

Predicting ICU Admission for COVID-19 Patients in Saudi Arabia: A Comparative Study of AdaBoost and Bagging Methods

Hamza Ghandorh¹, Mohammad Zubair Khan², Mehshan Ahmed Khan³, Yousef M. Alsafayan⁴,
Ahmed A. Alahmari⁵, Anas A. Khan⁶

College of Computer Science and Engineering, Taibah University Medina, Saudi Arabia^{1,2}
Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong Waurm Ponds, Australia³
General Directorate of Humanitarian Response, Saudi Red Crescent Authority, Riyadh, Saudi Arabia⁴
World Health Organization, EM/RGO/WHE/CPI, Egypt⁵
Global Center for Mass Gatherings, Medicine Ministry of Health, Riyadh, Saudi Arabia⁶
College of Medicine, King Saud University, Riyadh, Saudi Arabia⁶

Abstract—COVID-19's high fatality rate and accurately determining the mortality rate within a particular geographic region continue to be significant concerns. In this study, the authors investigated and assessed the performance of two advanced machine learning approaches, Adaptive Boosting (AdaBoost) and Bootstrap Aggregation (Bagging), as strong predictors of COVID-19-related intensive care unit (ICU) admissions within Saudi Arabia. These models may help Saudi health-care organizations determine who is at a higher risk of readmission, allowing for more targeted interventions and improved patient outcomes. The authors found AdaBoost-RF and Bagging-RF methods produced the most precise models, with accuracy rates of 97.4% and 97.2%, respectively. This work, like prior studies, illustrates the viability of developing, validating, and using machine learning (ML) prediction models to forecast ICU admission in COVID-19 cases. The ML models that have been developed have tremendous potential in the fight against COVID-19 in the health-care industry.

Keywords—COVID-19; adaptive boosting; bootstrap aggregation; prediction; ICU admission; Saudi Arabia; machine learning

I. INTRODUCTION

Since 2019, the coronavirus disease 2019 (COVID-19) pandemic has continued to spread globally [1]. The global impact of the COVID-19 pandemic has resulted in more than 170 million confirmed infections and 3.54 million fatalities, giving rise to significant public health and socioeconomic concerns. By the end of November 2021, more than two hundred million COVID-19 cases had been registered, more than five million deaths, and more than seven billion COVID-19 vaccines had been provided [2]. It was expected that between 0.02% and 0.82% of those infected with the virus would die [3]. Due to the high number of asymptomatic people, the death rate would increase.

Even though the World Health Organization (WHO) declared the epidemic to be over, low- and middle-income nations, such as Brazil or Ethiopia, are still experiencing the effects of the COVID-19 burden on health-care systems. With 1.5 million cases, Brazil saw around 12,000 deaths between January and September 2023 [4]. The pandemic has continued

to influence Ethiopias people and its economy, different from than earlier pandemics. This pandemic even briefly resulted in a global economic collapse and a near-total halt to social and economic activities [5]. Thus, not only acquiring knowledge and understanding the roots of the epidemic but also predicting its trajectory are of the utmost importance, especially in low- and middle-income nations.

The first instance of COVID-19 in Saudi Arabia was officially documented on March 2, 2020. Subsequently, as of September 12, 2020, a total of 325,050 confirmed cases were identified, out of which 301,836 individuals successfully recuperated, whereas 4,240 infections resulted in fatalities [6, 7, 8]. The COVID-19 pandemic presents a major public health risk to large-scale events, such as the Hajj, which draws an annual attendance of 2.5 million Muslim pilgrims from 150 countries, with foreign pilgrims accounting for 75% of the overall population [9].

In March 2020, Saudi Arabia witnessed a notable hospitalization rate of 71.6% among those who tested positive for COVID-19, accompanied by a fatality rate of 0.65% [10]. The aforementioned elevated rate has the potential to intensify further the economic strain caused by viral respiratory infections, leading to an approximate direct medical expenditure of SAR 48,551.36 (USD 12,947.03) per patient.

According to the official Saudi COVID-19 monitoring dashboard [11], an overview of confirmed critical case counts was described between May 2020 and May 2022. Within 24 months, the count of confirmed cases rose from less than 100 cases to around 700,000 cases, where critical cases ranged from less than 100 to 2,000 cases. Currently, it is estimated that there are 3,900 active cases.

Saudi Arabia encountered several public health issues during the COVID-19 pandemic, encompassing areas such as knowledge deficiencies, attitudes and behaviors, psychological implications, vaccine hesitancy, management of religious mass gatherings, and the application of travel limitations [12]. The aforementioned challenges exhibited a distinctiveness exclusive to Saudi Arabia, derived from its religious and cultural context.

The collective efforts to understand the COVID-19 pandemic have resulted in the creation of a large number of datasets. Yet the volume, risk factors, and complexity make predicting COVID-19 infection complicated. Factors such as demographics, medical history, symptoms, and real-time data updates add to the COVID-19 complications [13]. This level of complexity necessitates advanced computing methods and a significant amount of processing time. Thus, there is a need for innovative techniques, namely machine learning (ML), that can be utilized to investigate and forecast the severity of asymptomatic carriers, as well as the prospective death rate from recorded illnesses.

This study aims to investigate two advanced ML methods for predicting COVID-19 patients admission to the intensive care unit (ICU). The proposed model is built with well-known classification methods¹, namely the Adaptive Boosting (AdaBoost) and Bootstrap Aggregation (Bagging) methods. The proposed model was utilized on a private governmental dataset using the clinical COVID-19 characteristics of Saudi Arabian residents.

The authors have developed two hypotheses for this study: (1) The AdaBoost technique is expected to be superior to the Bagging method in terms of accuracy and precision in predicting ICU admission for COVID-19 patients; (2) the selection of features and their relevance to ICU admission prediction will influence the performance of the AdaBoost and Bagging algorithms. To our knowledge, there are limited studies investigating the application of advanced ML methods, namely AdaBoost and Bagging, on a local Saudi Arabia dataset. The contribution of this work is thus twofold:

- 1) To investigate the performance of two advanced ML methods, AdaBoost and Bagging, as predictors of COVID-19 related ICU admission.
- 2) To evaluate the proposed model, the authors applied it to a private government dataset in terms of size and number of clinical screening features.
- 3) To recognize the significance of feature selection in improving the efficiency of the applied ML classifiers under consideration.
- 4) To motivate the scientific community to employ different ML classifiers for improving ICU admission prediction in similar geographical regions.

The rest of the article is organized as follows: Section II demonstrates a few related works in which the AdaBoost and Bagging models were utilized for COVID-19 patient classifications and predictions. Section III describes the utilized dataset and proposed model, along with the applied evaluation schema. Section IV depicts a comparative analysis of the performance of the proposed model. Section V sheds light upon a further discussion, and Section VI provides a conclusion and outlines future directions.

II. RELATED WORK

This section demonstrates some recent work applying the AdaBoost and Bagging methods as strong predictors for

¹Methods, learners, classifiers, or techniques will be used interchangeably throughout the paper.

COVID-19 infection within a variety of geographical locations and clinical datasets.

Soui et al. [14] conducted a comparative study of various ML methods to identify an effective model for distinguishing COVID-19 cases from suspects. They applied their proposed model to two datasets: a dataset from the Wolfram Data Research Repository with 1,495 patients and a dataset from an external source with 99,232 samples. Many algorithms including forward floating selection, and non-dominated sorting genetic algorithm II— were utilized to choose the optimal subset of features. To thoroughly classify COVID-19 suspects, the authors applied various machine learning algorithms: MLP,² SVM,³ LR,⁴ DT,⁵ GB,⁶ RF,⁷ XGBoost,⁸ and AdaBoost, and they measured their performance. After SMOTE was applied to the datasets, the authors indicated RF outperformed all other classifiers in the first and second datasets, with an accuracy of 81.51% and 92.88%, respectively.

Darici [15] performed a comparative analysis between the AdaBoost-CNN and AdaBoost-ResNet-152 methods to not only autonomously extract image features from X-ray chest COVID-19 patients but also classify these images. The authors used datasets containing 2,905 photos from various sources, with an unequal distribution across classes, and the SMOTE was used to balance the number of photos in each class. Overall, 1024 features were fed into the AdaBoost method, and the authors chose SVM as the weak classifier. For automatic feature extraction, AdaBoost-CNN outperforms AdaBoost-ResNet-152. The best average accuracy result for AdaBoost-CNN model was 94.5%, whereas it was 89% for the AdaBoost-ResNet-152 model.

Mary et al. [16] aimed to predict COVID-19 severity by identifying and classifying COVID-19 cells in a chest X-ray dataset. The used dataset contains 10,000 images of chest X-rays, as well as CSV files, which were located at the Kaggle site. To extract and segment COVID-19 cells from the dataset, the authors proposed a Vulture-Based Adaboost-Feedforward Neural (VbAFN) method. To improve segmentation and classification accuracy, the authors employed a variety of optimization strategies, including CNN with Fuzzy, Fusion schemes, the BO-F methodology, CNN with VGG16, and Hidden Markov with U-net Architecture. When compared to previous studies, the authors reported the VbAFN scheme obtained an accuracy of 99%, with an error rate of 0.02.

Mazloumi et al. [17] investigated the use of blood samples, age, gender, and ICU admission to predict patient survival or death features in Wuhan, China. The authors examined various ML techniques from 306 infected Tangji Hospital patients. The SMOTE method for nominal and continuous variables was used to balance the dataset. The authors reported that DT, AdaBoost, RF, KNN,⁹ and SVM outperformed other ML methods in predicting COVID-19 patient survival or death, where DT achieved accuracy of 91.6% and AdaBoost achieved

²Multilayer Perceptron method

³Support Vector Machine method

⁴Linear Regression method

⁵Decision Tree method

⁶Gradient Boosting method

⁷Random Forest method

⁸Extreme Gradient Boosting method

⁹K-Nearest Neighbors method

accuracy of 91.3%. Additionally, the authors reported that age, LD, and leukocytosis features were the most critical criteria in measuring and analyzing COVID-19 survival.

Sharma et al. [18] aimed to effectively forecast the spread of COVID-19 in India using multivariate time series data. The authors employed two worldwide datasets from Kaggle and Indiastathealth sites. The datasets were aggregated between January 2020 and August 2021, with various features considered, such as the number of cases by date, confirmed cases by date, confirmed deaths, vaccination, and policy responses. To extract related COVID-19 features, an adaptive gradient LSTM model (AGLSTM) was used. RNN,¹⁰ LSTM,¹¹ LASSO regression, AdaBoost, LGB,¹² and KNN models are used as classification methods. The authors validated their model in two ways: local Indian case studies and data fusion and transfer-learning techniques. As a result, AGLSTM outperforms other ML methods, with an accuracy of 99.81% with little training time.

Solayman et al. [19] proposed an automated ML-driven COVID-19 identification tool to determine whether or not users were infected with COVID-19. Through answering symptom-related clinical questions, the tool filled the gap in earlier research by combining automated detection techniques with rapid prediction. The authors employed ML methods, including LR, RF, DT, KNN, SVM, AdaBoost, XGB, ANN,¹³ CNN, and LSTM to train and assess the proposed tool. The authors used a Middle Eastern-based open-source dataset with around two million patients with a focus on their patient information, symptoms, and COVID-19 test results. After dropping null values and feature engineering, the SMOTE approach was used to preprocess the dataset. As a result, other ML models were outperformed by the hybrid CNN-LSTM methods, with an accuracy rate of 96.34% and 85.49% after the use of SMOTE and no SMOTE techniques, respectively.

To extract and describe the chest CT characteristics of COVID-19 patients, Li et al. [20] proposed a COVID-19 early warning system, and it functioned upon various ML methods, including the XGBoost, LR, MLP, RF, and AdaBoost methods. The system utilized an aggregated adult CT imaging dataset from COVID-19 patients from three medical centers in Beijing, Wuhan, and Nanchang. The dataset included a variety of features, such as imaging ratings, clinical characteristics, and biomarker levels. With an accuracy of 82% and 84% (mean), the LR and XGBoost methods predicted the real probability of severe/critical COVID-19, respectively. The authors reported that general clinical markers such as blood oxygen saturation, age, and total lung involvement were found to be important predictors of critical COVID-19 patients.

Abegaz and Etikan [5] performed a case study to predict COVID-19 mortality in Ethiopia. The authors compared AdaBoost against weak classifiers including KNN, ANN, and SVM. The dataset used included two years of COVID-19 patient records that came from OurWorldInData and the John Hopkins University warehouse. The used datasets focused on a set of five features: the daily number of COVID-19 deaths,

TABLE I. DESCRIPTION OF THE 2019 CORONAVIRUS DISEASE (COVID-19) POSITIVE PATIENTS (CDPP) DATASET IN TERMS OF CATEGORICAL FEATURES

Feature	Description (Values)
ClassificationGroup	Epidemiological criteria (Case, Contact)
Outcome or Outcome_Modified	Admission outcome (Recovery No ICU, Recovery with history of ICU, Death)
age_65	Age above 65? (Y, N)
Gender	Patient gender (M, F)
Nationality	Nationality (Sa, Eg, Sd, etc.)
SYMPTOMATIC	Symptomatic? (Y, N)
HCW_totalpop	Occupation (Medical staff, military, others) (0 - 2)
comorbidity	Any comorbidity? (Y, N)
comorbidity or any_comorbidity	Any comorbidity? (Y, N)
morethan2comorbidities	Two or more comorbidities? (Y, N)
DM1	Diabetes? (Y, N)
HTN1	Hypertension? (Y, N)
CRF1	Chronic kidney disease? (Y, N)
cardiac1	Heart diseases? (Y, N)
asthma1	Asthma and chronic lung disease? (Y, N)
cancer _immunodeficiency1	Immunodeficiency? (Y, N)
C_lungdisease	Lung disease? (Y, N)
Smoking	Smoker? (Y, N)
Fever_PRESENT	Fever? (Y, N)
Cough_PRESENT	Cough? (Y, N)
SoreThroat_PRESENT	Sore throat? (Y, N)
RunnyNose_PRESENT	Runny nose? (Y, N)
Headacheonset	Headache frequency (0 - 1_3)
Myalgiaonset	Myalgia frequency (0 - 1_5)
GI symptomsonset	GI symptoms frequency (0 - 1_4)
SEVERITY	Patients conditions (0, 3)

daily new cases, bed capacity, mask use, and pneumonia status. The authors indicated the best coefficient determination for AdaBoost, KNN, ANN, and SVM were 94.2%, 86.2%, 86.3%, and 71.7%, respectively.

de Holanda et al. [4] aimed to forecast hospitalization and mortality outcomes for COVID-19 patients in Brazil. The purpose of the study is to support medical professionals and administrators in their decision-making. The demographics, medical history, immunization records, symptoms, and underlying illnesses of the patients were examined by the researchers using data from a publicly available dataset located at the OpenDataSus site. XGBoost, LR, AdaBoost, RF, SVM, KNN, DT, and NB,¹⁴ are among the 14 ML techniques applied to the dataset. In terms of hospitalization risk prediction, the gradient-boosting model was better than the others, with an accuracy rate of 71% and an AUC of 0.75.

III. MATERIAL AND METHODS

This section describes the utilized dataset and proposed model along with the applied evaluation schema.

¹⁰Recurrent Neural Network method

¹¹Long Short-Term Memory Networks method

¹²Light Gradient Boosting method

¹³Artificial Neural Networks method

¹⁴Naïve Bayes classifier

TABLE II. DESCRIPTION OF THE 2019 CORONAVIRUS DISEASE (COVID-19) POSITIVE PATIENTS (CDPP) DATASET IN TERMS OF NUMERICAL FEATURES

Feature	Description
LOSdays	Length of stay in days
dayofExposureifknown	Exposure period
Incubation	Incubation period
HEART_RATE	Heart beats per minute
RESPIRATORY	Number of breaths per minute
SBP	Systolic blood pressure
DBP	Diastolic blood pressure
SATURATION	Oxygen level
WHITE_CELLS	White cell count
CREATININE	Creatine phosphate count
LYMPHOCYTES	Lymphocyte count of less than 1,500 per 1 Mio. m ³
PLATELET	Platelet counts
NEUTROPHILS	White blood cell type level
BLOOD	Blood pressure

TABLE III. PROFILE INFORMATION OF THE 2019 CORONAVIRUS DISEASE (COVID-19) POSITIVE PATIENTS (CDPP) DATASET

Feature	Mean	Std. dev.	Feature values (Min—Max)
LOSdays	7.71	9.59	(0—60)
comorbidities	0.47	0.87	(0—5)
age_computed	36.59	15.57	(0—84)
dayofExposureifknown	8.5	6.3	(1—30)
Incubation	7.21	6.12	(1—30)
Temperature	37.05	2.37	(0—39.1)
HEART_RATE	89.82	14.19	(63—125)
RESPIRATORY	19.87	2.48	(14—30)
SBP	125.12	17.22	(60—188)
DBP	75.03	10.46	(57—116)
SATURATION	96.68	3.63	(69—100)
WHITE_CELLS	6.53	3.96	(0—17)
CREATININE	64.55	43.63	(0—145)
LYMPHOCYTES	24.43	13	(6.3—58.1)
PLATELET	246.52	92.53	(107—572)
NEUTROPHILS	48.68	32.52	(1.37—93.4)
BLOOD	12.5	16.3	(2—68)

A. Dataset

As an outcome of Saudi nationwide quantitative study of RT-PCR¹⁵ tests, a private COVID-19 Positive Patients (CDPP) dataset was aggregated. The CDPP dataset was curated under several Saudi authorities, including the Global Center for Mass Gatherings Medicine, the Saudi National Health Laboratory, and the Saudi Health Electronic Surveillance Network¹⁶ [21, 22]. The Saudi authorities employed local electronic health systems to facilitate essential indicators for health-care facility readiness and epidemiological surveillance. These indicators encompassed various aspects such as the health staff dashboard for isolation hospitals, reports on blood samples and sample carrier shipments, the supply dashboard, COVID-19 mortality reports, workforce information, and the blood bank dashboard.

The dataset comprised 639 records, with 44 features that

included clinical and demographic information about symptomatic and asymptomatic patients. There are three types of variables in the dataset: Boolean (15), categorical (15), and numerical (18). Table I, Table II, and Table III demonstrate an overview of the datasets features and its profile, respectively.

Null values were observed in the dataset. Some features do not have any null values, including ClassificationGroup, Outcome or Outcome_Modified, HCW_totalpop, Any_comorbidity, Morethan2comorbidities, DM1, HTN1, CRF1, Cardiac1, Asthma1, Cancer_immunodeficiency1, C_lungdisease, Age_65, Gender, and SEVERITY. Other features included null values with less than 70%, including LOSdays (1.3%), Smoking (0.2%), Nationality (5.6%), SYMPTOMATIC (57.7%), Fever_PRESENT (38%), Cough_PRESENT (65.1%), Headacheonset (21.6%), Myalgiaonset (21.6%), GIsymptomsonset (21.6%), DayofExposureifknown (64.9%), Incubation (65.1%), and SATURATION (59.9%). The remaining features included null values with more than 70% including SoreThroat_PRESENT (77.6%), RunnyNose_PRESENT (85%), HEART_RATE (82.2%), RESPIRATORY (83.6%), SBP (82.3%), DBP (82.2%), WHITE_CELLS (94.7%), CREATININE (96.9%), LYMPHOCYTES (96.4%), PLATELET (94.8%), NEUTROPHILS (95.6%), and BLOOD (97.2%).

To visualize the distribution and intensity of data points in the CDPP dataset, Fig. 1 presents a heatmap using the Spearman correlation coefficients. The observed values of DM1, HTN1, CRF1, cardiac1, and asthmal were correlated with comorbidity. No other significant correlations were observed.

B. Background

Boosting models were developed for handling classification difficulties before being used for regression problems. According to [23], boosting methods focus on a small number of weak classifiers (those that predict just marginally better than random) that are merged (i.e., boosted) to create an ensemble classifier with a lower generalized misclassification error rate [24]. Ensemble learning uses the same learning algorithm to train many predictive models, enhancing their accuracy and reliability over single-model instances. It frequently helps modelers understand the models fragility or reliance on specific data points, which can aid in determining which fresh data sets should be gathered and with what priority. Ensemble learners often utilized Bagging, Boosting, and Stacking [25].

1) *AdaBoost*: The Adaptive Boosting (AdaBoost) method creates a series of weak classifiers, with the best classifier picked based on the current sample weights after each iteration. Fig. 2 contains an overview of the AdaBoost method.

Samples classified inaccurately in the k^{th} iteration receive a higher weight in the $(k + 1)^{st}$ iteration, whereas samples classified correctly receive a lower weight in the subsequent iteration. Difficult data are given more weight until the classifier finds a model that correctly classifies them. As a result, each iteration of the classifier must learn a new aspect of the data, focusing on regions containing complex samples. For each iteration, a stage weight is calculated based on the error rate of the iteration [24]. The final hypothesis h_f is a weighted majority vote of the hypotheses of weak learner t where t is

¹⁵Reverse Transcription Polymerase Chain Reaction

¹⁶Site: <https://hesn.moh.gov.sa/webportal/>

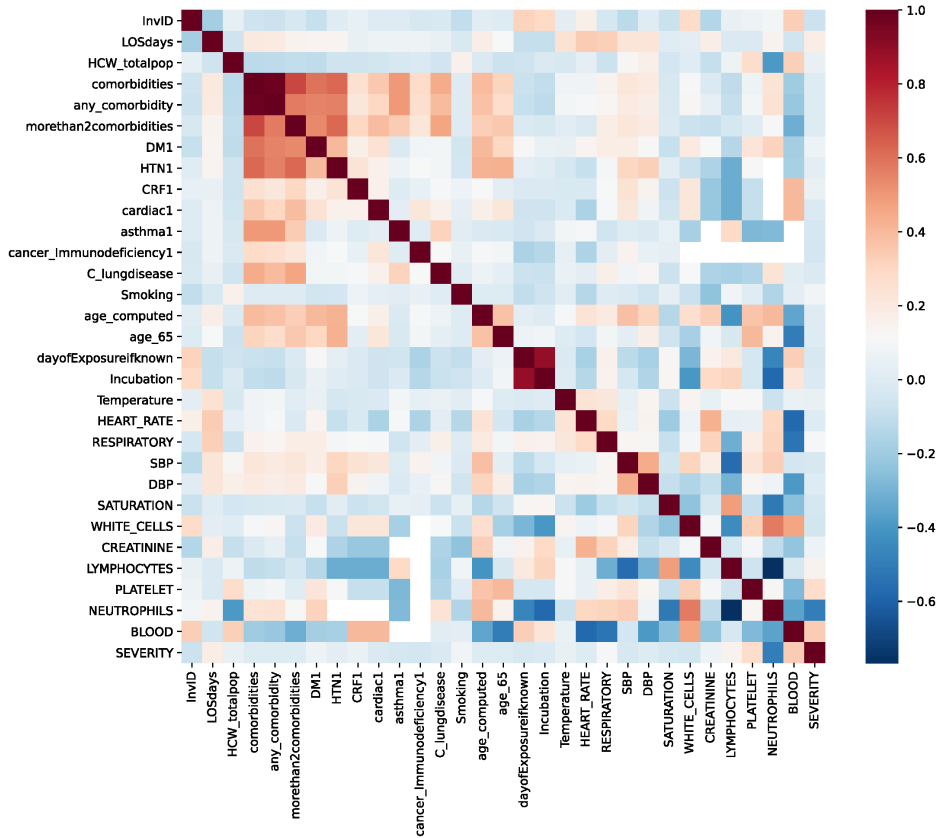


Fig. 1. Heatmap of correlation coefficient of the 2019 coronavirus disease (COVID-19) positive patients (CDPP) dataset features using the spearman coefficient.

$$h_f(i) = \begin{cases} 1, & \sum_{t=1}^T (\log \frac{1}{\beta_t}) h_t(i) \geq \frac{1}{2} \sum_{t=1}^T \log \frac{1}{\beta_t} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

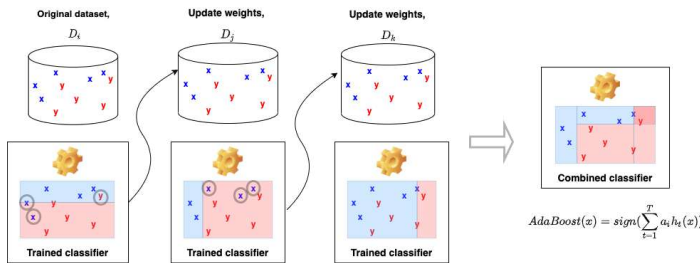


Fig. 2. An Adaptive Boosting (AdaBoost) model assigns weights to weak and solid classifiers and the distributions samples in a way that classifiers are driven to focus on complicated data point-related observations.

There are many advantages of the AdaBoost method, including: (1) fast, simple, and easy classification method; to create; (2) it has little to no configurable parameters; (3) it does not require prior knowledge of the weak learners; (4) it shows the model to be combined with other methods for finding weak hypotheses; and (5) it can create valid weak hypotheses consistently when enough data are provided. Overall, it provides a set of theoretical guarantees. However, when there are inadequate data, weak hypotheses are highly complicated, or weak hypotheses are too fragile. Thus the model may underperform [27].

2) *Bagging*: The Bootstrap Aggregation (Bagging) method was one of the first ensemble approaches produced [28]. Fig. 3 depicts an overview of the Bagging model.

the weight assigned to a hypothesis h_t calculated through Eq. 1 [26]:

Bagging is a general approach to constructing an ensemble model that employs bootstrapping in conjunction with regres-

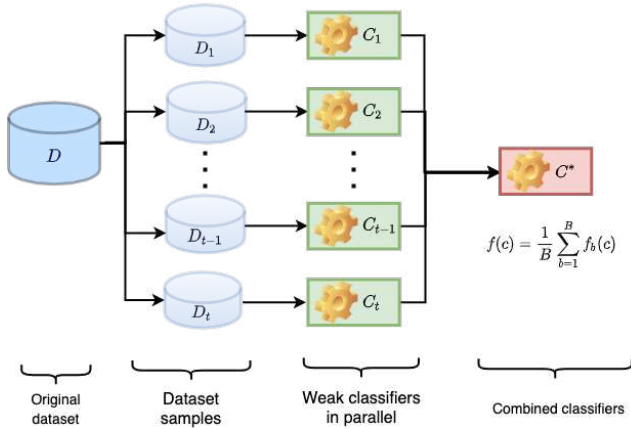


Fig. 3. A Bootstrap Aggregation (Bagging) model predicts a new data sample d , and the forecasts are averaged to yield the final model.

sion models. By combining multiple models (i.e., learners) trained on various subsamples of the same data set, Bagging decreases the variance of predictions. The Bagging method creates several data sets from the original data, trains various classifiers on each data set, then integrates these models to give a single response value [25]. Each model in the ensemble learners is then used to predict a new sample, and the forecasts are averaged to yield the prediction of the bagged model [24]. In the Bagging method, it consists of the following simple steps as follows:

```

for  $l \in$  learners
  leftmargin=0.8cm
    generate a bootstrap sample of  $d \in$  dataset
    train an unpruned tree  $l$  on the  $d$ 
end

```

Bagging models provide various advantages over non-Bagged models. First, through its aggregation process, Bagging effectively minimizes the variance of a forecast. Another benefit of Bagging models is they can produce their internal estimate of predicted performance that matches well with either cross-validation or test set estimates. Although Bagging increases predictive performance for unstable models in most cases, a Bagged model is substantially less interpretable than a non-Bagged model [24].

C. Proposed Model

The current rapid and exponential increase in the number of infected patients has necessitated an accurate estimation of suitable ML models' potential outcomes. In this study, the authors investigated the ability to predict the severity of the asymptomatic carriers and the possible death rates using two advanced ML methods, AdaBoost and Bagging, within Saudi Arabia. The model utilized the 2019 Coronavirus Disease (COVID-19) Positive Patients (CDPP) dataset (refer to Subsection III-A). The CDPP dataset is arranged utilizing these classifiers and weak classifiers (DT, RF, and SVM) under the test method of 10-fold cross-validation.

Base learners results are assessed by comparing the results obtained from popular classifiers: AdaBoost-DT, AdaBoost-

Algorithm 1 AdaBoost method pseudocode.

Input: Dataset $D = \{(a_1, c_1), (a_2, c_2), \dots, (a_N, c_N)\}$, Base Learner L , and number of learning iteration T

Output: $H(a) = \text{sign} \sum_{t=1}^T \alpha_t h_t(a)$

```

1 Initialize equal weights to all training samples  $w_i = \frac{1}{N}$ ,  $i = 1, 2, 3, \dots, N$ 
for  $t = 1$  to  $T$  do
2   (a) Train a base learner  $h_t$  from  $D$  using  $D_t$  to training sample using  $w_i$ 
      $h_t = L(D, D_t)$ 
     (b) Compute error of  $h_t$  as
     
$$\text{err}_t = \frac{\sum_{i=1}^N w_i I(h_t(a_i) \neq c_i)}{\sum_{i=1}^N w_i}$$

     (c) Compute the weight of  $h_t$  as
     
$$\alpha_t = \log\left(\frac{1 - \text{err}_t}{\text{err}_t}\right)$$

     (d) Set  $w_i \leftarrow w_i \cdot \exp[\alpha_t I(h_t(a_i) \neq c_i)]$ 
3 end

```

Algorithm 2 Bagging method pseudocode.

Input: Base Learner L , Bootstrap Samples X_l , $X_l = \{x^t, r^t\}_{t=1}^N$

Output: Voted Best Base Learner $g^*(x)$

```

4 Generate  $l = 1, 2, \dots, L$  with  $|X_l| = N$  by sampling  $\frac{1}{N}$  with replacement
   Train  $L$  for  $X_l \Rightarrow g_l(x)$ 
   Use voting (average or median with regression) of multiple base learners
   
$$g_{\text{bag}}(x) = \frac{1}{L} \sum g_l(x)$$


```

RF, Bagging-SVM, Bagging-DT, Bagging-RF, and AdaBoost-SVM. An overview of the AdaBoost method is presented in Pseudocode 1, and an overview of the Bagging method is presented in Pseudocode 2. The execution time for all the classifiers was not more than 0.05 seconds. Different execution measures are utilized to assess the error rate and accuracy of chosen classifiers. The models performance has been evaluated in terms of accuracy based on the confusion matrix [29].

The study utilized ensemble model techniques, AdaBoost and Bagging, in combination to distinguish between COVID-19 and common viral features. These methods successfully integrated multiple features, leading to a high level of accuracy while reducing execution and training times. Additionally, these methods are known to be less biased compared to traditional ML methods.

The proposed model architecture consists of five main phases, namely data acquisition, preprocessing, feature extraction, feature selection, and classification. Fig. 4 presents an overview of the proposed model.

1) *Dataset preprocessing:* The Synthetic Minority Over-Sampling Technique (SMOTE) [30], which oversamples the synthetics in the minority class and duplicates the same entities without adding new information, was applied to the dataset to correct this imbalance. During training, the Outcome feature served as both an independent and dependent variable. There were 563 patients in each class after applying SMOTE.

Three classes made up the dataset: Death, Recovery No

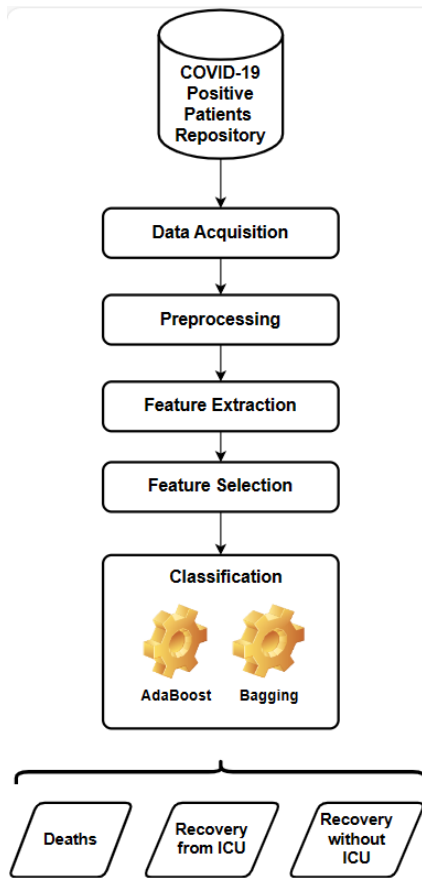


Fig. 4. Architecture of the proposed model to investigate and forecast the incidence and the potential death risk of the asymptomatic carriers.

ICU, and Active ICU or Recovery with History of ICU. The patient classifications were maintained as follows: 563 to the Recovery No ICU class, 65 to the Active ICU or Recovery with History of ICU class, and 11 to the Death class. Boolean variables with values of Y and N, respectively, include Fever_PRESENT, SoreThroat, and RunnyNose_PRESENT. Y is substituted with 1 and N with 0 notations to unify the code system. An individual's age is represented by two entries, 1 and 2, in the field Age_65, for instance. 0 is used to replace empty entries in the event the symptom is absent. Binary code systems are used by other boolean variables such as Any_comorbidity, DM1, HTN1, CRF1, cardiac1, asthma, Cancer_immunodeficiency, and C_lungdisease. The variables median was substituted for missing values in numerical features such as Smoking, LOSdays, Comorbidities, and more. There were also missing values for categorical features such as ClassificationGroup, Gender, Nationality, and Outcome. The scales of other numerical features, such as Headacheonset, GIsymptomsonset, and Myalgiaonset, varied and did not all contribute equally to the models fit. To scale values on a single scale, the MinMax scaler was applied to each variable. The median of each variable was used in place of the categorical features, such as Myalgiaonset, GIsymptomsonset, and Headacheonset to represent an individual read of a symptom every two days. Other factors, namely, client name and InvID were removed because these variables had no role in the

TABLE IV. A CONFUSION MATRIX SAMPLE

		Actual Class	
		P	N
Predicted Class	P	T_P	F_P
	N	F_N	T_N

TABLE V. DESCRIPTION OF THE PERFORMANCE EVALUATION CRITERIA

Criteria	Representation (%)
Accuracy	$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} * 100$
Precision	$Precision = \frac{T_P}{T_P + F_P} * 100$
AUC or Recall	$AUC = \frac{T_P}{T_P + F_N} * 100$
F-score	$F\ score = 2 * \frac{Precision * Recall}{Precision + Recall} * 100$

classification stage.

2) *Experimental setup*: Both ML methods, AdaBoost and Bagging, were implemented using Python 3.7 in a web-based on-demand service platform. The platform is referred to as Google Colaboratory (Colab)¹⁷, and is designed for ML tasks and data analysis. The necessary libraries, including Pandas, sklearn, NumPy, Seaborn, SciPy, Keras, ELI5, and TensorFlow, were included. The execution utilized a CPU with Intel Xeon 2.20 GHz, 1 GB of RAM, and 69 GB of storage on Google Drive.

3) *Performance evaluation metrics*: Accuracy is critical when predicting ICU admission based on various COVID-19 symptoms and patients clinical features. To investigate the accuracy of the suggested approach, the authors calculated the confusion matrix parameters. Table IV shows a sample of the confusion matrix.

T_P , T_N , F_P , and F_N are true positive, true negative, false positive, and false negative values, respectively. T_P characterizes a data point that was anticipated to be in a selected class, and it was found in it. T_N describes a data point that was not anticipated to be in a selected class, and it was not found in it. F_P describes a data point anticipated to be in a selected class, but it was not found in it. F_N characterizes a data point that was not anticipated to be in a selected class, and it was found in it.

The most commonly used performance metrics for classification problems include accuracy, precision, AUC or recall, and F-score. Table V demonstrates the performance criteria. Vidyalya [31] stated accuracy is the simple ratio between the number of correctly classified points to the total number of points, whereas precision is the fraction of the correctly classified instances from the total classified instances. "F-score is the harmonic mean of precision and recall.". The area under the curve (AUC) or recall is the ratio of the true positive samples to the sum of the true positive and false negative samples.

Errors metrics are used as quantitative measures to demonstrate how predictive models perform. The authors calculated four error metrics: Root Mean Squared Error (RMSE), Relative Squared Error (RSE), Mean Absolute Error (MAE), and Relative Absolute Error (RAE). Table VI demonstrates the used

¹⁷Colab Site: <https://colab.research.google.com/>

TABLE VI. DESCRIPTION OF THE ERROR METRICS

Criteria Title	Representation (%)
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}} * 100$
Relative Squared Error (RSE)	$RSE = \frac{\sum_{i=1}^n (p_i - a_i)^2}{\sum_{i=1}^n (\bar{a} - a_i)^2} * 100$
Mean Absolute Error (MAE)	$MAE = \frac{\sum_{i=1}^n p_i - a_i }{n} * 100$
Relative Absolute Error (RAE)	$RAE = \frac{\sum_{i=1}^n p_i - a_i }{\sum_{i=1}^n \bar{a} - a_i } * 100$

TABLE VII. PERFORMANCE METRICS OF THE ML CLASSIFIERS, ADABOOST, AND BAGGING METHODS

Classifiers	Accuracy	Precision	AUC	F-Score
AdaBoost-SVM	0.310	0.096	0.500	0
DT	0.882	0.883	0.814	0.839
AdaBoost-DT	0.884	0.885	0.824	0.845
Bagging-DT	0.945	0.945	0.912	0.923
SVM	0.949	0.950	0.912	0.937
Bagging-SVM	0.959	0.958	0.936	0.946
RF	0.966	0.967	0.936	0.953
Bagging-RF	0.972	0.972	0.953	0.961
AdaBoost-RF	0.974	0.974	0.955	0.964

evaluation criteria in this work. RMSE calculates the mean magnitude of the error, where a is the actual target, and p is the predicted target. RSE compares the sum of the models errors to a simple predictor (using the average). MAE is the average of all absolute errors. The square root of the relative squared error is calculated by RAE.

IV. RESULTS

This section demonstrates the study results from the applied ML classifiers and the used evaluation schema.

A. Performance of ML Methods

Using the accuracy, precision, AUC, and f-score values, the authors evaluated the ML classifiers output. Table VII and Fig. 5 contain an overview of the performance of the ML classifiers was calculated.

Fig. 6 and Fig. 7 depict the accuracy and precision measures. Fig. 8 shows the AUC measure, and Fig. 9 depicts the f-score measure. From Fig. 6 and Fig. 8, several ML classifiers, namely, AdaBoost-RF, Bagging-RF, Bagging-SVM, Bagging-DT, and AdaBoost-DT, applied for classification on the CDPP dataset yielded accuracy of 97.4%, 97.2%, 95.9%, 94.5%, and 88.4%, respectively. AdaBoost-RF and Bagging-RF provided accuracy of 97.4% and 97.2% and AUC value of 95.5% and 95.3% greater than other variants of the CDPP dataset by alleviating data inconsistencies, respectively. In contrast, AdaBoost-SVM provided the worse accuracy of 30.9% and AUC value of 50% of the CDPP dataset.

Table VIII presents a comparison between related COVID-19 predictive models. These models attained an average accuracy of 90.3%, whereas our proposed model provided an accuracy of 97.4% when considering the used dataset.

TABLE VIII. COMPARISON OF THE PROPOSED MODEL WITH THE RELATED MODELS

ML Architecture	ML Methods	Accuracy (%)
de Holanda et al. [4]	Ensemble learners	71
Li et al. [20]	Ensemble learners	84
Mazloumi et al. [17]	Ensemble learners	91.6
Soui et al. [14]	Ensemble learners	92.88
Abegaz and Etikan [5]	AdaBoost	94.2
Darici [15]	AdaBoost	94.5
Solayman et al. [19]	CNN-LSTM	96.34
Ghandorh et al. [32]	Ensemble learners	97.93
Proposed approach	AdaBoost and Bagging	97.4

TABLE IX. P-VALUES OF THE ML CLASSIFIERS AGAINST THE ADABOOST AND BAGGING-METHODS

ML Methods	t-test	p-value
RF - AdaBoost-SVM	8.132	0.00004
SVM - AdaBoost-SVM	7.327	0.0001
DT - AdaBoost-SVM	5.895	0.0004
RF - AdaBoost-DT	3.215	0.0123
DT - AdaBoost-RF	-2.931	0.0190
DT - Bagging-RF	-2.609	0.0312
DT - Bagging-SVM	-2.152	0.0636
SVM - AdaBoost-RF	-2.041	0.0756
SVM - AdaBoost-DT	1.811	0.1078
RF - Bagging-SVM	1.646	0.1383
RF - Bagging-DT	1.595	0.1495
SVM - Bagging-RF	-1.569	0.1553
DT - Bagging-DT	-1.380	0.2050
SVM - Bagging-SVM	-0.785	0.4553
RF - Bagging-RF	0.368	0.7227
DT - AdaBoost-DT	0.172	0.8678
RF - AdaBoost-RF	0.000	0.9997
SVM - Bagging-DT	0.000	0.9998

Using the t-test while examining possible significant differences by the ML classifiers, the authors calculated p-values among a set of weak ML classifiers, namely the RF, DT, and SVM methods, against the AdaBoost and Bagging methods. Table IX shows the p-values obtained. By comparing the RF, SVM, and DT methods against the AdaBoost-SVM method, the p-value was less than 0.05, indicating a significant difference and resulting in the rejection of the null hypothesis. Comparing the RF method against the AdaBoost-DT method, with a p-value below 0.05, there is clear evidence of a significant difference, leading to a rejection of the null hypothesis. In addition, when the authors compared the DT method against the AdaBoost-RF and Bagging-RF methods, the authors found a significant difference, as the p-value fell below 0.05, thereby leading to the rejection of the null hypothesis.

By comparing the DT, RF, and SVM methods against the Bagging-SVM method, the p-value was larger than 0.05, indicating insufficient evidence to reject the null hypothesis and suggesting no significant difference. Comparing the RF, DT, and SVM methods against the Bagging-DT method, the p-value exceeding 0.05 implies a lack of substantial evidence to reject the null hypothesis, suggesting the absence of a significant difference. Moreover, when the authors compared

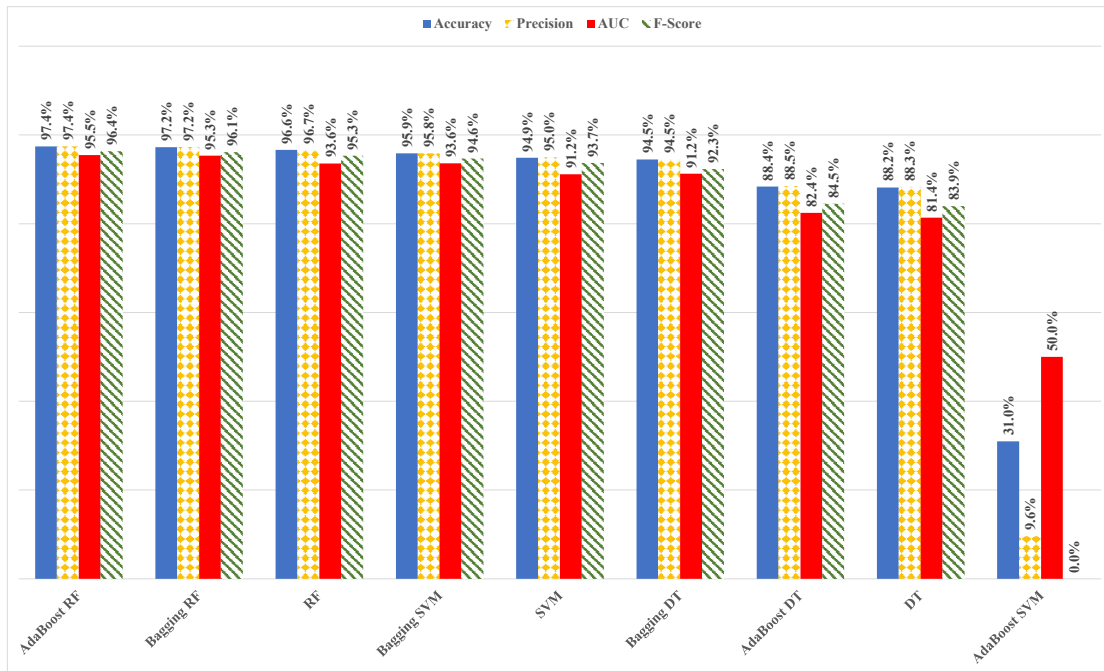


Fig. 5. Performance metrics of the ML classifiers, AdaBoost, and Bagging methods.

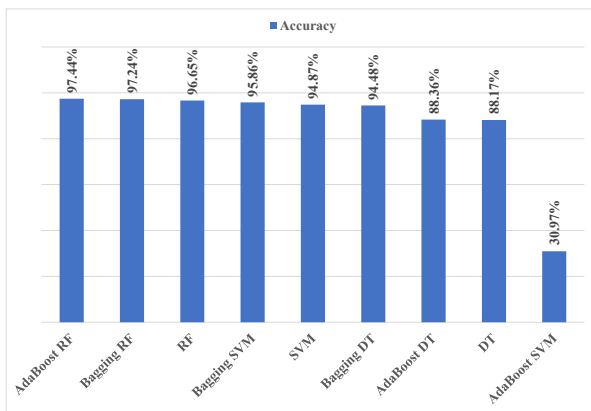


Fig. 6. Accuracy scores of the applied ML classifiers, AdaBoost and Bagging methods.

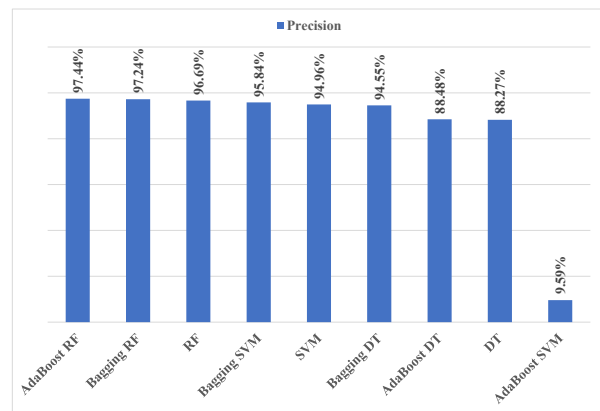


Fig. 7. Precision scores of the applied ML classifiers, AdaBoost and Bagging methods.

the SVM and RF methods against the AdaBoost-RF method, the authors found insufficient evidence exists to reject the null hypothesis based on the p-value being larger than 0.05, implying no significant difference.

B. Error Rates

The accuracy of the ML classifiers was ensured compared to the evaluation of classifier error rates by Table X or Fig. 10. In Fig. 11 – Fig. 14, the different error rates obtained for different classifiers are shown, respectively. Using the RMSE, MAE, RAE, and RSE rates, was calculated the error rate of each predictor.

The RMSE rate computes the median value of the absolute differences between observed and predicted values. MAE is a statistical method used to determine the average absolute

TABLE X. AN OVERVIEW OF DIFFERENT ERROR METRICS GIVEN BY THE ML CLASSIFIERS, ADA BOOST, AND BAGGING METHODS

Classifiers	RMSE	MAE	RAE	RSE
AdaBoost SVM	1.32	1.74	1.55	1.63
DT	0.65	0.43	0.33	0.81
AdaBoost DT	0.63	0.40	0.31	0.78
Bagging DT	0.44	0.20	0.15	0.55
SVM	0.42	0.18	0.14	0.52
Bagging SVM	0.39	0.15	0.12	0.48
RF	0.37	0.13	0.10	0.45
Bagging RF	0.33	0.11	0.08	0.41
AdaBoost RF	0.32	0.10	0.08	0.40

difference between expected and observed values, where each differences weight remains constant. RAE and RSP rates are

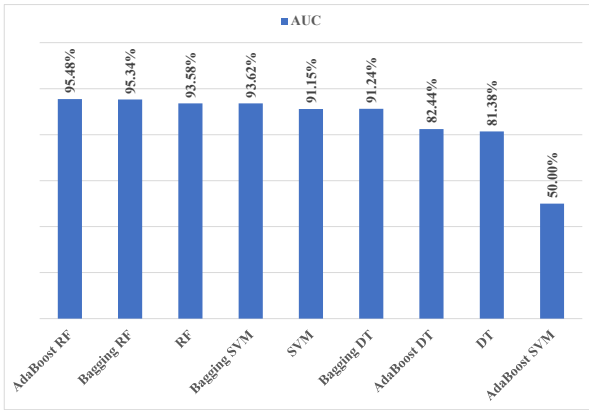


Fig. 8. AUC scores of the applied ML classifiers, AdaBoost and Bagging methods.

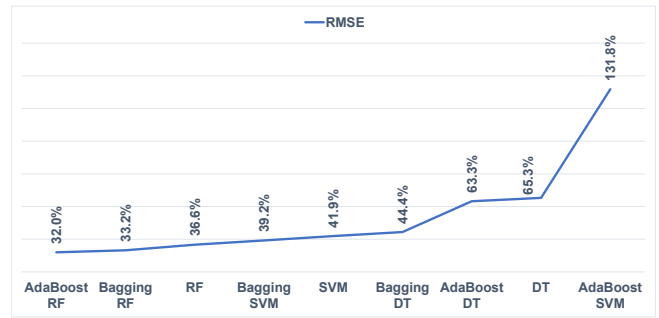


Fig. 11. Root Mean Squared Error (RMSE) rate of the applied ML classifiers, AdaBoost, and Bagging methods.

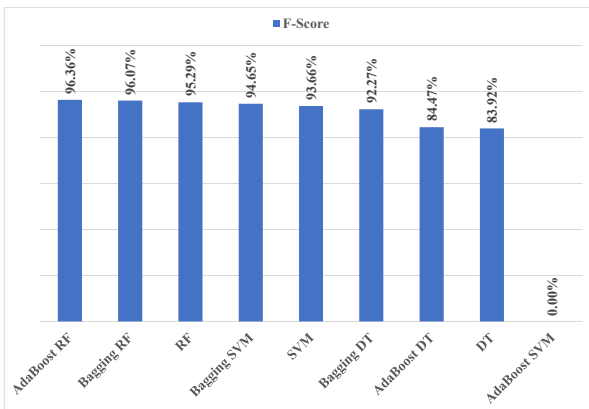


Fig. 9. F score scores of the applied ML classifiers and AdaBoost and Bagging methods.

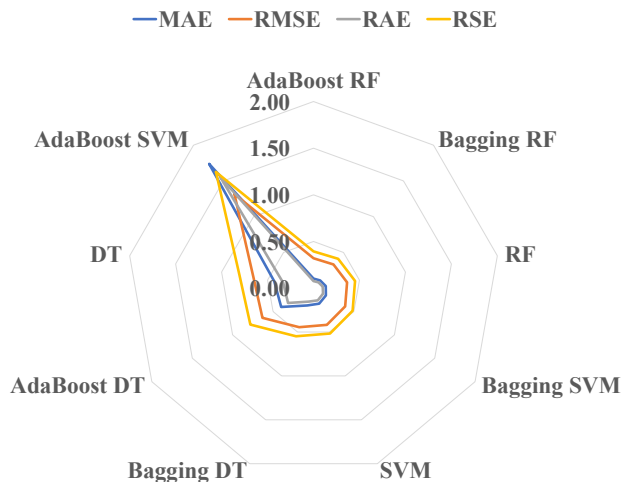


Fig. 10. An overview of different error metrics for the ML classifiers, AdaBoost, and Bagging Methods.

equivalent, determined by dividing MAE by simple classifier error received. A lower RAE value enhances prediction accu-

racy. The RSP rate offers more accurate results by normalizing data values obtained from simple classifiers, providing the squared error of forecasts relative to the mean of each data value.

From Fig. 11, the authors can see the AdaBoost-RF and Bagging-RF methods gave RMSE rates of 32%—33.2% with accurate classification of COVID-19 ICU recoveries, death, and recoveries, whereas the AdaBoost-SVM method indicated an RMSE rate of 131.8% with inaccurate classification. Other ML models, namely the Bagging-SVM, Bagging-DT, and AdaBoost-DT methods, fell between the RMSE rates of 39% and 63% of the remaining ML classifiers.

From Fig. 12, the AdaBoost-RF and Bagging-RF methods gave the lowest MAE value at 10% and 11%, whereas the AdaBoost-SVM method gave a 173.8% MAE value. Other ML models, namely the Bagging-SVM, Bagging-DT, and AdaBoost-DT methods, fell between 15.4% and 40% MAE values of the remaining ML classifiers.

Fig. 13 shows the AdaBoost-RF method attained 7.6%, a superior RAE value, whereas the AdaBoost-SVM method showed a worst RAE value of 155%. Other ML models, the Bagging-SVM, Bagging-DT, and AdaBoost-DT methods, fell between 11.8% and 31.5% RAE values of the remaining ML classifiers.

As shown in Fig. 14, the AdaBoost-RF and Bagging-RF methods held an RSE rates of 39.5% and 41%, whereas the AdaBoost-SVM method maintained RSE rate of 162.6%. Other ML models, such as the Bagging-SVM, Bagging-DT, and AdaBoost-DT methods, fell between RSE rates of 48.4% and 78.1% of the remaining ML classifiers.

C. Confusion Matrices

To thoroughly break down the proposed models performance and identify whether a ML model might be biased toward a specific class, the authors calculated the confusion matrices for the Bagging-DT, Bagging-SVM, Bagging-RF, and AdaBoost-RF methods. Fig. 15 - Fig. 18 exhibit the confusion matrices for Bagging-DT, Bagging-SVM, Bagging-RF, and AdaBoost-RF methods, respectively.

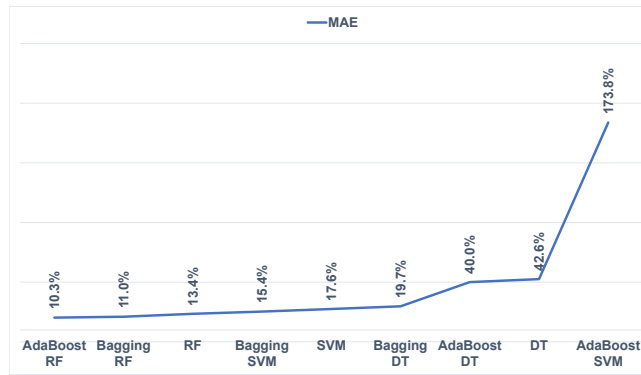


Fig. 12. Mean Absolute Error (MAE) rate of the applied ML classifiers, AdaBoost, and Bagging methods.

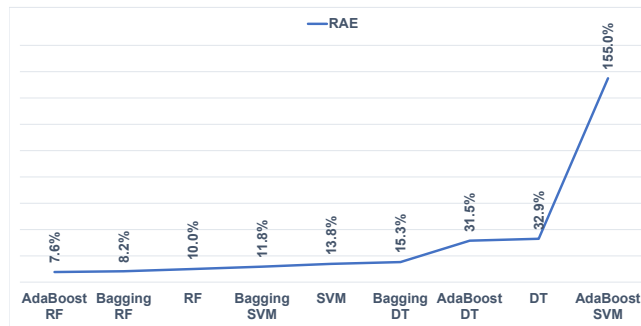


Fig. 13. Relative Absolute Error (RAE) rate of the applied ML classifiers, AdaBoost, and Bagging methods.

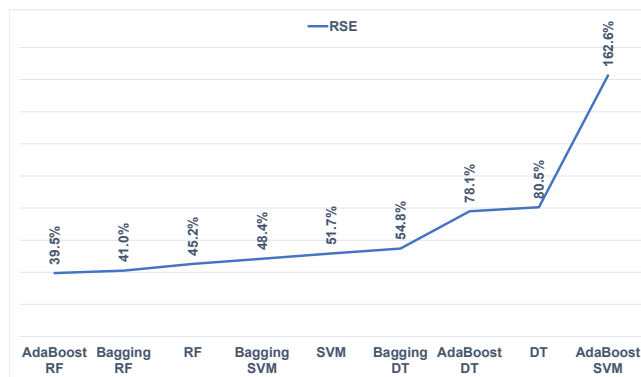


Fig. 14. Relative Squared Error (RSE) rate of the applied ML classifiers, AdaBoost, and Bagging methods.

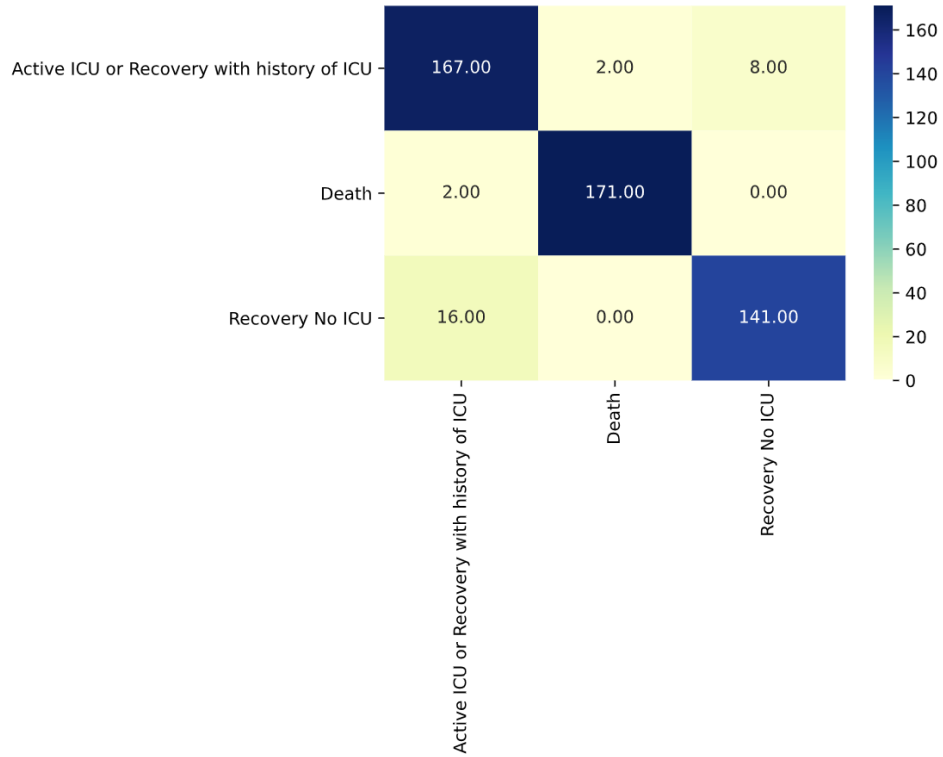


Fig. 15. Confusion matrix computed from the Bagging-DT method.

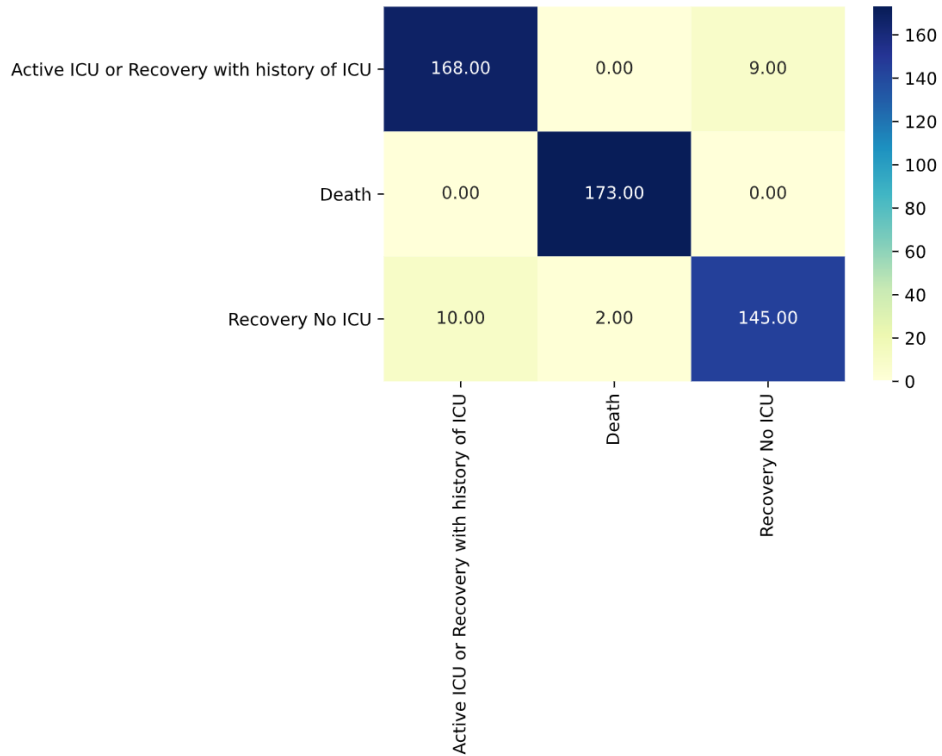


Fig. 16. Confusion matrix computed from the Bagging-SVM method.



Fig. 17. Confusion matrix computed from the Bagging-RF method.

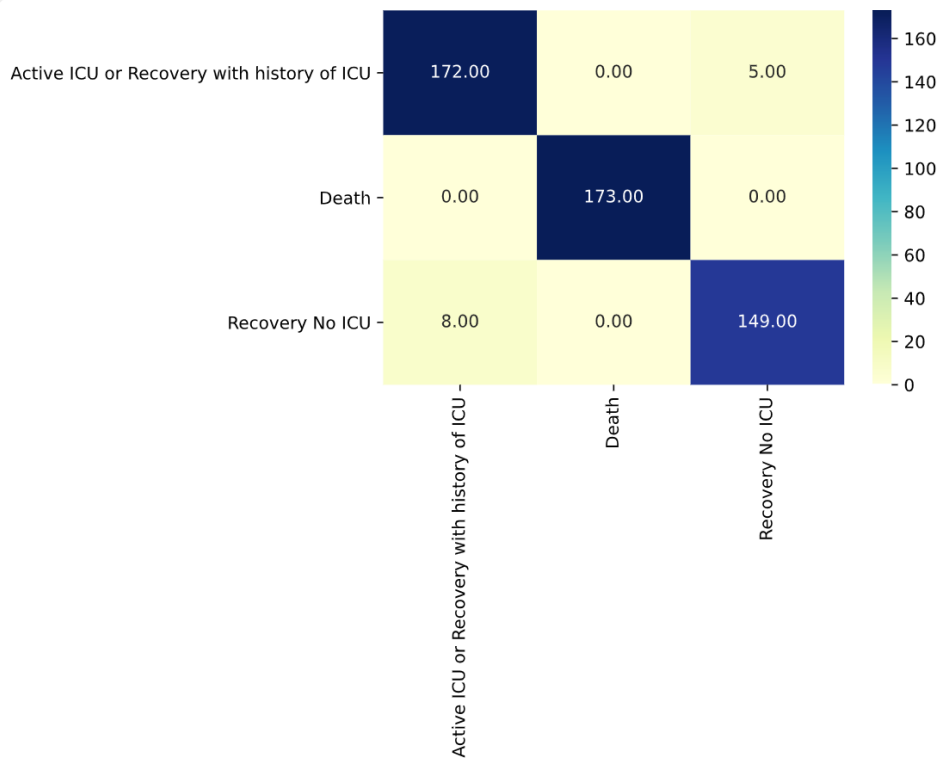


Fig. 18. Confusion matrix computed from the AdaBoost-RF method.

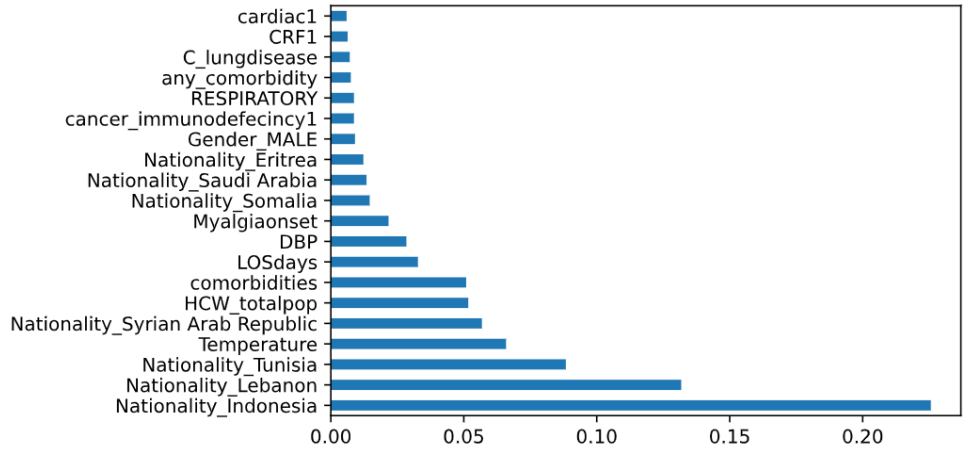


Fig. 19. Feature importance computed from the Bagging-DT method that was fitted to the dataset.

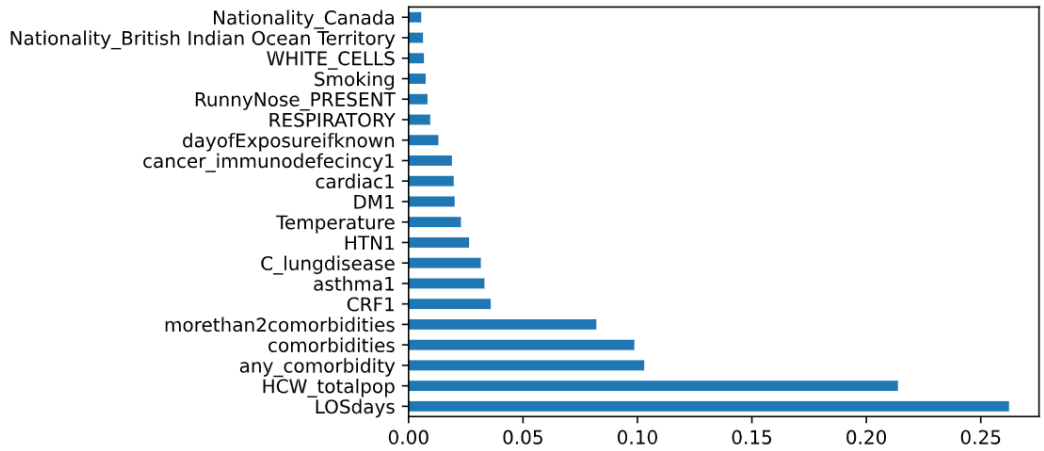


Fig. 20. Feature importance computed from the Bagging-SVM method that was fitted to the dataset.

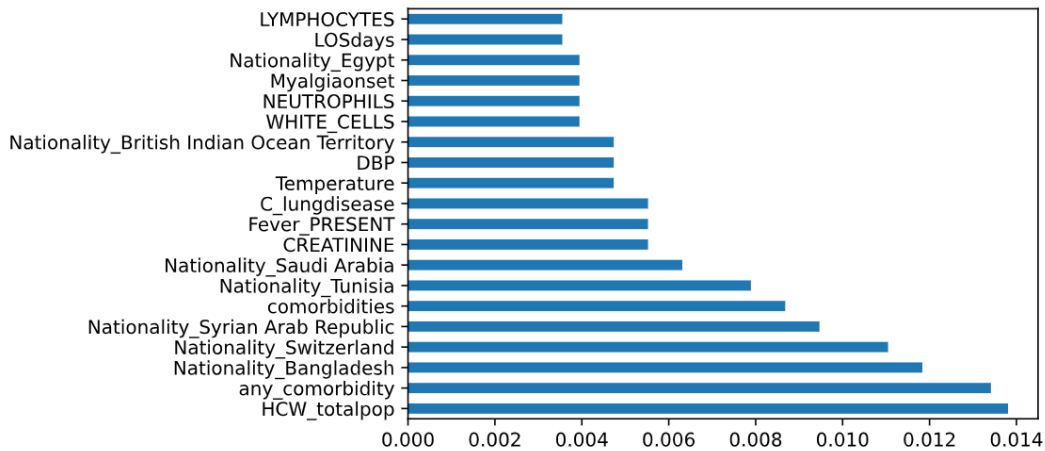


Fig. 21. Feature importance computed from the Bagging-RF method that was fitted to the dataset.

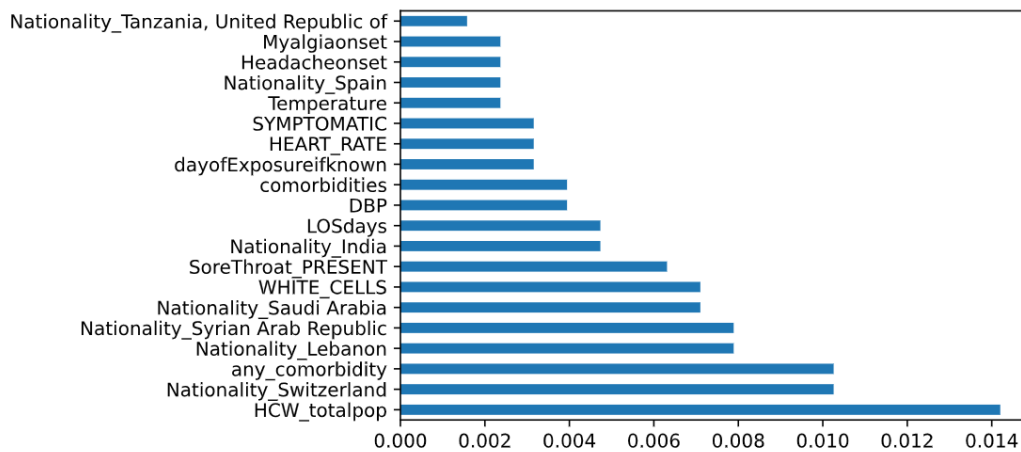


Fig. 22. Feature importance computed from the AdaBoost-RF method that was fitted to the dataset.

D. Feature Importance

From Fig. 19 – Fig. 22, the authors can see the feature importance of the dataset. From Fig. 19, Nationality_Indonesia is the most important feature among all the dependent features, whereas some diseases (e.g., cardiac1, CRF1, and C_lungdisease) are the least important features of the dataset. From Fig. 20, LOSdays is the most important feature among all the dependent features, whereas some nationalities (e.g., Nationality_Canada, Nationality_British Indian Ocean Territory) are the least important features of the dataset. From Fig. 21, HCW_totalpop is the most important feature among all the dependent features, whereas LYMPHOCYTES is the least important feature of the dataset. From Fig. 22, HCW_totalpop is the most important feature among all the dependent features, whereas Nationality_Tanzania is the least important feature of the dataset.

V. DISCUSSION

The findings in this study emphasize the importance of advanced ML in health-care decision-making for better predictive capabilities and resource allocation. Traditional methods have trouble gathering complex health-care data patterns, and this was especially true during COVID-19. ML methods can help health-care practitioners extract insights and construct predictive models, identify patients who need extra support, optimize resource allocation, reduce readmission rates, and improve patient care quality.

This study successfully validated the findings of a previous investigation [32], bringing greater credence to its conclusions. The previous study employed the same dataset features, approach, and assessment schema, and included a variety of weak classifiers including, NC,¹⁸ KNN, SVM, DT, RF, ANN, and Ensemble learners methods. By reproducing these experimental settings, the current study not only confirmed earlier findings but also reinforced their robustness and generalizability.

Several factors can be attributed to the limitations of the study. Firstly, the used dataset was relatively small, due to

not only its limiting scope to a single geographical territory but also it was intended for health-care facility readiness and epidemiological surveillance. Moreover, there were a few administrative challenges that had an impact on the curating of the dataset. These challenges include: 1) insufficient systematic procedures regarding the collection, storage, and sharing of medical data, 2) applying data privacy and security measures to protect sensitive and personally identifiable patient information, 3) applying data standardization processes were necessary to consolidate inputs from different sources that used diverse data formats and coding systems, and 4) ensuring data completeness to rectify potential errors in data entry and inconsistent recording practices. As a result of these challenges, different methods or research groups may utilize the data to varying extents, affecting its applicability and reliability.

To more fully understand the COVID-19 implications, there are still many missing puzzles. In China, at least 5% of COVID-19 patients develop severe illness and become critically sick, with critically ill patients having an ICU death rate of 50%–60%. Early detection and treatment of warning symptoms can minimize mortality and increase cure rates [20]. Specialized tests, such as lactate dehydrogenase level and total blood count, are utilized in developing countries like Iran to assess patient deterioration. Although these tests are not specific, they can be used in combination with RT-PCR tests, the most commonly used test for COVID-19 identification, to improve accuracy [17]. In addition, individuals infected with COVID-19 are more prone to develop neurological and mental diseases such as dementia and psychosis, even two years after diagnosis. Adults had a greater risk of mental diseases or anxiety resulting from COVID-19; however, for those with other respiratory infections, this risk decreased to baseline levels after two months. Even two years after the first infection, the risk of cognitive damage remained significant six months after infection. Oxford University researchers discovered mental problems, strokes, and dementia in COVID-19-infected people. The Lancet journal released research that revealed a worldwide increase in serious depression and anxiety disorders [18].

Studying the impact of the illness on age extremes (elderly,

¹⁸Nearest Centroid method

pediatrics) revealed special attention should be given to the elderly. The epidemiological profiles, clinical characteristics, risk factors, and final outcomes for COVID-19 cases have been extensively documented. One of the early studies showed young men are the most affected and that cough, fever, and sore throat were the most dominant clinical manifestations of the disease [21]. Nevertheless, another recent study documented risk factors associated with unfavorable outcomes in COVID-19 cases. This included male gender, elderly (above age 60), and specific comorbidities (cardiac and chronic respiratory diseases) [33, 34, 35].

To manage the negative impact of COVID-19 on Saudi Arabia's territories, many technological and administrative initiatives have taken place. First, a nationwide plan was developed following WHO-suggested frameworks to delay the recording of the first case in the country. The plan included (1) structuring a ministerial committee to make proper decisions and monitor their implications, (2) sending messages related to the disease via different media platforms and engaging the community, (3) deploying rapid response teams and disease surveillance, (4) controlling points of entry, (5) escalating laboratory capabilities, (6) sharing protocols and guidelines, (7) providing COVID-19 cases with management and establishing surge capacity plans, (8) providing logistic support, and (9) ensuring health-care services for non-COVID-19 cases [36]. In addition, the implementation of the eight pillars of response has been carried out by Saudi Arabia by the Operational Planning Guidelines to Support Country Preparedness and Response provided by the WHO. This implementation includes the integration of digital technologies to enhance the effectiveness of preparedness and response efforts. These technologies have been widely implemented on a global scale to address a range of pandemic-related objectives, such as preventative measures and contact tracing efforts [37].

To make it more comprehensive, effective, and integrated, many Saudi health-care-related projects have prioritized innovation, financial sustainability, illness prevention, and increased access to health care. For instance, SEHA virtual hospital launched in 2022 integrating over 150 institutions with over 30 specialized health services [38].

VI. CONCLUSIONS AND DIRECTIONS FOR FUTURE

Using the AdaBoost and Bagging methods, the authors investigated the utilization of the both ML models to correctly predict ICU rates in COVID-19 patients. The models have a high degree of accuracy, sensitivity, and positive predictive value. The models generated using AdaBoost-RF and Bagging-RF demonstrated the highest levels of precision among all the models, with an accuracy of 97.4% and 97.2% respectively. These models could assist health-care institutions in identifying who is at a higher risk of readmission, allowing for more targeted interventions and improved patient outcomes. Similar to previous research, this work demonstrates the feasibility of creating, validating, and utilizing ML predictive models for forecasting ICU admission in cases of COVID-19 infection. The models have the potential to be integrated into decision-support systems for semi-autonomous diagnostic equipment, enabling them to screen and diagnose potential outbreaks quickly. Subsequent research endeavors should focus on the development of ML prediction models aimed at identifying

individuals who are susceptible to experiencing severe consequences as a result of influenza, emphysema, or pulmonary fibrosis.

A. Data Availability Statement

The dataset presented in this study is private property of the Global Center for Mass Gatherings Medicine.

B. Funding

This project received no funding.

C. Conflict of Interest

All authors declared there are no conflicts of interest.

ACKNOWLEDGEMENT

This project resulted from a collaboration between the Global Center for Mass Gatherings Medicine and the College of Computer Science and Engineering in Taibah University, starting from October 1, 2020, until October 6, 2023.

REFERENCES

- [1] World Health Organization, "Interactive timeline of WHO's response to covid-19," <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline>, 2023, accessed: October 6, 2023.
- [2] —, "Who coronavirus COVID-19 dashboard," <https://covid19.who.int/>, 2023, accessed: October 4, 2023.
- [3] G. Meyerowitz-Katz and L. Merone, "A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates," *Epidemiology*, preprint, May 2020. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/33007452/>
- [4] W. D. de Holanda, L. C. e Silva, and Álvaro Alvares de Carvalho César Sobrinho, "Machine learning models for predicting hospitalization and mortality risks of COVID-19 patients," *Expert Systems with Applications*, vol. 240, p. 122670, 2024.
- [5] K. H. Abegaz and I. Etikan, "Boosting the performance of artificial intelligence-driven models in predicting COVID-19 mortality in ethiopia," *Diagnostics*, vol. 13, no. 4, p. 658, 2023.
- [6] A. A. Khan, H. M. Alahdal, R. M. Alotaibi, H. S. Sonbol, R. H. Almaghrabi, Y. M. Alsofayan, S. M. Althunayyan, F. A. Alsaif, S. S. Almudarra, K. I. Alabdulkareem, A. M. Assiri, and H. A. Jokhdar, "Controlling covid-19 pandemic: A mass screening experience in saudi arabia," *Frontiers in Public Health*, vol. 8, 2021. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fpubh.2020.606385>
- [7] Saudi Ministry of Health, "Saudi ministry of health news," <https://www.moh.gov.sa/en/Ministry/MediaCenter/News/Pages/News-2020-03-02-002.aspx>, 2020, accessed: October 6, 2023.
- [8] F. Alqahtani, A. Khan, J. Alowais, T. Alaama, and H. Jokhdar, "Bed surge capacity in saudi hospitals during the covid-19 pandemic," *Disaster Medicine and Public Health Preparedness*, vol. 16, no. 6, p. 2446–2452, 2022.

- [9] A. Khan, K. L. Bieh, A. El-Ganainy, S. Ghallab, A. Assiri, and H. Jokhdar, "Estimating the COVID-19 risk during the Hajj pilgrimage," *Journal of Travel Medicine*, vol. 27, no. 8, p. taaa157, 09 2020. [Online]. Available: <https://doi.org/10.1093/jtm/taaa157>
- [10] A. A. Khan, Y. AlRuthia, B. Balkhi, S. M. Alghadeer, M.-H. Tamsah, S. M. Althunayyan, and Y. M. Alsafyan, "Survival and estimation of direct medical costs of hospitalized covid-19 patients in the kingdom of saudi arabia," *International Journal of Environmental Research and Public Health*, vol. 17, no. 20, 2020. [Online]. Available: <https://www.mdpi.com/1660-4601/17/20/7458>
- [11] Saudi Ministry of Health, "COVID-19 dashboard: Saudi arabia," <https://covid19.moh.gov.sa/>, 2023, accessed: October 4, 2023.
- [12] H. A. Sheerah, Y. Almuzaini, and A. Khan, "Public Health Challenges in Saudi Arabia during the COVID-19 Pandemic: A Literature Review," *Healthcare*, vol. 11, no. 12, p. 1757, Jun. 2023. [Online]. Available: <https://www.mdpi.com/2227-9032/11/12/1757>
- [13] L. Wynants, B. Van Calster, G. S. Collins, R. D. Riley, G. Heinze, E. Schuit, E. Albu, B. Arshi, V. Bellou, M. M. J. Bonten, D. L. Dahly, J. A. Damen, T. P. A. Debray, V. M. T. De Jong, M. De Vos, P. Dhiman, J. Ensor, S. Gao, M. C. Haller, M. O. Harhay, L. Henckaerts, P. Heus, J. Hoogland, M. Hudda, K. Jenniskens, M. Kammer, N. Kreuzberger, A. Lohmann, B. Levis, K. Luijken, J. Ma, G. P. Martin, D. J. McLernon, C. L. A. Navarro, J. B. Reitsma, J. C. Sergeant, C. Shi, N. Skoetz, L. J. M. Smits, K. I. E. Snell, M. Sperrin, R. Spijker, E. W. Steyerberg, T. Takada, I. Tzoulaki, S. M. J. Van Kuijk, B. C. T. Van Bussel, I. C. C. Van Der Horst, K. Reeve, F. S. Van Royen, J. Y. Verbakel, C. Wallisch, J. Wilkinson, R. Wolff, L. Hooft, K. G. M. Moons, and M. Van Smeden, "Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal," *BMJ*, p. m1328, 2020. [Online]. Available: <https://www.bmj.com/lookup/doi/10.1136/bmj.m1328>
- [14] M. Soui, N. Mansouri, R. I. Alhamad, M. Kessentini, and K. Ghédira, "Nsga-ii as feature selection technique and adaboost classifier for COVID-19 prediction using patient's symptoms," *Nonlinear Dynamics*, vol. 106, pp. 1453 – 1475, 2021. [Online]. Available: <https://www.semanticscholar.org/paper/d66b8689a4489c9ea67d8246fe88dbc8b4035cff>
- [15] M. Darici, "Performance analysis of combination of CNN-based models with Adaboost algorithm to diagnose covid-19 disease," *Journal of Polytechnic*, vol. null, p. null, 2021. [Online]. Available: https://www.semanticscholar.org/paper/Performance-Analysis-of-Combination-of-CNN-based-to-Darici/0f3c5f9c37077787a2ff4e7d658203f20d31ce6e?utm_source=direct_link
- [16] S. R. Mary, V. Kumar, K. J. P. Venkatesan, R. S. Kumar, N. P. Jagini, and A. Srinivas, "Vulture-based AdaBoost-feedforward neural frame work for COVID-19 prediction and severity analysis system," *Interdisciplinary Sciences, Computational Life Sciences*, vol. 14, pp. 582 – 595, 2022.
- [17] R. Mazloumi, S. R. Abazari, F. Nafarieh, A. Aghsami, and F. Jolai, "Statistical analysis of blood characteristics of COVID-19 patients and their survival or death prediction using machine learning algorithms," *Neural Computing and Applications*, vol. 34, no. 17, pp. 14 729–14 743, 2022. [Online]. Available: <https://link.springer.com/10.1007/s00521-022-07325-y>
- [18] S. Sharma, Y. K. Gupta, and A. K. Mishra, "Analysis and prediction of COVID-19 multivariate data using deep ensemble learning methods," *International Journal of Environmental Research and Public Health*, vol. 20, no. 5943, 2023.
- [19] S. Solayman, S. A. Aumi, C. S. Mery, M. Mubassir, and R. Khan, "Automatic COVID-19 prediction using explainable machine learning techniques," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 36 – 46, 2023.
- [20] Q.-Y. Li, Z.-Y. An, Z.-H. Pan, Z.-Z. Wang, Y.-R. Wang, X.-G. Zhang, and N. Shen, "Severe/critical COVID-19 early warning system based on machine learning algorithms using novel imaging scores," *World Journal of Clinical Cases*, vol. 11, no. 12, pp. 2716–2728, 2023. [Online]. Available: <https://www.wjgnet.com/2307-8960/full/v11/i12/2716.htm>
- [21] Y. M. Alsafyan, S. M. Althunayyan, A. A. Khan, A. M. Hakawi, and A. M. Assiri, "Clinical characteristics of COVID-19 in Saudi Arabia: A national retrospective study," *Journal of Infection and Public Health*, vol. 13, no. 7, pp. 920–925, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1876034120304925>
- [22] A. A. Alahmari, A. A. Khan, A. Elganainy, E. L. Almohammadi, A. M. Hakawi, A. M. Assiri, and H. A. Jokhdar, "Epidemiological and clinical features of covid-19 patients in saudi arabia," *Journal of Infection and Public Health*, vol. 14, no. 4, pp. 437–443, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1876034121000071>
- [23] R. E. Schapire, "The strength of weak learnability," *Machine Learning*, vol. 5, no. 2, pp. 197–227, Jun. 1990. [Online]. Available: <http://link.springer.com/10.1007/BF00116037>
- [24] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York, NY: Springer New York, 2013, vol. 1. [Online]. Available: <http://link.springer.com/10.1007/978-1-4614-6849-3>
- [25] R. E. Schapire, *The Boosting Approach to Machine Learning: An Overview*. New York, NY: Springer New York, 2003, ch. 8, pp. 149–171. [Online]. Available: http://link.springer.com/10.1007/978-0-387-21579-2_9
- [26] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S002200009791504X>
- [27] R. E. Schapire and Y. Freund, "A Brief Introduction to Boosting," in *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, vol. 2. Stockholm, Sweden: Morgan Kaufmann Publishers Inc., 1999, pp. 1401–1406.
- [28] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available:

- <http://link.springer.com/10.1023/A:1010933404324>
- [29] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sensing of Environment*, vol. 62, no. 1, pp. 77 – 89, 1997. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0034425797000837>
- [30] K. W. Bowyer, N. V. Chawla, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *CoRR*, vol. abs/1106.1813, 2011. [Online]. Available: <http://arxiv.org/abs/1106.1813>
- [31] R. Vidiyala, "Performance metrics for classification machine learning problems," <https://towardsdatascience.com/performance-metrics-for-classification-machine-learning-problems-97e7e774a007>, 2020, accessed on October 6, 2023.
- [32] H. Ghandorh, M. Z. Khan, R. Alsufyani, M. Khan, Y. M. Alsofayan, A. A. Khan, and A. A. Alahmari, "An icu admission predictive model for COVID-19 patients in Saudi Arabia," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2021.0120764>
- [33] A. Khan, S. Althunayyan, Y. Alsofayan, R. Alotaibi, A. Mubarak, M. Arafat, A. Assiri, and H. Jokhdar, "Risk factors associated with worse outcomes in COVID-19: a retrospective study in Saudi Arabia," *Eastern Mediterranean Health Journal*, vol. 26, no. 11, pp. 1371–1380, Nov. 2020. [Online]. Available: <https://applications.emro.who.int/emhj/v26/11/1020-3397-2020-2611-1371-1380-eng.pdf>
- [34] F. Alamri, Y. Alsofayan, Y. AlRuthia, A. Alahmari, Y. Almuzaini, F. A. Gazalah, F. Alradini, T. Alaama, and A. Khan, "Predictors of hospitalization among older adults with COVID-19 in Saudi Arabia: A cross-sectional study of a nationally representative sample," *Risk Management and Healthcare Policy*, vol. 14, pp. 875–886, Mar. 2021. [Online]. Available: <https://www.dovepress.com/predictors-of-hospitalization-among-older-adults-with-covid-19-in-saud-peer-reviewed-fulltext-article-RMHP>
- [35] Y. Almuzaini, F. Alsohime, S. Subaie, M. Temsah, Y. Alsofayan, F. Alamri, A. Alahmari, H. Alahdal, H. Sonbol, R. Almaghrabi, M. Nahhas, and A. Khan, "Clinical profiles associated with SARS-CoV-2 infection and complications from coronavirus disease-2019 in children from a national registry in Saudi Arabia," *Annals of Thoracic Medicine*, vol. 16, no. 3, p. 280, 2021. [Online]. Available: <http://www.thoracicmedicine.org/text.asp?2021/16/3/280/321905>
- [36] A. Khan, Y. Alsofayan, A. Alahmari, J. Alowais, A. Algwizani, H. Alserehi, A. Assiri, and H. Jokhdar, "COVID-19 in Saudi Arabia: the national health response," *Eastern Mediterranean Health Journal*, vol. 27, no. 11, pp. 1114–1124, Dec. 2021. [Online]. Available: <https://applications.emro.who.int/EMHJ/V27/11/1020-3397-2021-2711-1114-1124-eng.pdf>
- [37] A. Khan, A. Alahmari, Y. Almuzaini, N. Alturki, A. Aburas, F. A. Alamri, M. Albagami, M. Alzaid, T. Alharbi, R. Alomar, M. A. Tayli, A. M. Assiri, and H. A. Jokhdar, "The role of digital technology in responding to COVID-19 pandemic: Saudi Arabia experience," *Risk Management and Healthcare Policy*, vol. 14, pp. 3923–3934, Sep. 2021. [Online]. Available: <https://www.dovepress.com/the-role-of-digital-technology-in-responding-to-covid-19-pandemic-saud-peer-reviewed-fulltext-article-RMHP>
- [38] Vision2030 Editing Staff, "Health sector transformation program," Vision 2030 - Kingdom of Saudi Arabia, Riyadh, 2023. [Online]. Available: <https://www.vision2030.gov.sa/en/vision-2030/vrp/health-sector-transformation-program/#:~:text=This%20enhanced%20system%20prioritizes%20innovation,and%20adhering%20to%20international%20standards.>