge from ICU class for **Model 3** with feature selection.

K-fold	Accuracy	Precision	Recall	F1-Score
3	90.72%	90.80%	90.72%	90.78%
5	89.26%	89.27%	89.26%	89.20%
7	85.32%	85.04%	85.32%	84.98%
10	86.99%	87.05%	86.99%	87.00%
15	82.69%	83.39%	82.69%	82.86%

Table 11

The result of predicting the number of days to discharge from ICU class for **Model 4**.

K-fold	Accuracy	Precision	Recall	F1-Score
3	93.13%	93.42%	93.13%	93.19%
5	92.58%	92.56%	92.58%	92.54%
7	88.94%	89.55%	88.94%	88.96%
10	91.10%	91.39%	91.10%	91.20%
15	85.01%	85.44%	85.01%	85.12%

Table 12

The result of predicting the number of days to discharge from ICU class for **Model 4** with feature selection.

K-fold	Accuracy	Precision	Recall	F1-Score
3	93.81%	93.79%	93.81%	93.78%
5	92.19%	92.23%	92.19%	92.13%
7	87.45%	87.53%	87.45%	87.45%
10	88.81%	88.85%	88.81%	88.72%
15	86.56%	87.18%	86.56%	86.71%

Table 13

Results of investigating the effect of feature selection on the dataset.

Algorithm	Accuracy	Precision	Recall	F1-Score
Model 1 (RF)	87.27%	87.27%	87.17%	87.14%
Model 2 (GB)	81.17%	81.25%	81.17%	81.03%
Model 3 (XGBoost)	87.00%	87.11%	87.00%	86.96%
Model 4 (Ensemble)	93.81%	93.79%	93.81%	93.78%

respectively. The rest of the classifiers yielded slightly lower performance. For all developed models, feature selection did not contribute to improving performance. The feature selection process was applied in

Table 14

Results of using the complete features set to predict the number of days to discharge from ICU.

Model	Accuracy	Precision	Recall	F1-Score
Model 1 (RF)	94.16%	94.14%	94.16%	94.14%
Model 2 (GB)	88.14%	88.17%	88.14%	88.08%
Model 3 (XGBoost)	91.41%	91.49%	91.41%	91.42%
Model 4 (Ensemble)	93.13%	93.42%	93.13%	93.19%

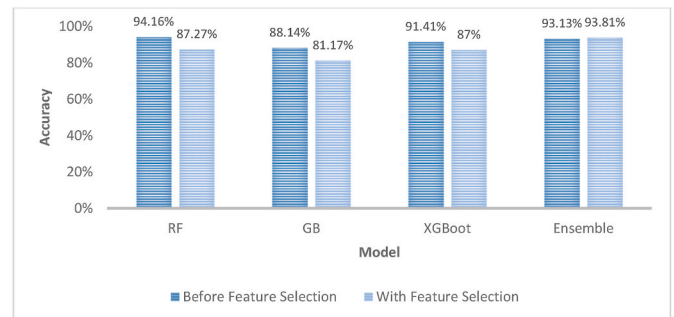


Fig. 3. Prediction models accuracy before and after feature selection.

hope of reducing overfitting, reducing training time, and improving accuracy by eliminating irrelevant features. However, by applying the feature selection, the performance was slightly decreased as the achieved accuracy for the majority of the classifiers was less compared to using the full set of features. This indicates that all used features contribute to the final decision and are all important.

Model 1 results were also compared to the other works in the literature presented in [10,20,24] where different datasets were used with ML prediction models. The comparison confirmed that the implemented model in this work achieved higher accuracy. Table 15 shows a comparison between the related works and the current research in terms of the method, performance measurements, features, geographical location, and dataset size.

The second goal of this work was to identify the features that are most relevant to patients' need for ICU and their stay length. Therefore, the feature correlation with the number of days to discharge from the ICU was extracted from the heatmap. Table 16 shows the top 10 features that are highly correlated with the LoS for COVID-19 patients in Saudi Arabia. The obtained correlation results show that age, C-Reactive Protein, nasal oxygen support days, organ dysfunction or failure, LDH, blood oxygen saturation, headache, autoimmune disease, D-dimer, and ferritin are the features that can strongly indicate the need for patients to be in ICU. These obtained results are aligned generally with the identified features relevant to ICU admission in the literature which was presented earlier in Table 2.

6. Conclusion and future work

Artificial intelligence technologies have been widely employed in the battle against the COVID-19 pandemic as they support medical services with great tools to predict, monitor, track, and contain the pandemic. This work aimed to provide a model to help in managing the medical resources and facilitate the services for COVID-19 patients with severe conditions. The work presented in this paper aimed first to develop a prediction model to predict the need for COVID-19 patients to be in the ICU and to estimate the duration in days of their stay. Four prediction models were developed using ML classifiers: RF, GB, XGBoost, and Ensemble as those four classifiers are known to perform well for the targeted purpose. The developed model achieved high accuracy of 94.16% with Random Forest which provides a reliable tool to help hospitals and healthcare providers to predict the period where the ICU

Table 15
Comparison of obtained results with the related studies.

Ref	Method	Measurements	Features	Dataset location	Dataset size
[10]	LR	AUC: 0.89	Clinical data	Manisa, Turkey	1668
[19]	LR	MAE: <1.25 patients/day, <0.5 ventilators/day	LOS in hospital, and duration of using the ventilator	Miami-USA	–
[20]	Nonparametric mixture cure model	–	Age, and gender	Spain	10,454
[21]	Distributional regression model	Accuracy = 80%	Age, and gender	Switzerland	2411/ 557
[22]	DPE and CPE	DPE and CPE estimates of ICU-ALoS (95% CI)	Age, and gender	ZHWU	59
[23]	SVM	Accuracy: 1. Prediction of ICU admission (83%,84%) 2. Prediction of ICU death (92%, 98%) MAE of the prediction of LoS in ICU (0.723)	Demographic, and clinical data	Wuhan, China	733
[24]	Statistical methods (t-test, Chi-square)	AUC: 0.84	Demographic and clinical data	Zhejiang Tertiary	75
Ours	RF, GB, XGBoost, Ensemble	Accuracy: 94.16%	Demographic and clinical data	Saudi Arabia	895

Note: Area Under the Curve (**AUC**), Acute Respiratory Distress Syndrome (**ARDS**), Electronic Healthcare Records (**EHR**), Discharged Patient Estimation (**DPE**), Censored Patient Estimation (**CPE**), Linear Regression (**LR**), Mean Absolute Error (**MAE**), Support Vector Machine (**SVM**), University of Iowa Hospitals and Clinics (**UIHC**), University of Miami UHealth Tower (**UHT**), Zhongnan Hospital of Wuhan University (**ZHWU**).

Table 16

Top 10 features that are highly correlated with the LoS for COVID-19 patients in Saudi Arabia.

Feature	Rank
Age	1
C-Reactive Protein CRP	2
Nasal Oxygen Support days	3
Organ dysfunction or failure	4
LDH	5
Blood Oxygen saturation (SATS)	6
Headache	7
Autoimmune disease	8
D-dimer	9
Ferritin	10

beds will be available to organize and control the process of receiving new patients in the ICU. The second goal was to find the features related to the need for ICU for COVID-19 patients in the eastern province of Saudi Arabia. The obtained results showed that age, CRP, nasal oxygen support days are the top features related to the need for ICU admission and the length of stay in the unit.

The future direction is to utilize the implemented model and the collected data to estimate and predict the need for COVID-19 patients for nasal oxygen support to better manage these resources. Also, the implemented model can be used with a larger dataset collected from the main regions of Saudi Arabia to be able to generalize the model to the kingdom or worldwide.

Funding

This work was supported in part by the Imam Abdulrahman bin Faisal University/College of Computer Science & Information Technology through the Deputyship for Research & Innovation, Ministry of Education, Saudi Arabia under Project Covid19-2020-067-CSI.

Institutional review board statement

Not applicable.

Informed consent statement

Informed consent was obtained from all subjects involved in the study.

Data availability statement

Data was obtained from King Fahd Hospital of Imam Abdulrahman Bin Faisal University and are available from the authors with the permission of King Fahd Hospital of the University, IRB No. IRB-2020-09-189.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors extend their appreciation to the Deputyship of Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number Covid19-2020-067-CSIT at Imam Abdulrahman bin Faisal University/ College of Computer Science and Information Technology.

Summary

Since the World Health Organization WHO declared the Coronavirus disease 2019 (COVID-19) as a pandemic in March 2020, all countries around the world devoted their efforts and resources to contain the rapid spread of the disease and limit its hazardous effects on individuals' lives. One of the challenges that faces most hospitals and healthcare providers is managing the hospital resources and identifying when these resources will be available to serve new patients. Intensive Care Unit (ICU) beds are of the resources that witnessed high demand during the pandemic as COVID-19 is a respiratory disease and some patients develop complications when getting affected, which requires ICU admission. This work reviewed the published work that aimed at estimating the length of stay in the ICU for COVID-19 patients. It has been found that no yet published study estimated the ICU length of stay in the Saudi Arabia dataset. This research aims at utilizing machine learning techniques to develop a prediction model to assess the need for COVID-19 patients for ICU admission and to estimate the number of days needed. A dataset of 895 positive COVID-19 patients at King Fahad University hospital in Dammam in Saudi Arabia was collected. It contained 47 clinical and demographic features collected at admission time, from the medical history of the patient, and updated daily until admitted to the ICU. Data were prepared and pre-processed to be used with the models, KNN-

imputation was applied to fill in the missing data. The imbalanced data issue was resolved by applying the Synthetic Minority Oversampling Technique (SMOTE).

Four prediction models were developed using four ML classifiers: Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), and Ensemble. The dataset was split into 2 folds, 80% for training and 20% for testing. The same data was used with all models to compare the performance. A number of experiments were conducted with different k-fold. Feature selection was applied using the Boruta algorithm to reduce the number of features. However, using the full set of features proved to provide higher accuracy.

The conducted experiments show that the highest accuracy could be achieved by applying the RF model for prediction as the accuracy was 94.16% which is high accuracy compared to the other published works for the same purpose. Feature correlation with the length of stay in ICU was obtained, and the results showed that age, CRP, nasal oxygen support days are the top indicators that are related to the patient's need for the ICU admission. This work hopes to pave the path for hospitals and healthcare providers in Saudi Arabia to manage their resources to encounter the pandemic in the best manner.

References

- Zhuang Z, Cao P, Zhao S, Han L, He D, Yang L. The shortage of hospital beds for COVID-19 and non-COVID-19 patients during the lockdown of Wuhan, China. *Ann Transl Med* 2021;9. <https://doi.org/10.21037/atm-20-5248>. 200–200.
- Sen-Crowe B, Sutherland M, McKenney M, Elkbuli A. A closer look into global hospital beds capacity and resource shortages during the COVID-19 pandemic. *J Surg Res* 2021;260:56–63. <https://doi.org/10.1016/j.jss.2020.11.062>.
- Bravata DM, Perkins AJ, Myers LJ, Arling G, Zhang Y, Zillich AJ, Reese L, Dysangco A, Agarwal R, Myers J, Austin C, Sexson A, Leonard SJ, Dev S, Keyhani S. Association of intensive care unit patient load and demand with mortality rates in US department of Veterans Affairs hospitals during the COVID-19 pandemic. *JAMA Netw Open* 2021;4:e2034266. <https://doi.org/10.1001/jamanetworkopen.2020.34266>.
- Tyrrell CSB, Mytton OT, Gentry SV, Thomas-Meyer M, Allen JLY, Narula AA, McGrath B, Lupton M, Broadbent J, Ahmed A, Mavrodaris A, Abdul Pari AA. Managing intensive care admissions when there are not enough beds during the COVID-19 pandemic: a systematic review. *Thorax* 2021;76:302–12. <https://doi.org/10.1136/thoraxjnl-2020-215518>.
- Jaziri R, Alnahdi S. Choosing which COVID-19 patient to save? The ethical triage and rationing dilemma. *Ethics, Med. Public Heal.* 2020;15:100570. <https://doi.org/10.1016/j.jemep.2020.100570>.
- Denis. In: Campbell Health policy, editor. UK coronavirus victims have lain undetected at home for two weeks, 1–4; 2020. <https://www.theguardian.com/world/2020/jun/07/uk-coronavirus-victims-have-lain-undetected-at-home-for-two-weeks>. [Accessed 25 July 2021].
- Alsunaidi SJ, Almuhaideb AM, Ibrahim NM, Shaikh FS, Alqudaihi KS, Alhaidari FA, Khan IU, Aslam N, Alshahrani MS. Applications of big data analytics to control COVID-19 pandemic. *Sensors* 2021;21:2282. <https://doi.org/10.3390/s21072282>.
- Estiri H, Strasser ZH, Murphy SN. Individualized prediction of COVID-19 adverse outcomes with MLHO. *Sci Rep* 2021;11:5322. <https://doi.org/10.1038/s41598-021-84781-x>.
- Heldt FS, Vizcaychipi MP, Peacock S, Cinelli M, McLachlan L, Andreotti F, Jovanović S, Dürichen R, Lipunova N, Fletcher RA, Hancock A, McCarthy A, Pointon RA, Brown A, Eaton J, Liddi R, Mackillop L, Tarassenko L, Khan RT. Early risk assessment for COVID-19 patients from emergency department data using machine learning. *Sci Rep* 2021;11:4200. <https://doi.org/10.1038/s41598-021-83784-y>.
- Gunduz R, Yildiz BS, Ozdemir IH, Cetin N, Ozen MB, Bakir EO, Ozgur S, Bayturan O. CHA2DS2-VASc score and modified CHA2DS2-VASc score can predict mortality and intensive care unit hospitalization in COVID-19 patients. *J Thromb Thrombolysis* 2021. <https://doi.org/10.1007/s11239-021-02427-1>.
- Yu L, Halalau A, Dalal B, Abbas AE, Ivascu F, Amin M, Nair GB. Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. *PLoS One* 2021;16:1–18. <https://doi.org/10.1371/journal.pone.0249285>.
- Roimi M, Gutman R, Somer J, Ben Arie A, Calman I, Bar-Lavie Y, Gelbshtein U, Liverant-Taub S, Ziv A, Eytan D, Gorfine M, Shalit U. Development and validation of a machine learning model predicting illness trajectory and hospital utilization of COVID-19 patients: a nationwide study. *J. Am. Med. Informatics Assoc.* 2021:1–9. <https://doi.org/10.1093/jamia/ocab005.00>.
- Subudhi S, Verma A, Patel AB, Hardin CC, Khandekar MJ, Lee H, McEvoy D, Stylianopoulos T, Munn LL, Dutta S, Jain RK. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *Npj Digit. Med.* 2021;4:1–7. <https://doi.org/10.1038/s41746-021-00456-x>.
- Chang R, Elhusseiny KM, Yeh Y-C, Sun W-Z. COVID-19 ICU and mechanical ventilation patient characteristics and outcomes—a systematic review and meta-analysis. *PLoS One* 2021;16:e0246318. <https://doi.org/10.1371/journal.pone.0246318>.
- Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. *Comput Struct Biotechnol J* 2021;19: 2833–50. <https://doi.org/10.1016/j.csbj.2021.05.010>.
- Aktar S, Ahamad MM, Rashed-Al-Mahfuz M, Azad A, Uddin S, Kamal A, Alyami SA, Lin P-I, Islam SMS, Quinn JM, Eapen V, Moni MA. Machine learning approach to predicting COVID-19 disease severity based on clinical blood test data: statistical analysis and model development. *JMIR Med. Informatics* 2021;9:e25884. <https://doi.org/10.2196/25884>.
- Izquierdo JL, Ancochea J, Soriano JB. Clinical characteristics and prognostic factors for intensive care unit admission of patients with COVID-19: retrospective study using machine learning and natural language processing. *J Med Internet Res* 2020;22:e21801. <https://doi.org/10.2196/21801>.
- Mangayarkarasi R, Vanmathi C, Zubair Khan M, Noorwali A, Jain R, Agarwal P. COVID19: forecasting air quality index and particulate matter (PM2.5). *Comput Mater Continua (CMC)* 2021;67:3363–80. <https://doi.org/10.32604/cmc.2021.014991>.
- Epstein RH, Dexter F. A predictive model for patient census and ventilator requirements at individual hospitals during the Coronavirus disease 2019 (COVID-19) pandemic: a preliminary technical report, vol. 2019; 2020. <https://doi.org/10.7759/cureus.8501>. Cureus.
- López-Cheda A, Jácome MA, Cao R, De Salazar PM. Estimating lengths-of-stay of hospitalized COVID-19 patients using a non-parametric model: a case study in Galicia (Spain). *Epidemiol Infect* 2021. <https://doi.org/10.1017/S0950268821000959>.
- Henzi A, Kleger G-R, Hilty MP, Wendel Garcia PD, Ziegel JF. Probabilistic analysis of COVID-19 patients' length of stay in Swiss intensive care units. *PLoS One* 2021; 16:e0247265. <https://doi.org/10.1371/journal.pone.0247265>.
- Lapidus N, Zhou X, Carrat F, Riou B, Zhao Y, Hejblum G. Biased and unbiased estimation of the average length of stay in intensive care units in the Covid-19 pandemic. *Ann. Intensive Care* 2020;10:135. <https://doi.org/10.1186/s13613-020-00749-6>.
- Dan T, Li Y, Zhu Z, Chen X, Quan W, Hu Y, Tao G, Zhu L, Zhu J, Jin Y, Li L, Liang C, Wen H, Cai H. Machine learning to predict ICU admission, ICU mortality and survivors' length of stay among COVID-19 patients: toward optimal allocation of ICU resources. In: 2020 IEEE int. Conf. Bioinforma. Biomed. IEEE; 2020. p. 555–61. <https://doi.org/10.1109/BIBM49941.2020.9313292>.
- Hong Y, Wu X, Qu J, Gao Y, Chen H, Zhang Z. Clinical characteristics of Coronavirus Disease 2019 and development of a prediction model for prolonged hospital length of stay. *Ann Transl Med* 2020;8. <https://doi.org/10.21037/atm.2020.03.147>. 443–443.
- Su L, Xu Z, Chang F, Ma Y, Liu S, Jiang H, Wang H, Li D, Chen H, Zhou X, Hong N, Zhu W, Long Y. Early prediction of mortality, severity, and length of stay in the intensive care unit of sepsis patients based on sepsis 3.0 by machine learning models. *Front Med* 2021;8.
- Gupta VK, Gupta A, Kumar D, Sardana A. Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model. *Big Data Min. Anal.* 2021;4:116–23. <https://doi.org/10.26599/BDMA.2020.9020016>.
- Xu Z, Wang Z. A risk prediction model for type 2 diabetes based on weighted feature selection of random forest and XGboost ensemble classifier, 2019 elev. Int. Conf. Adv. Comput. Intell. 2019:278–83.
- Sinha NK, Khulal M, Gurung M, Lal A. Developing A web based system for breast cancer prediction using XGboost classifier. *Int J Eng Res* 2020;9:852–6.
- Ayaru L, Ypsilantis P-P, Nanapragasam A, Choi RC-H, Thillanathan A, Min-Ho L, Montana G. Prediction of outcome in acute lower gastrointestinal bleeding using gradient boosting. *PLoS One* 2015;10:e0132485. <https://doi.org/10.1371/journal.pone.0132485>.
- Tin Kam Ho. Random decision forests. In: Proc. 3rd int. Conf. Doc. Anal. Recognit. IEEE Comput. Soc. Press; 1995. p. 278–82. <https://doi.org/10.1109/ICDAR.1995.598994>.
- Yiu T. Understanding random forest. *Towar. Data sci.* 2019;1–10. <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>. [Accessed 25 July 2021].
- Kho J. Why random forest is my favorite machine learning model, towar. *Data sci.* 2018. p. 1–7. <https://towardsdatascience.com/why-random-forest-is-my-favorite-machine-learning-model-b97651fa3706>. [Accessed 25 July 2021].
- Mujtaba H. What is gradient boosting and how is it different from AdaBoost? *Gt. Learn* 2020:1–8. <https://www.mygreatlearning.com/blog/gradient-boosting/>. [Accessed 25 July 2021].
- IBM Cloud Education. Learn about boosting algorithms and how they work. IBM Cloud Learn Hub; 2021. p. 1–7. <https://www.ibm.com/cloud/learn/boosting>. [Accessed 25 July 2021].
- Kumar A. Introduction to gradient boosting algorithm. *Anal. Vidhya.* 2020:1–10. <https://medium.com/analytics-vidhya/introduction-to-the-gradient-boosting-algorithm-c25c653f826b>. [Accessed 25 July 2021].
- Brownlee J. Extreme gradient boosting (XGBoost) ensemble in Python. *Mach. Learn. Mastery.* 2020:1–23. <https://machinelearningmastery.com/extreme-gradient-boosting-ensemble-in-python/>. [Accessed 25 July 2021].
- Chen T, Guestrin C. XGBoost. In: Proc. 22nd ACM SIGKDD int. Conf. Knowl. Discov. Data min. New York, NY, USA: ACM; 2016. p. 785–94. <https://doi.org/10.1145/29399672.2939785>.
- Bramer M. Ensemble classification. https://doi.org/10.1007/978-1-4471-7493-6_14; 2020. 209–220.
- Opitz D. Classifier ensembles. 1999. <http://www.cs.cmu.edu/afs/cs/project/jair/pub/volume1/opitz99a-html/node2.html>. [Accessed 25 July 2021].

- [40] Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inform. Decis. Mak.* 2019;19:281. <https://doi.org/10.1186/s12911-019-1004-8>.
- [41] Pan Y, Liu D, Deng L. Accurate prediction of functional effects for variants by combining gradient tree boosting with optimal neighborhood properties. *PLoS One* 2017;12:e0179314. <https://doi.org/10.1371/journal.pone.0179314>.
- [42] Ishaq A, Sadiq S, Umer M, Ullah S, Mirjalili S, Rupapara V, Nappi M. Improving the prediction of heart failure patients' survival using SMOTE and effective data mining techniques. *IEEE Access* 2021;9:39707–16. <https://doi.org/10.1109/ACCESS.2021.3064084>.
- [43] Smitha N, Bharath R. Performance comparison of machine learning classifiers for fake news detection. In: 2020 second int. Conf. Inven. Res. Comput. Appl. IEEE; 2020. p. 696–700. <https://doi.org/10.1109/ICIRCA48905.2020.9183072>.
- [44] Hsu C-N, Liu C-L, Tain Y-L, Kuo C-Y, Lin Y-C. Machine learning model for risk prediction of community-acquired acute kidney injury hospitalization from electronic health records: development and validation study. *J Med Internet Res* 2020;22:e16903. <https://doi.org/10.2196/16903>.
- [45] Wei L, Xing P, Zeng J, Chen J, Su R, Guo F. Improved prediction of protein–protein interactions using novel negative samples, features, and an ensemble classifier. *Artif Intell Med* 2017;83:67–74. <https://doi.org/10.1016/j.artmed.2017.03.001>.
- [46] Beretta L, Santaniello A. Nearest neighbor imputation algorithms: a critical evaluation. *BMC Med. Inform. Decis. Mak.* 2016;16:74. <https://doi.org/10.1186/s12911-016-0318-z>.
- [47] Rahman MM, Davis DN. Addressing the class imbalance problem in medical datasets. *Int. J. Mach. Learn. Comput.* 2013;3:224–8. <https://doi.org/10.7763/IJMLC.2013.V3.307>.