

Machine Learning-Driven Prediction of ICU Admissions for COVID-19 Patients

A.M. Mutawa*, *Member, IEEE*

Abstract—In the wake of the COVID-19 pandemic, efficiently allocating ICU resources for critical patients has become crucial, especially for those with chronic conditions. This study harnesses machine learning (ML) to forecast ICU admissions among COVID-19 patients in Kuwait, analyzing a dataset of 4399 patients to identify pivotal predictors for ICU needs. Employing cross-validation and Synthetic Minority Over-sampling Technique (SMOTE) to tackle data imbalance, the predictive variables were refined using backward feature selection with logistic regression and evaluated model interpretability with Shapley additive explanations (SHAP). The Support Vector Machine (SVM) model outperformed other models with an area under the curve (AUC) of 0.91, and the Extra Tree (ET) model showed better performance with an accuracy of 96.42%. Critical predictors included demographics, clinical outcomes like shortness of breath, elevated d-dimer levels, and abnormal chest X-rays. This research not only underscores the potential of ML in critical healthcare decision-making during pandemics but also highlights its role in discovery science, suggesting broader applications in healthcare and other scientific domains. The study advances medical informatics by integrating ML with healthcare, offering insights into disease dynamics and improving resource allocation strategies.

Index Terms— COVID-19, Cross-validation, Feature selection, Intensive care unit, Machine learning.

I. INTRODUCTION

THE COVID-19 pandemic has led to unprecedented levels of mortality, morbidity, and economic impact. By 25th July 2022, the number of reported cases of COVID-19 worldwide was 576.58 million, with 6.4 million deaths [1]. Given the scale of the problem and the variable clinical course, identifying and targeting individuals at the highest risk becomes important for efficiently allocating limited resources [2]. Even though the immediate crisis of COVID-19 may have abated, the tools and insights garnered during its peak remain pivotal. Future health emergencies, such as infectious outbreaks or other challenges, will benefit from the strategic methodologies honed during COVID-19, ensuring their continued relevance. While occasionally predictable, contagious disease outbreaks always carry uncertainty regarding their exact character and scope. Research undertaken during the COVID-19 period fortifies the preparedness for such unpredictable challenges. Refined predictive models, innovative techniques, and strategic insights are invaluable assets for managing prospective health crises.

The methodologies and ML models fine-tuned during the

pandemic are versatile and poised to be employed across a spectrum of healthcare domains. The strategic insights, predictive analytics, and data management skills cultivated during this period can be effectively repurposed for diverse diseases, conditions, and health emergencies. Age and medical conditions increase the likelihood of severe COVID-19 and its consequences [3]. Acute or chronic COVID-19 disease persists in people without vaccines or treatments and those with underlying conditions that decrease immunity to vaccinations [4].

An ICU is a specialized critical care ward dedicated to treating and caring for patients with serious diseases that substantially threaten their lives. An ICU admission policy for older people would consider their unique requirements and enhance admission options to improve their results and experiences [5, 6]. Critically ill emergency department patients require enhanced interaction and translation assistance, according to research by Olds et al. [7].

Medical organizations should monitor ICU treatment usage and ensure that ICU admission decisions are patient-focused while considering the resources and restrictions of each hospital [8]. The virus mutation with COVID-19 had severe changes in people, affecting individual treatments and vaccinations. Hence, data analysis provides essential information for future pandemics [9]. According to Ungar et al. [10], multiple medical conditions are substantially related to ICU admission among children diagnosed with COVID-19.

The relationship between diabetes as a significant risk factor and COVID-19 disease severity was examined in the literature using both "classic" statistical analysis techniques [11-14] and ML techniques [15, 16]. The diabetes burden is incredibly high in Kuwait, with an overall adjusted prevalence of 19.1% for diabetes and 13.5% for prediabetes [17], making this especially relevant in the context of this study.

ML significantly enhances the potential for strategic handling of clinical research [18, 19]. Several studies have examined the COVID-19-related ICU admission risk variables and employed Artificial Intelligence (AI) and ML techniques to develop models for predicting disease severity based on laboratory investigations, imaging, or clinical note data. A review of previous studies is mentioned in Table 1.

TABLE I

REVIEW OF ML MODELS FOR PREDICTION IN THE FIELD OF HEALTH CARE.

Reference and year	Model	Target type	Dataset type	Results
[20] 2023	LR, DT, RF, XGB, Light GB, ensemble	In-hospital mortality of lung cancer patient	1176 lung cancer patient	Ensemble model with 89% accuracy
[21] 2023	LR, RF, XGB	ICU transfer prediction with dichotomous variables	532 COVID-19 patients	LR with AUC = 0.748
[22] 2022	DT, SVM, MLP, KNN	Predicting intubation risk with 54 variables	1225 COVID-19 patients	DT with 93.8% accuracy
[23] 2022	Different DT algorithms	ICU prediction with 53 variables	512 COVID-19 patients	J-48 with 81.9% accuracy
[24] 2022	14 ML classifiers	Prediction of deterioration with 33 variables	1079 COVID-19 patients	AUROC of CatBoost = 0.84
[25] 2022	Light GB, LR	Prediction of mortality	1571 COVID-19 patients	Light GBM with AUC = 0.88
[26] 2022	log-binomial mixed-effects regression	Prediction of ICU admission and mortality	793 pregnant COVID-19 patients	AUROC = 0.73
[27] 2022	LR, CART, C4.5, C5.0	Prediction of ICU admission	228 Myasthenia gravis patients	C5.0 with AUC = 0.814
[28] 2021	18 different ML algorithms	Prediction of ICU admission and mortality	5308 COVID-19 patients	Ensemble models with F1-score = 0.81
[14] 2021	LR	Prediction of mortality with 57 variables	247 COVID-19 diabetic and prediabetic patients	AUC = 0.889
[29] 2021	Multivariable LR	Prediction of ICU admission and mortality	356 COVID-19 patients	AUC = 0.77
[30] 2021	RF, MLP, SVM, ET, GB, Adaboost	Prediction of ICU and ventilation	212 COVID-19 patients	AUC = 0.80 with RF model
[31] 2021	GB, LR, RF, SVM, DT	Prediction of ICU with CBC data	1218 COVID-19 patients	AUC = 0.88 with ensemble models
[32] 2021	MLP, RF, XGB,	Prediction of ICU admission with 165 variables	3623 COVID-19 patients	AUC = 0.83 with XGB model
[16] 2020	LR, RF, XGB, ANN, majority voting, CNN	Prediction of mortality	9954 COVID-19 diabetic patients with	AUC = 0.97 with CNN

			clinical notes	
Abbreviations: AUROC: area under the receiver operating characteristic curve, DT: Decision tree, SVM: Support vector machine, MLP: Multilayer perceptron, KNN: K-nearest neighbors, ANN: Artificial neural network, LR: Logistic regression, CatBoost: categorical boosting, GB: Gradient boosting, XGB: extreme gradient boosting, AUC: area under the curve, CART: classification and regression tree, RF: Random forest, CBC: complete blood count, ET: Extra Trees classifier				

An unbalanced classification issue has skewed instances across majority classes. Predictive modeling is complex with imbalanced classifications as most ML algorithms assume equal instances for each class, thereby making minority class models unreliable—a concern as the minority class is usually more significant than the majority class and particularly vulnerable to classification errors [33, 34]. The Synthetic Minority Over-sampling Technique (SMOTE) is a data augmentation approach employed in ML to overcome the class imbalance problem [35]. SMOTE creates a balanced dataset with equal numbers of majority and minority cases. It improves skewed data in ML models and is a common but simple and effective oversampling method [36].

This study explores the application of ML algorithms to develop prediction models for ICU admission using a rich set of clinical and laboratory parameters among COVID-19 patients, distinguishing between those with and without diabetes. One major challenge faced in medical datasets is the issue of data imbalance, in which one class (e.g., patients requiring ICU) is significantly underrepresented. This imbalance can introduce a bias in the predictive model, often leading to inaccurate predictions, especially for the minority class. SMOTE is a standard imbalance solution. The model makes better, more impartial predictions by producing synthetic samples in the feature space to balance the minority class.

While applying these techniques provides a robust foundation for this analysis, selecting the most informative variables is equally crucial. Not all variables might be essential for accurate predictions, prompting the need for effective feature selection methods.

Detailed methodologies, including the application of SMOTE and feature selection procedures, are elaborated in the methods section. Based on the Population, Intervention, Comparison, And Outcome (PICO) approach, the research questions were framed as follows:

- Research question 1 (RQ1): In patients diagnosed with COVID-19, how effectively do different ML models predict ICU admissions?
- Research question 2 (RQ2): In patients diagnosed with COVID-19, which essential clinical and laboratory variables, when incorporated into this ML model, most significantly predict the need for ICU admission?
- Research question 3 (RQ3): Among COVID-19 patients known to have diabetes, how do the essential variables differentiate in their predictive power for ICU admission compared to the broader COVID-19 patient

population?

The proposed study is motivated by the urgent need for a precise and efficient predictive model tailored for a specific demographic to inform healthcare responses. The main contributions of this paper are:

- 1) **A Comprehensive Dataset:** By consolidating a broad spectrum of data, from demographic details to specific clinical findings, one of the most exhaustive datasets related to ICU requirements is offered amidst the COVID-19 pandemic in Kuwait.
- 2) **ML-based Approach with Feature Selection:** Recognizing the potential pitfalls of using raw, unfiltered data, backward feature selection with logistic regression is employed, enhancing the accuracy and efficiency of the prediction models. Support vector machine (SVM), extra trees (ET), decision tree (DT), logistic regression (LR), and random forest (RF) were the ML algorithms used in this study.
- 3) **Addressing Data Imbalance:** In medical datasets, data imbalances can often lead to skewed results. The adoption of SMOTE ensures a balanced and unbiased model prediction.
- 4) **Transparent Feature Importance with SHAP:** This model does not just predict; it also informs. By integrating Shapley additive explanations (SHAP) values, the relative significance of each predictor is elucidated, bridging the gap between ML and clinical interpretation.

II. MATERIALS AND METHODS

A. Study Design

A use-case scenario explains the context of the study, as depicted in Fig 1. The use-case diagram consists of actors or persons interacting with the system, its functioning, and the connection link or the relationship between them. The patient with COVID-19 is one of the actors. The patients were admitted to the hospital and consulted by the doctors. They were tested for COVID-19. The other two actors in the use case (doctors and nurses) recorded their symptoms and clinical records. All the medical records, including the ICU admission, were created in a database maintained by an admin (actor). The "include" parameter in the connection link requires another use case to perform the task. The patient was discharged from the hospital based on the doctor's decision. Once the dataset was ready, the ML model developer (actor) accessed the data and preprocessed it, including cleaning, imputing, and standardizing. The best medical record variables were selected using the backward variable selection method. The ML model parameters were tuned, and the best parameter was picked for training and validation to implement and deploy the model. The deployed model can then be employed in the clinical setting to predict new patient data. The model contained preprocessing steps and had the best variables with the best ML model parameters.

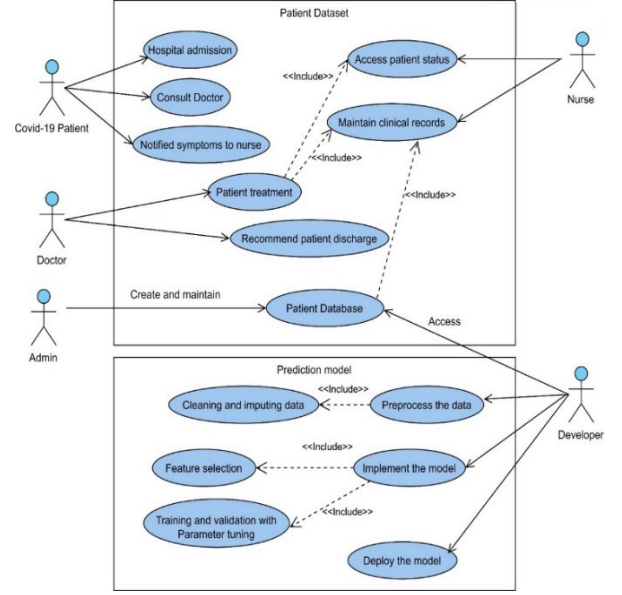


Fig. 1. Use-case scenario of the study.

B. Dataset

Data were collected from patients admitted from Kuwait Hospitals between 4 May and 26 August 2020. The Permanent Committee for Coordination of Medical and Health Research at the Ministry of Health (MOH) obtained all the legal approvals for sample collection and patient surveys with consent forms. Patient data were collected retrospectively, and the inclusion criteria of the validation cohort were based on admission and discharge data availability. Patients who did not have this information were not included in the validation group. All patients were tested for COVID-19, and the polymerase chain reaction (PCR) result was positive. A total of 4555 patient data was collected. From this, 156 individuals were disqualified from the study because of excessive missing values. So, the total number of the complete cohort employed in this study was 4399. The mean age of the patients was 42, and the first-third quartile (Q1–Q3) shows ages from 31–54. Table 2 represents the demographic information according to age, gender, and ICU admission requirements.

TABLE II

THE STUDY COHORT ACCORDING TO AGE, GENDER, AND ICU ADMISSION

Condition	Age	Male	Female
ICU admission-required	Up to 30	7	1
	Between 31 and 50	97	6
	Between 51 and 65	91	18
	Above 65	46	15
ICU admission-not required	Up to 30	679	380
	Between 31 and 50	1371	489
	Between 51 and 65	569	302
	Above 65	173	155

The variables hypertension and diabetes mellitus (DM) were calculated based on the results of a few study parameters. Suppose a patient has systolic blood pressure greater than 140 or diastolic blood pressure (DBP) greater than 90; in that case, they are said to have hypertension (value=1), otherwise not (value=0). The glucose parameter shows the random glucose level of hospital patients. Estimating glucose levels based on a random sample is called random glucose level. In all cases, it is not a fasting sample. The blood glucose level can be measured at any time of the day by using this test. A patient has DM (value=1) if the glucose ≥ 11.1 mmol/L or HbA1c ≥ 6.5 mmol/L; otherwise, non-diabetes (DM value=0). If the patient had less glucose value but had a positive DM history, that person was under medication [37].

The demographic predictors included age, gender, height, and nationality. Each patient's symptoms during COVID-19 were marked here, including weakness, sore throat, headache, SOB, abdominal pain, fever, expectoration, diarrhea, asthma, cough, and history of bariatric surgery. The diabetes result considered at the hospital was the history of DM, glucose level, and hemoglobin A1c or glycated hemoglobin (HbA1cA) value. Specific clinical tests were also considered. These were temperature, chest x-ray (CXR), creatinine, phosphate, pulse, respiratory rate, oximeter, prothrombin time, potassium (K), sodium (Na), magnesium (Mg), estimated glomerular filtration rate (eGFR), c-reactive protein (CRP), urate, procaltitonin, total protein (T.Protein), total bilirubin (T.Bili), serum albumin (S.Albumin), carbon dioxide (CO2), chloride (Cl), blood urea nitrogen (BUN), calcium (Ca), activated partial thromboplastin clotting time (APTT), alanine transaminase (ALT), alkaline phosphatase (ALP), d-dimer, ferritin level, gamma-glutamyl transferase (GGT), and international normalized ratio (INR).

Blood test results such as lactate dehydrogenase (LDH), platelet count, neutrophil%, lymphocyte%, monocyte%, basophil%, eosinophil%, hemoglobin (HB), White blood cell (WBC), red blood cell (RBC), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean cell hemoglobin concentration (MCHC), red cell distribution width (RDW), and mean platelet volume (MPV) were also considered. Table 3 provides the complete description of each variable used in this study.

TABLE III
THE PARAMETERS OF THE COVID-19 DATASET

Variables	N	Mean \pm Std. Deviation	Missing data	
			Count	Percent
Age	4399	42.16 \pm 16.73	0	0
SBP	1389	129.94 \pm 18.17	3010	68.4
DBP	1388	76.95 \pm 10.97	3011	68.4
Temp	1341	37.04 \pm 0.63	3058	69.5
eGFR	4043	97.17 \pm 29.69	356	8.1
CRP	2524	32.89 \pm 74.84	1875	42.6
Procalcitonin	1470	3.57 \pm 28.46	2929	66.6

Creatinin	3908	96.22 \pm 108.81	491	11.2
Phosphate	4258	1.28 \pm 0.40	141	3.2
Urate	4179	297.94 \pm 109.26	220	5
T.Protein	4288	65.12 \pm 7.36	111	2.5
T.Bili	4160	14.38 \pm 22.74	239	5.4
S.Albumin	4359	34.84 \pm 7.41	40	0.9
Respiratory rate	1280	21.96 \pm 5.21	3119	70.9
Pulserate	1435	87.77 \pm 15.03	2964	67.4
Pulse oximetry	1355	96.41 \pm 4.95	3044	69.2
Prothrombin Time	3233	14.14 \pm 4.22	1166	26.5
Na	4356	137.70 \pm 3.98	43	1
Mg	4209	0.83 \pm 0.11	190	4.3
LDH	3861	116.75 \pm 406.87	538	12.2
K	4353	4.27 \pm 0.54	46	1
INR	3234	1.05 \pm 0.31	1165	26.5
GGT	3975	45.13 \pm 73.21	424	9.6
Ferritin level	3805	102.52 \pm 259.21	594	13.5
DDimer	1223	539.07 \pm 1470.42	3176	72.2
CO2	4362	25.46 \pm 3.13	37	0.8
CL	4363	102.77 \pm 5.53	36	0.8
Ca	4269	2.23 \pm 0.19	130	3
BUN	4362	5.87 \pm 7.52	37	0.8
APTT	3206	33.46 \pm 12.91	1193	27.1
ALT	4152	47.95 \pm 216.65	247	5.6
ALP	4321	82.87 \pm 69.97	78	1.8
WBCs	4395	8.03 \pm 4.74	4	0.1
Neutrophils	4398	57.24 \pm 16.18	1	0
Lymphocytes	4380	26.01 \pm 17.64	19	0.4
Monocytes	4384	8.79 \pm 3.25	15	0.3
Basophil	4384	0.44 \pm 0.28	15	0.3
Eosinophil	4398	2.27 \pm 2.50	1	0
RBCs	4398	4.78 \pm 0.89	1	0
HB	4398	129.08 \pm 23.76	1	0
HCT	4384	0.40 \pm 0.07	15	0.3
MCV	4384	83.60 \pm 7.28	15	0.3
MCHC	4384	324.20 \pm 17.47	15	0.3
MCH	4384	27.14 \pm 2.78	15	0.3
RDW	4383	13.79 \pm 2.34	16	0.4
MPV	4248	10.38 \pm 1.77	151	3.4
Platelets count	4398	287.35 \pm 117.14	1	0

Glucose	4282	8.87 ± 5.10	117	2.7
Hba1c	510	8.94 ± 2.44	3889	88.4
Gender	4399	-	0	0
Nationality	4399	-	0	0
CXR	1462	-	2937	66.8
Weakness	1513	-	2886	65.6
Headache	1513	-	2886	65.6
Sore throat	1513	-	2886	65.6
SOB	1513	-	2886	65.6
Abdominal pain	1513	-	2886	65.6
Fever	1513	-	2886	65.6
Expectoration	1513	-	2886	65.6
Diarrhea	1513	-	2886	65.6
Cough	1513	-	2886	65.6
Asthma	1513	-	2886	65.6
Bariatric Surgery	1513	-	2886	65.6
Hypertension	1389	-	3010	68.4
DM history	1513	-	2886	65.6
DM	4299	-	100	2.3
ICU admission (Target)	4399	-	0	0

Abbreviations: SBP: Systolic Blood Pressure, DBP: Diastolic Blood Pressure, CXR: Chest X-Ray, DM: Diabetes Mellitus, SOB: Shortness of breath, WBC: White Blood cell, RBC: Red Blood cell, HB: Hemoglobin, HCT: Hematocrit, MCV: Mean corpuscular volume, MCHC: Mean cell hemoglobin concentration, MCH: Mean corpuscular hemoglobin, eGFR: Estimated Glomerular Filtration Rate, CRP: C-reactive protein, T.Protein: Total protein, T.Bili: Total bilirubin, S.Albumin: serum albumin, Na: Sodium, Mg: Magnesium, LDH: Lactate dehydrogenase, K: Potassium, Hba1c: Hemoglobin A1c, INR: International normalized ratio, GGT: Gamma-glutamyl transferase, CO2: Carbon dioxide, Cl: Chloride, Ca: Calcium, BUN: Blood urea nitrogen, APTT: Activated Partial Thromboplastin Clotting Time, ALT: Alanine transaminase, ALP: Alkaline phosphatase, RDW: Red cell distribution width, MPV: Mean platelet volume, ICU: Intensive Care unit

C. Methodology

This section outlines the comprehensive methodology employed in this study. The dataset is detailed in the beginning, encompassing 4399 COVID-19 patient records with 66 variable columns encapsulating diverse demographic and clinical details. The variables were selected after preliminary data preprocessing, which includes cleaning, imputation, and normalization. This step ensured the retention of only the most relevant and non-collinear variables, using techniques such as Pearson correlation and the chi-square test. The data underwent further processing with the selected variables to tackle class imbalances via the SMOTE method. Subsequently, ML models, including SVM, ET, DT, LR, and RF, were applied. A thorough evaluation followed using a 3-fold cross-validation (CV) and metrics such as precision, recall, and F1-score. This

methodology ensures a robust and systematic analysis, as elucidated in the subsequent sections.

Fig 2 depicts the methodology used in this work. The dataset consists of 4399 COVID-19 patient records and 66 variable columns that indicate demographic details such as age, gender, nationality, COVID-19 symptoms, chronic diseases, diabetes, blood culture results, pulse rate, oximeter, and other clinical results, as mentioned in Table 3. The data was preprocessed before moving it to the classification models. The best variables were chosen from the variables with the help of a wrapper variable selection method. The data was split into 60% for train validation using a 3-fold CV method, and the remaining 40% was set as unseen data for external validation.

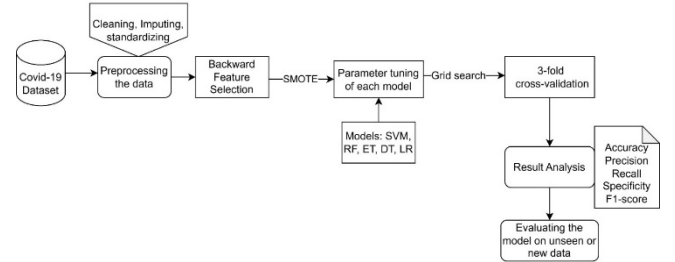


Fig. 2. Study methodology.

D. Preprocessing

The first stage was to clean the data before processing it into the model. In the next step, the missing values were replaced by an imputing method. The variables in Table 3 that showed more than 75% missing values were excluded from the study [23]. The KNN imputing (K-nearest neighbors) method was used for missing values [38]. It is the most straightforward algorithm for imputation and is based on Euclidean distance. The number of neighbors was set as 5 in this study, the default value in the package used. It meant that an average of five nearest neighbors replaced the missing value. The Euclidean distance was defined in equation (1) as

$$distance = \sqrt{weight * squared\ distance\ from\ f} \quad (1)$$

Where f is the present feature, and weight is defined in equation (2) as

$$weight = \frac{total\ number\ of\ features}{number\ of\ features\ having\ values} \quad (2)$$

The missing and imputed data were analyzed to ensure robustness. After the cells were filled by the KNN method, the data was preprocessed to make all the data in a particular range.

E. Variable Selection

There were 65 independent variables (after removing missing values greater than 75%) and a target variable (ICU admission required-1, and ICU admission not required-0). ML models might be unable to make the required predictions, utilizing all the variables used in the dataset. The highly correlated independent variables were removed based on the

Pearson correlation for the continuous variables, as shown in Fig 3. The correlation coefficient value between 0.7 and 1 was said to have a strong relationship between the independent variables (either positive or negative) [39]. The positive correlation implies that both variables move in the same direction. In contrast, the negative value shows that the variable is in the opposite direction. A strong relation was identified in variables between eGFR-creatinine, prothrombin time-INR, RBC-HB, RBC-HCT, HB-HCT, and MCV-MCH. These variables provide the same information, potentially leading to overfitting in ML modeling. Therefore, the variables creatinine, INR, RBC, HB, and MCV were removed from the dataset to mitigate this.

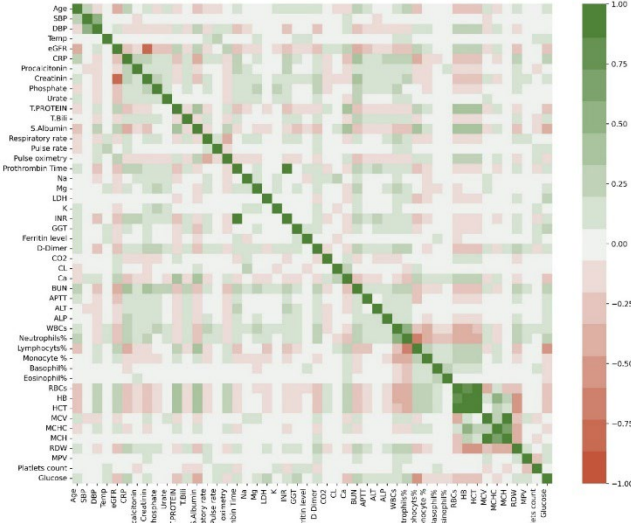


Fig. 3. Correlation between each continuous variable.

The chi-square test for independence determines the association and significance of binary variables (mentioned in Table 3). The association was checked on the independent variables. The chi-square test null hypothesis states no statistical difference between the variables observed. It is accepted when the p-value is greater than 0.05. The chi-square p-value for all the independent variables is depicted in Appendix A, Table A1. The weakness, history of bariatric surgery, headache, abdominal pain, diarrhea, and expectoration were deleted from the dataset as they have a p-value greater than 0.05 (accepted null hypothesis); thus, it shows no difference between the observed variables.

A backward elimination wrapper method determines the relevant predictors with continuous data [40, 41]. Logistic Regression (LR) was used as a statistical method to find the p-value [42]. It determines how likely a variable will contribute to a target variable under a particular hypothesis. The null hypothesis is that there is no effect on the target variable based on the other selected variables, which is accepted when the p-value is greater than 0.05. The variables are removed when the null hypothesis is accepted. Suppose the variable age has a p-value of 0.001; age is confirmed to contribute to the target variable ICU admission prediction and selected for the classification model. The variable selection procedure is shown

in the flowchart (Fig 4). These selected variables are used for the external validation test model. The complete details of LR results are depicted in Appendix A Table A2.

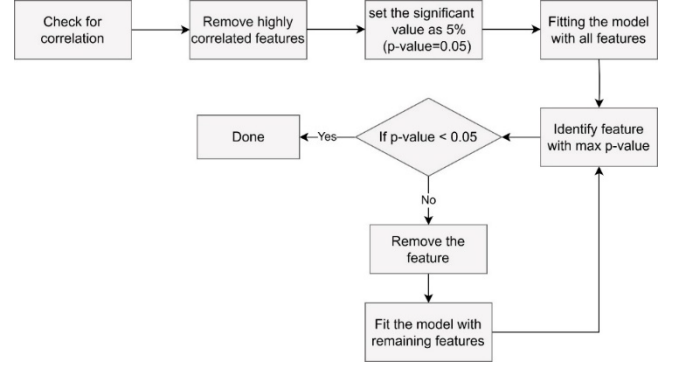


Fig. 4. Flowchart showing the variable selection procedure.

F. Machine Learning Algorithms and Evaluation

The data with selected variables is considered for the train-test-split method. The data is highly imbalanced, and the SMOTE method was employed to solve the issue. It creates more labels for the minority class by over-sampling, thus increasing model performance. It generates new instances in the minority class (ICU admission required) based on the variable space by interpolating positive samples lying next to each other. It was performed only on the CV model in this study. The performance of an ML model was evaluated using a 3-fold CV with a grid-search technique, which also helps determine the ideal combination of hyperparameters to utilize. It does this by first dividing the dataset into three folds, then training the model on two of those folds, and then testing it on the one-fold that is left over. This procedure was performed three times, and the findings were determined to be optimum after each fold was considered for assessing the model's performance. The ML algorithms used in this research were SVM with linear and radial basis function (RBF) kernels, RF, DT, LR, and ET models [43]. The best parameter was picked based on accuracy and then used to evaluate the dataset with a stratified 3-fold CV to prevent the over-fitting problem. The bias and variance of the model were also assessed to ensure the model does not overfit. Python with the Scikit-Learn library was used in this work [44].

Metrics such as precision or positive predictive rate, recall or sensitivity, specificity, accuracy, and F1-score were used to evaluate the predicted model. The accuracy shows how well the model performs across classes. However, class distribution imbalances may make it a poor metric. Suppose the number of non-ICU admissions far outnumbers the number of ICU admissions. In that case, the model may overlook the ICU admission instances by forecasting everything as non-ICU. In unbalanced class distributions, accuracy, recall, specificity, and F1-score may give more essential insights into the model's performance. The evaluation was based on the equation provided below.

$$Precision = \frac{trPs}{trPs + faPs} \quad (3)$$

$$Recall = \frac{trPs}{trPs + faNg} \quad (1)$$

$$F1 - Score = \frac{2trPs}{2trPs+faPs+faNg} \quad (2)$$

$$Accuracy = \frac{trPs+trNg}{trPs+trNg+faPs+faNg} \quad (3)$$

$$Specificity = \frac{trNg}{trNg+faPs} \quad (4)$$

where trPs is true positive, faPs is false positive, faNg is false negative, and trNg is true negative. Precision measures how well the model can predict who will be admitted to the ICU. A model with high precision is good at reducing false positives, which means that it correctly predicts ICU admission, the positive class. Recall measures how well the model can catch all important ICU entry cases. A high recall implies that the model predicts most ICU patient cases well. Specificity measures its ability to prevent false negatives. High specificity means the model can reliably detect non-ICU admissions by limiting false positives for negative instances. The F1-score is a harmonic means of the precision and recall of values of the model.

The receiver operating characteristic (ROC) curve is a practical approach for binary classification assessment. The performance of each model was examined for each threshold by plotting the true positive rate and the false positive rate. The model's performance improves when the area under the curve (AUC) metric approaches one [45]. Then, SHAP is used to find the contribution of each variable toward the prediction [46]. The SHAP library, specifically, supports most ML libraries, including Scikit-Learn, TensorFlow, and PyTorch.

III. RESULTS

A. Data Analysis

The total data collected was 4399, of which 31% (1366) were female, and 69% (3033) were male. The total number of ICU admissions required is 281 (6%), whereas 4118 (94%) did not require admission. The missing values were identified and updated using KNN imputing. The imputed and missing data sets were analyzed using the Hosmer and Lemeshow Test in logistic regression analysis. If the p-value exceeds 0.05, the null hypothesis can be accepted; thus, the model fits the data. Table 4 shows the study of missing and imputed data. The Hosmer-Lemeshow test indicates that the model fits the missing and imputed data well regarding the specified goodness-of-fit test.

TABLE IV

HOSMER-LEMESHOW TESTS FOR GOODNESS-OF-FIT TEST WITH MISSING AND IMPUTED DATA

Dataset	Chi-square	df	Sig. (p-value)
With missing	6.086	8	0.638
With imputed	2.583	8	0.958

Using the min-max function in Scikit-Learn, the data was standardized according to the mean and variance of 60% of the data (2639 patient data). The data count for the target label was ICU admission required=172 and ICU not required=2467. Here, under-sampling was not preferred as it would lose some relevant information. So, by using the SMOTE method, over-sampling was performed. The minority class was ICU-required

and over-sampled to the majority class's size (ICU admission not required). The GridSearchCV was used to tune the parameters, and the best parameter was saved for each model. Afterward, the dataset was evaluated in a stratified 3-fold CV manner with the best parameter for each classifier.

B. Model Performance with Variable Selection

The model was implemented based on selected variables in 60% of the data. The highly correlated variables were removed, and the backward elimination method using the logistic regression was employed, as shown in Fig 4. Variables were reduced to 20 from 65 according to the significant value (less than 0.05). Fig 5 depicts the selected predictors and their distribution. Gender, age, nationality, DBP, CXR, sore throat, SOB, fever, T.Protein, T.Bili, S.Albumin, respiratory rate, pulse rate, pulse oximeter, LDH, D-dimer, WBCs, HCT, MCHC, and DM history were the selected variables.

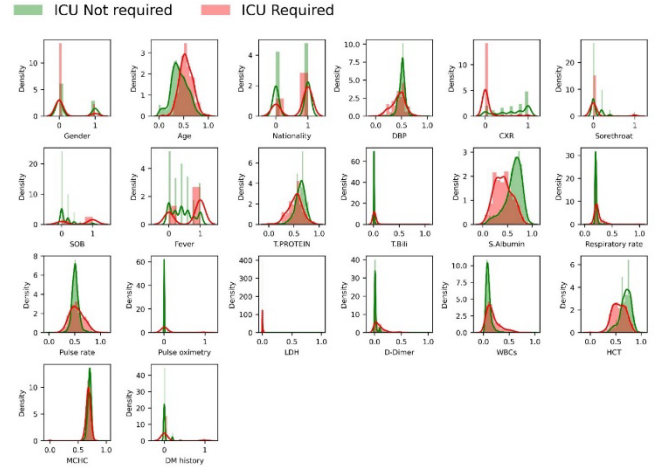


Fig. 5. Distribution plot of the selected variables.

The models were tuned with Grid search 3-fold CV, and the best parameter was chosen for fitting each classifier based on accuracy scoring. The tuned parameter is depicted in Table 5. The best parameter for the DT model was max depth=20 and alpha=0.001. The parameters C with 24 and gamma with values 2-2 showed the best accuracy for SVM with RBF kernel. The number of estimators for the ET classifier was 100, and the RF classifier was 350 and used max feature as log2. Similarly, the attributes were adjusted for LR classifiers, with the regularization parameter C=2 and the tolerance value (tol) 0.001. The SVM was also trained according to the Linear kernel, and the tuned parameters were C=2 and tolerance value (tol) 0.0001.

TABLE V

ML MODELS PARAMETERS FOR GRID-SEARCH WITH 3-FOLD CROSS-VALIDATION

Model	Tuning parameters	Range
DT	Max-depth	From 2 to 22
	Alpha	From 0.1 to 0.0001
SVM-Linear	C	From 2^{-1} to 2^4
	tol	From 0.1 to 0.0001
RF	N-estimator	From 100 to 500

	Max-features	Auto, sqrt, log ₂
LR	C	0.01, 0.1, 1, 2
	tol	From 0.1 to 0.00001
SVM-RBF	C	From 2 ⁰ to 2 ⁴
	Gamma	From 2 ⁻² to 2 ⁻⁵
ET	N-estimator	From 100 to 500
	Max-features	Auto, sqrt, log ₂

The evaluation results with all metrics for each ML model are shown in Table 6. The model was trained and evaluated based on the best-tuned parameters for each model. The ET classifier achieved the highest accuracy, with 98.09%, and showed better precision (0.9816), recall or sensitivity (0.9809), F1-score (0.9813), and specificity (0.9809).

TABLE VI

FULL EVALUATION METRICS OF ALL MODELS IN THE CROSS-VALIDATION OF 3-FOLD

Model	Acc \pm std.	Pre \pm std.	Rec \pm std.	F1 \pm std.	Spe \pm std.
DT	0.956 \pm 0.014	0.957 \pm 0.019	0.956 \pm 0.019	0.956 \pm 0.019	0.956 \pm 0.019
LR	0.901 \pm 0.022	0.901 \pm 0.018	0.901 \pm 0.022	0.901 \pm 0.020	0.901 \pm 0.022
SVM-Linear	0.913 \pm 0.001	0.914 \pm 0.021	0.913 \pm 0.026	0.913 \pm 0.023	0.913 \pm 0.026
SVM-RBF	0.959 \pm 0.001	0.961 \pm 0.024	0.959 \pm 0.026	0.959 \pm 0.025	0.959 \pm 0.026
ET	0.981 \pm 0.003	0.982 \pm 0.017	0.981 \pm 0.018	0.981 \pm 0.018	0.981 \pm 0.018
RF	0.974 \pm 0.003	0.974 \pm 0.016	0.974 \pm 0.017	0.974 \pm 0.016	0.974 \pm 0.017

The bias and variance of the CV models are depicted in Fig 6. Using a low-bias model will lead to overfitting, indicating that new data predictions will not be accurate. A model with a high bias will underfit the data, causing it to make consistently bad predictions. The RF and ET are ensemble models with the lowest variance compared to other models. The SVM model fits more complex data distribution. The LR and SVM-Linear had high bias; hence, they are less sensitive to noise in the dataset but less accurate on training data. The learning curve of the model with misclassification error or loss is shown in Fig 7.

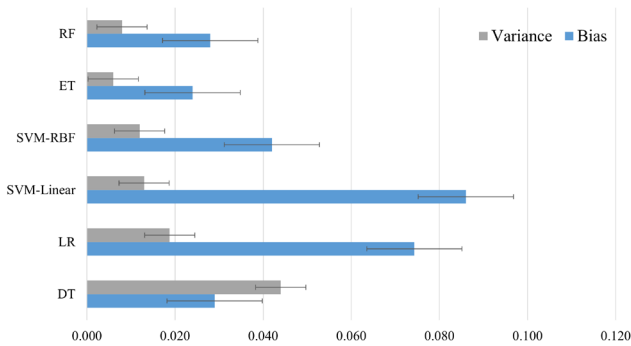


Fig. 6. Training and validation error of models using cross-validation.

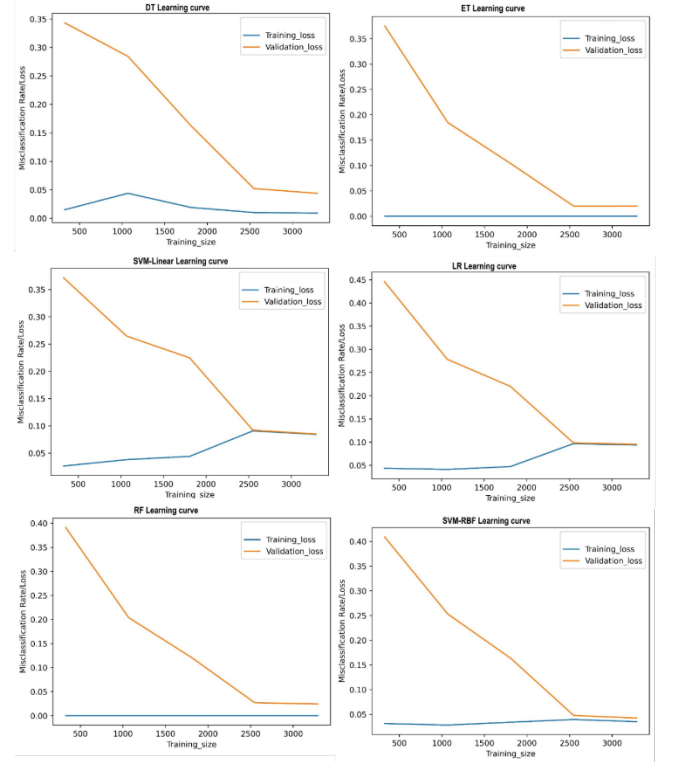


Fig. 7. The learning curve of a good fit model.

The models were tested externally with new unseen data (40% of actual data). The selected variables using the variable selection method were applied to the latest test data with preprocessing. The SMOTE method was not applied to the unseen data because a real dataset cannot guarantee that the data is balanced; hence, the unseen data was unbalanced. The unseen data has an unbalanced target variable with 1651 patient data for ICU admission not required, and 109 patient counts for ICU admission required. As the external test data was imbalanced, the recall or sensitivity, precision, F1-score, and specificity values were measured. The ET has better accuracy with 96.42%, whereas LR and SVM show high precision values with 0.99. The complete evaluation results are displayed in Table 7.

TABLE VII

EXTERNAL VALIDATION OF THE TESTED MODELS ON 40% UNSEEN IMBALANCED DATA

Model	Accuracy	Precision	Recall	F1-score	Specificity
DT	0.9290	0.9769	0.9467	0.9616	0.6606
LR	0.8972	0.9940	0.8958	0.9423	0.9174
SVM-Linear	0.8989	0.9913	0.9001	0.9435	0.8807
SVM-RBF	0.9449	0.9905	0.9503	0.9700	0.8624
ET	0.9642	0.9853	0.9764	0.9808	0.7798
RF	0.9574	0.9864	0.9679	0.9771	0.7982

With varying classification thresholds, the ROC curve shows the trade-off between true (sensitivity) and false positives. The greater the AUC, the better the model can distinguish between the two classes. Fig 8 shows the AUC score of all ML models

based on unseen test data. The SVM-RBF and LR show better AUC scores, with 0.91 each.

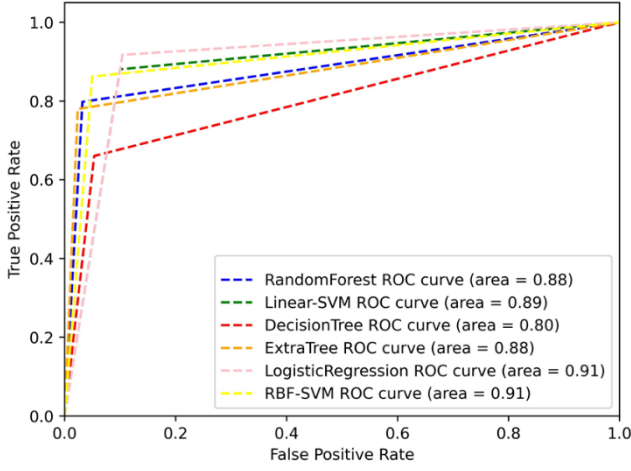


Fig. 8. The receiver operating characteristic curve with AUC score for unseen test data.

IV. DISCUSSION

The main goal of this study was to predict ICU admission requirements based on some clinical parameters of COVID-19 patients. The variables were selected based on LR with a backward selection method. Removing unnecessary or redundant variables allows variable selection approaches to decrease the number of independent variables. The list of variables is then reduced to those most essential to the ML models, helping to train the model faster, improve its performance, and reduce the chance of overfitting the data [47, 48]. Gender, age, nationality, DBP, CXR, sore throat, SOB, fever, T.Protein, T.Bili, S.Albumin, respiratory rate, pulse rate, pulse oximeter, LDH, D-dimer, WBCs, HCT, MCHC, and DM history were the selected variables.

When predicting ICU admissions, it is important to include the AUC-ROC score as a crucial assessment parameter, along with recall, accuracy, F1-score, and other metrics. The AUC score quantifies the discriminatory power of a model in distinguishing between patients who require ICU admission and those who do not, using all feasible categorization levels. A greater AUC value implies that the model is better at accurately distinguishing between different categories. In contrast to accuracy, recall, or specificity, which are influenced by a particular decision threshold, the AUC-ROC score assesses the performance of a model over all potential thresholds. This is especially valuable in medical scenarios such as predicting ICU admissions, where the ideal threshold may require adjustment depending on the clinical circumstances or the availability of ICU resources.

The ET model has superior performance in predicting ICU admissions, achieving the greatest accuracy of 97.64% and the highest F1-score of 0.9808. The SVM-RBF model has higher precision and specificity in comparison to the ET model, indicating a more cautious approach in forecasting ICU admissions. Consequently, it yields fewer false positives but may potentially overlook more real positives than the ET model.

The SVM-RBF model had the greatest AUC score of 0.91, suggesting its superior ability to differentiate between patients who require ICU admission and those who do not, regardless of the threshold used. These findings indicate that the SVM-RBF model, though not having the highest accuracy (but shows good performance with 94.49% accuracy), achieves the optimal trade-off between sensitivity (true positive rate) and specificity (true negative rate) among the studied models.

SVM is the most popularly used ML model [49]. If the classes are separated, SVM performs well and is not subject to overfitting. It can also handle more high-dimensional data than other ML models [50]. The model predicts the best hyperplane between the classes. The selection of ML models such as DT, RF, ET, SVM, and others varied according to the data size, problems, and variable sets [51]. The complete details of the previous work related to ML algorithms are mentioned in Table 8.

TABLE VIII

COMPARISON OF ML MODELS FOR PREDICTING ICU ADMISSION

Reference and year	Model	Dataset type	Results
[21] 2023	LR	532 COVID-19 patients with dichotomous variables	AUC = 0.748
[52] 2023	ANN	248 COVID-19 patients with five variables	95.97% accuracy
[23] 2022	DT algorithm J-48	512 COVID-19 patients with 53 variables	81.9% accuracy
[28] 2021	Ensemble model	5308 COVID-19 patients	F1-score = 0.81
[29] 2021	Multivariable LR	356 COVID-19 patients	AUC = 0.77
[30] 2021	RF	212 COVID-19 patients	AUC = 0.80
[31] 2021	Ensemble model	1218 COVID-19 patients with CBC data	AUC = 0.88
[32] 2021	XGB	3623 COVID-19 patients with 165 variables	AUC = 0.83
Proposed Study	SVM-RBF	4399 COVID-19 patients with 20 variables	AUC=0.91

The world health sector has faced difficulties due to the COVID-19 epidemic, creating a lack of ICU resources and health workers. Creating AI programs to identify whether a patient needs ICU admission helps the health workers arrange the resources. Many studies predicted ICU admission. Most are based on demographic details, blood results, albumin, oxygen tests, chronic diseases, and CXR to evaluate intensive care [29-31, 53]. Even the dataset was updated during the hospitalization, which helps predict early risk detection [24]. The mortality or survival prediction studies specify that variables such as potassium, chloride, oxygen level, blood

count test, D-dimer, and age are significant predictors [25, 28, 54, 55].

A study was done by Famiglini et al. [31] using SMOTE, but the predictors were based only on blood count data. The model used an ensemble algorithm with a combination of extreme GB (XGB), RF, and LR classifiers and achieved a score of 0.88 area under the curve (AUC) [31]. The paper by Shanbehzadeh et al. shows the prediction with the DT algorithm with AUC 0.822 [23]. The chi-square test predicts the relevant variables and found that the top predictors are thromboplastin time, age, and diabetes. Another study by the authors in the paper (9) indicates that absolute eosinophil counts from the CBC results are the best predictors. Age, pH, oxygen saturation, and chloride were significant predictors of ICU admission [54]. The AUC score achieved 0.917 using an artificial neural network (ANN). In the cases of pregnant women, the variable body mass index (BMI), CRP, neutrophil %, and respiratory level help the model predict the necessity of ICU admission [26]. Using a graphical calculating device, a nomogram, the authors in the paper [56] described that the model efficiently predicts the ICU requirement based on five predictors: CRP, AST, LDH, and platelet count.

Subudhi et al. [28] made the prediction model with an RF classifier with important markers such as CRP, most clinical blood results, oxygen level, chloride, D-dimer, and procalcitonin. According to the papers [32, 57], the XGB algorithm predicts better and has an AUC of 0.98 and 0.83, respectively. High albumin present in the body can also factor in ICU admittance [58]. Additionally, DM history, age, gender, congestive heart failure (CHF), and stroke also help in prediction [59]. The LR model shows a better forecast with an AUC of 0.74, according to the paper [60], and the significant markers are oxygen level, procalcitonin, lymphocyte count, and LDH. In addition to these factors, CRP and ferritin are also included using the deep learning model [61]. Rather than using clinical variables, the paper [62] used the National Early Warning Score and Rapid Emergency Medicine Score to predict ICU admittance.

The variable contribution for prediction is identified using the SHAP plot. The plot is shown in Fig 9. The 20 selected variables are evaluated to determine their contribution to the target prediction. The SHAP plot also shows the variable importance and the impact on the prediction variable. A variable showing the positive side of the SHAP value positively contributes to the ICU admission requirement. The blue (low value) and red (high value) indicate how the variables contribute to predicting ICU admission concerning their values. The plot is based on the forecast that ICU admission is required. The variables are arranged as most important from top to bottom. The top three predictors are CXR, d-dimer, and SOB. The low and high CXR values positively and negatively contribute to ICU prediction, respectively. Thus, having a normal X-ray reduces the need for ICU admission because a normal CXR is coded as one and abnormal as zero. The d-dimer and SOB have a highly positive prediction when they have high values. For example, suppose a patient has a high d-dimer; in that case, he can be considered for ICU admission by looking into other variables. A high value of S.Albumin, pulse oximeter, HCT, T.Protein, sore throat, LDH, gender, DBP, and MCHC

have a high negative contribution. Hence, it contributes much towards unrequired ICU admissions. On the other hand, fever, respiratory rate, age, WBC, nationality, T.Bili, pulse rate, and DM history have a high positive contribution when the values are high; hence, they show that the high value of these variables positively contributes to ICU admission.

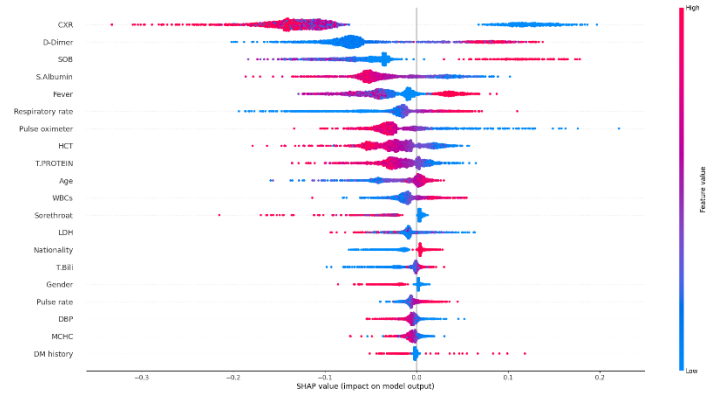


Fig. 9. SHAP value summary plot of 20 important variables for ICU admission required prediction.

Certain illnesses and treatments need gender information. For example, men have cardiac problems in common, whereas women have autoimmune illnesses. Age has a significant role in the occurrence of many illnesses. When thinking about specific diseases or treatments, knowing a patient's nationality might be a vital piece of information. Having high blood pressure raises the likelihood of developing heart disease. CXR helps identify pneumonia and other disorders. Fever, sore throat, and SOB might indicate viruses, infections, or allergies. When red blood cells (RBCs) die, they release bilirubin. Albumin maintains blood volume and pressure. Low albumin, total protein levels, and a high bilirubin level can be signs of kidney failure or liver disease. ICU admission is necessary if the patient has high respiratory and pulse rates and low pulse oximeters. Elevated d-dimer readings might indicate a blood clot. When cells are injured, they produce an enzyme called LDH. Elevated LDH levels have been linked to cardiac arrest and stroke. Patients with high diabetes are also at higher risk for various diseases. Low levels of HCT and MCHC in the blood can indicate anemia. In contrast, a high WBC count indicates a sign of infection.

The SHAP plot (Fig 9) shows that non-Kuwaitis are likelier to be admitted to ICU, probably due to economic, cultural, or other factors. Race and ethnicity vary among non-Kuwaitis, and they have health disparities due to heat and air pollution [63]. This study supports the findings of Lat et al. [64] that men are more likely to be admitted to the ICU than women are. In this study, gender and nationality significantly affect ICU admission ($p < 0.001$). Larger-scale experiments are still needed to verify these results, which can be considered a future enhancement.

The dataset was again divided according to the random glucose level to check how the model predicts ICU admission for diabetic patients. It analyses how patients with and without diabetes require the ICU. Out of the COVID-19 patient's random glucose levels, 61% were low, and 39% had a high value and belonged to the diabetic group. For the dataset with

diabetic patients, the target variable (ICU prediction) achieved 98.60% with the SVM-RBF model and 97.83% with the non-diabetic dataset. The variable contribution based on diabetic patients' prediction of ICU admission is depicted in Fig 10. The top 10 variables for diabetic patients that contribute to the prediction of ICU admission required are a low value of CXR, high value of SOB, d-dimer, lymphocytes, age, respiratory rate, CRP, low value of T.Protein, and platelet count.

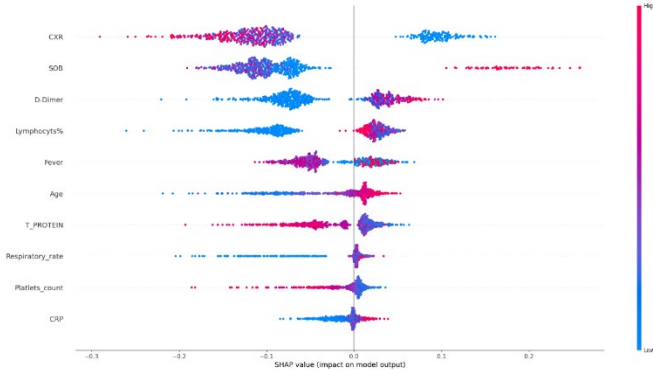


Fig. 10. Variable importance for the prediction of ICU admission based on the diabetic patient dataset.

The proposed study has certain strengths. One is the methodology used. The CV solves the overfitting problem, and SMOTE resolves the imbalance problem. Hence, the majority and minority class problems are not misclassified. The second is that the study uses almost all specific health results, including demographic details, for prediction. For the variable importance and contribution, the SHAP plot is analyzed. For missing data, the study employs the KNN imputation method. Not all variables are necessary for the final prediction. So, backward variable elimination determines the relevant variables. The developed SVM model could help notify the health workers in preparing ICU resources at the proper time for patients with COVID-19. As a result, the resources can be utilized appropriately.

V. CONCLUSION

In this study, the data is leveraged from Kuwaiti hospitals to predict ICU admissions for COVID-19 patients, employing preprocessing techniques, class distribution balancing via SMOTE, variable selection through backward selection with logistic regression, and missing value imputation with KNN. The findings, validated by the Hosmer-Lemeshow test, demonstrate the superior performance of the SVM classifier, achieving 94.49% accuracy, AUC of 0.91, and an F1-score of 0.97. This research not only elucidates the critical predictors for ICU admission but also showcases the model's decision-making process, aligning with the discovery science's goal of enhancing interpretability in machine learning models. The model's ability to effectively allocate ICU resources during disease outbreaks exemplifies the potential of ML in improving healthcare responses and resource optimization, contributing significantly to the fields of computational scientific discovery and intelligent data analysis. By identifying patients at higher risk for ICU admission, the approach assists healthcare professionals in resource allocation, emphasizing the broader

applicability of such ML-based models in discovery science across various scientific domains. This work underscores the importance of machine learning in facilitating scientific discoveries and enhancing healthcare strategies, promising substantial advancements in both discovery science and applied healthcare.

A. Limitations and Future Works

The few hospitals from which the data is collected form one of the limitations of this work. Collecting more data from different Gulf Cooperation Council hospitals is a future enhancement, thereby implementing the model with the deep learning concept. Many variable selection methods exist, such as filter, wrapper, and embedded. So, implementing and comparing the proposed model with different variable selection methods is another future work. Another future enhancement is to generate a model based on data from Kuwaiti diabetes patients and check whether it is clinically significant.

APPENDIX A

TABLE A.I

CHI-SQUARE P-VALUE FOR THE ASSOCIATION OF INDEPENDENT CATEGORICAL VARIABLE

Variables	Gender	Nationality	CXR	Weakness	Headache	Sore throat
Gender	-					
Nationality	<0.001	-				
CXR	0.006	<0.001	-			
Weakness	0.126	0.087	<0.001	-		
Headache	0.016	0.275	<0.001	<0.001	-	
Sore throat	<0.001	<0.001	<0.001	<0.001	<0.001	-
SOB	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Abdominal pain	0.016	0.351	<0.001	<0.001	<0.001	0.255
Fever	<0.001	0.002	<0.001	<0.001	<0.001	<0.001
Expectoration	0.037	0.076	<0.001	<0.001	<0.001	<0.001
Diarrhea	<0.001	0.006	<0.001	<0.001	<0.001	<0.001
Cough	<0.001	0.002	<0.001	<0.001	<0.001	<0.001
Asthma	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Bariatric Surgery	0.015	0.122	0.316	<0.001	<0.001	<0.001
Hypertension	0.003	<0.001	<0.001	<0.001	<0.001	<0.001
DM history	<0.001	0.004	<0.001	<0.001	0.151	<0.001
DM	<0.001	<0.001	<0.001	<0.001	0.032	<0.001

TABLE A.I (continued)

CHI-SQUARE P-VALUE FOR THE ASSOCIATION OF INDEPENDENT CATEGORICAL VARIABLE

Variables	SOB	Abdominal pain	Fever	Expectoration	Diarrhea	Cough
Gender						
Nationality						
CXR						

Weakness						
Headache						
Sore throat						
SOB	-					
Abdominal pain	<0.001	-				
Fever	<0.001	<0.001	-			
Expectoration	<0.001	0.601	<0.001	-		
Diarrhea	<0.001	<0.001	<0.001	0.001	-	
Cough	<0.001	<0.001	<0.001	<0.001	<0.001	-
Asthma	<0.001	0.279	<0.001	0.001	<0.001	<0.001
Bariatric Surgery	0.172	<0.001	<0.001	0.856	0.177	<0.001
Hypertension	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
DM history	<0.001	0.843	<0.001	<0.001	0.107	<0.001
DM	<0.001	0.026	<0.001	<0.001	0.269	<0.001

TABLE A.I (continued)

CHI-SQUARE P-VALUE FOR THE ASSOCIATION OF INDEPENDENT CATEGORICAL VARIABLE

Variables	Asthma	Bariatric Surgery	Hypertension	DM history
Gender				
Nationality				
CXR				
Weakness				
Headache				
Sore throat				
SOB				
Abdominal pain				
Fever				
Expectoration				
Diarrhea				
Cough				
Asthma	-			
Bariatric Surgery	0.004	-		
Hypertension	<0.001	0.033	-	
DM history	<0.001	0.991	<0.001	-
DM	<0.001	0.992	<0.001	<0.001

TABLE A.II

THE LOGISTIC REGRESSION RESULTS ON THE BACKWARD FEATURE ELIMINATION METHOD

Variable	Coef	Std err	p-value	97% CI	
				Lower	Upper
Gender	-0.887	0.235	0.000	-1.347	-0.428
Age	2.528	0.617	0.000	1.319	3.737
Nationality	0.820	0.200	0.000	0.429	1.212
DBP	-2.161	0.841	0.010	-3.811	-0.513
CXR	-3.277	0.430	0.000	-4.120	-2.434
Sore throat	-1.331	0.426	0.002	-2.165	-0.497
SOB	1.594	0.213	0.000	1.178	2.011
Fever	0.441	0.221	0.045	0.009	0.874
T.Protein	-2.509	0.690	0.000	-3.861	-1.158
T.Bili	5.794	2.729	0.034	0.446	11.143
S.Albumin	-2.050	0.503	0.000	-3.036	-1.065
Respiratory rate	6.185	1.703	0.000	2.847	9.524

Respiratory rate					
Pulse rate	2.426	0.829	0.003	0.801	4.052
Pulse oximeter	3.216	0.798	0.000	1.653	4.781
LDH	-47.636	12.031	0.000	-71.216	-24.055
D-Dimer	4.441	1.160	0.000	2.168	6.714
WBCs	2.425	0.862	0.005	0.735	4.116
HCT	-1.778	0.789	0.024	-3.325	-0.233
MCHC	-3.739	1.261	0.003	-6.212	-1.267
DM history	2.171	0.466	0.000	1.257	3.085

ACKNOWLEDGMENT

The author thanks all the people involved in this study and the clinicians from the hospitals in Kuwait.

STATEMENT REGARDING ETHICAL APPROVAL

The Permanent Committee for Coordination of Medical and Health Research at the Ministry of Health, Kuwait, approved the study for sample collection and survey of the patients with consent forms (approval number: 559, dated 03-05-2020).

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

DATA AVAILABILITY

The data required to reproduce the above findings cannot be shared for ethical reasons.

CONFLICT OF INTEREST

The author has no conflict of interests to declare.

REFERENCES

- [1] Worldometer. "COVID Live - Coronavirus Statistics." <https://www.worldometers.info/coronavirus/> (accessed 27 July 2022, 2022).
- [2] A. Supady *et al.*, "Allocating scarce intensive care resources during the COVID-19 pandemic: practical challenges to theoretical frameworks," *The Lancet Respiratory Medicine*, vol. 9, no. 4, pp. 430-434, 2021, doi: 10.1016/S2213-2600(20)30580-4.
- [3] M. Cascella, M. Rajnik, A. Aleem, S. Dulebohn, and R. Di Napoli. "Features, evaluation, and treatment of coronavirus (COVID-19)." <https://www.statpearls.com/ArticleLibrary/viewarticle/52171> (accessed 2 August 2023).
- [4] W. M. El-Sadr, A. Vasan, and A. El-Mohandes, "Facing the New Covid-19 Reality," *New England Journal of Medicine*, vol. 388, no. 5, pp. 385-387, 2023, doi: 10.1056/NEJMp2213920.
- [5] C. Foley, M. Bloomer, and A. M. Hutchinson, "Factors that influence intensive care admission decisions for older people: A systematic review," *Australian Critical Care*, vol. 36, no. 2, pp. 274-284, 2023/03/01/ 2023, doi: 10.1016/j.aucc.2021.12.006.
- [6] N. Asmari *et al.*, "Machine learning-derived analysis of intensive care unit mortality in patients with ICU acquired infections," *Journal of Critical Care*, vol. 81, p. 154746, 2024/06/01/ 2024, doi: <https://doi.org/10.1016/j.jccr.2024.154746>.
- [7] P. K. Olds, N. Musinguzi, B. P. Geisler, and J. E. Haber, "Evaluating disparities by social determinants in hospital admission decisions for patients with COVID-19 quaternary hospital early in the pandemic," *Medicine*, vol. 102, no. 10, p. e33178, 2023, doi: 10.1097/md.00000000000033178.
- [8] T. S. Valley *et al.*, "Hospital factors that influence ICU admission decision-making: a qualitative study of eight hospitals," *Intensive Care*

- Medicine*, vol. 49, no. 5, pp. 505-516, 2023/05/01 2023, doi: 10.1007/s00134-023-07031-w.
- [9] K. Miyashita *et al.*, "Changes in the characteristics and outcomes of COVID-19 patients from the early pandemic to the delta variant epidemic: a nationwide population-based study," *Emerging Microbes & Infections*, vol. 12, no. 1, p. 2155250, 2023/12/31 2023, doi: 10.1080/22221751.2022.2155250.
 - [10] S. P. Ungar *et al.*, "Hospital and ICU Admission Risk Associated With Comorbidities Among Children With COVID-19 Ancestral Strains," *Clinical Pediatrics*, vol. 0, no. 0, p. 00099228221150605, doi: 10.1177/00099228221150605.
 - [11] A. Alshukry *et al.*, "Clinical characteristics and outcomes of COVID-19 patients with diabetes mellitus in Kuwait," *Heliyon*, vol. 7, no. 4, p. e06706, 2021.
 - [12] H. Ali, A. Al-Shammari, B. Alahmad, and F. Al-Mulla, "Advancing risk analysis of COVID-19 clinical predictors: the case of fasting blood glucose," *Frontiers in Medicine*, vol. 8, p. 636065, 2021.
 - [13] J. M. Dennis *et al.*, "Type 2 Diabetes and COVID-19-Related Mortality in the Critical Care Setting: A National Cohort Study in England, March–July 2020," *Diabetes Care*, vol. 44, no. 1, pp. 50-57, 2020, doi: 10.2337/dc20-1444.
 - [14] H. Sourij *et al.*, "COVID-19 fatality prediction in people with diabetes and prediabetes using a simple score upon hospital admission," (in eng), *Diabetes Obes Metab*, vol. 23, no. 2, pp. 589-598, Feb 2021, doi: 10.1111/dom.14256.
 - [15] R. Anand *et al.*, "Predicting Mortality in Diabetic ICU Patients Using Machine Learning and Severity Indices," *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, vol. 2017, pp. 310-319, 05/18 2018.
 - [16] J. Ye, L. Yao, J. Shen, R. Janarthanam, and Y. Luo, "Predicting mortality in critically ill patients with diabetes using machine learning and clinical notes," *BMC Medical Informatics and Decision Making*, vol. 20, no. 11, p. 295, 2020/12/30 2020, doi: 10.1186/s12911-020-01318-4.
 - [17] A. Alkandari *et al.*, "Adult diabetes and prediabetes prevalence in Kuwait: data from the cross-sectional Kuwait diabetes epidemiology program," *Journal of clinical medicine*, vol. 9, no. 11, p. 3420, 2020.
 - [18] A. Núñez Reiz, M. A. Armengol de la Hoz, and M. Sánchez García, "Big Data Analysis and Machine Learning in Intensive Care Units," *Medicina Intensiva (English Edition)*, vol. 43, no. 7, pp. 416-426, 2019/10/01/ 2019, doi: 10.1016/j.medic.2019.06.012.
 - [19] A. Ö. Yeşilyurt *et al.*, "Prediction of mortality in COVID-19 acute respiratory distress syndrome patients during the critical first three days in ICU," *Journal of Critical Care*, vol. 81, p. 154674, 2024/06/01/ 2024, doi: <https://doi.org/10.1016/j.jcrc.2024.154674>.
 - [20] T. Huang, D. Le, L. Yuan, S. Xu, and X. Peng, "Machine learning for prediction of in-hospital mortality in lung cancer patients admitted to intensive care unit," *Plos one*, vol. 18, no. 1, p. e0280606, 2023.
 - [21] P. Wendland *et al.*, "Machine learning models for predicting severe COVID-19 outcomes in hospitals," *Informatics in Medicine Unlocked*, p. 101188, 2023.
 - [22] Z. A. Varzaneh, A. Orooji, L. Erfannia, and M. Shanbehzadeh, "A new COVID-19 intubation prediction strategy using an intelligent feature selection and K-NN method," *Informatics in Medicine Unlocked*, vol. 28, p. 100825, 2022, doi: 10.1016/j.imu.2021.100825.
 - [23] M. Shanbehzadeh, R. Nopour, and H. Kazemi-Arpanahi, "Using decision tree algorithms for estimating ICU admission of COVID-19 patients," *Informatics in Medicine Unlocked*, p. 100919, 2022.
 - [24] O. Noy *et al.*, "A machine learning model for predicting deterioration of COVID-19 inpatients," *Scientific reports*, vol. 12, no. 1, pp. 1-9, 2022.
 - [25] T. Kuno, Y. Sahashi, S. Kawahito, M. Takahashi, M. Iwagami, and N. N. Egorova, "Prediction of in-hospital mortality with machine learning for COVID-19 patients treated with steroid and remdesivir," *Journal of Medical Virology*, vol. 94, no. 3, pp. 958-964, 2022.
 - [26] E. Kalafat *et al.*, "An internally validated prediction model for critical COVID-19 infection and intensive care unit admission in symptomatic pregnant women," *American Journal of Obstetrics and Gynecology*, vol. 226, no. 3, pp. 403. e1-403. e13, 2022.
 - [27] C.-C. Chang *et al.*, "Utilization of Decision Tree Algorithms for Supporting the Prediction of Intensive Care Unit Admission of Myasthenia Gravis: A Machine Learning-Based Approach," *Journal of Personalized Medicine*, vol. 12, no. 1, p. 32, 2022.
 - [28] S. Subudhi *et al.*, "Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19," *NPJ digital medicine*, vol. 4, no. 1, pp. 1-7, 2021.
 - [29] S. Schalekamp *et al.*, "Model-based prediction of critical illness in hospitalized patients with COVID-19," *Radiology*, vol. 298, no. 1, pp. E46-E54, 2021.
 - [30] D. Patel *et al.*, "Machine learning based predictors for COVID-19 disease severity," *Scientific reports*, vol. 11, no. 1, pp. 1-7, 2021.
 - [31] L. Famiglini, G. Bini, A. Carobene, A. Campagner, and F. Cabitza, "Prediction of ICU admission for COVID-19 patients: a Machine Learning approach based on Complete Blood Count data," presented at the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), Aveiro, Portugal, 2021. [Online]. Available: <https://doi.org/10.1109/CBMS52027.2021.00065>.
 - [32] R. Aznar-Gimeno *et al.*, "A clinical decision web to predict ICU admission or death for patients hospitalised with COVID-19 using machine learning algorithms," *International Journal of Environmental Research and Public Health*, vol. 18, no. 16, p. 8677, 2021.
 - [33] X. Guo, Y. Yin, C. Dong, G. Yang, and G. Zhou, "On the Class Imbalance Problem," presented at the 2008 Fourth International Conference on Natural Computation, 18-20 Oct. 2008, 2008. [Online]. Available: <https://doi.org/10.1109/ICNC.2008.871>.
 - [34] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *Journal of Big Data*, vol. 6, no. 1, p. 27, 2019/03/19 2019, doi: 10.1186/s40537-019-0192-5.
 - [35] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321-357, 2002, doi: <https://doi.org/10.1613/jair.953>.
 - [36] A. Fernández, S. García, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *Journal of artificial intelligence research*, vol. 61, pp. 863-905, 2018, doi: 10.1613/jair.1.11192.
 - [37] A. D. A. P. P. Committee, "2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2022," *Diabetes Care*, vol. 45, no. Supplement_1, pp. S17-S38, 2021, doi: 10.2337/dc22-S002.
 - [38] O. Troyanskaya *et al.*, "Missing value estimation methods for DNA microarrays," *Bioinformatics*, vol. 17, no. 6, pp. 520-525, 2001.
 - [39] B. Ratner, "The correlation coefficient: Its values range between+ 1/- 1, or do they?," *Journal of targeting, measurement and analysis for marketing*, vol. 17, no. 2, pp. 139-142, 2009.
 - [40] M. Mehmood, N. Alshammari, S. A. Alanazi, and F. Ahmad, "Systematic Framework to Predict Early-Stage Liver Carcinoma Using Hybrid of Feature Selection Techniques and Regression Techniques," *Complexity*, vol. 2022, 2022, doi: 10.1155/2022/7816200.
 - [41] J. Park, K. Li, and H. Zhou, "k-fold Subsampling based Sequential Backward Feature Elimination," presented at the International Conference on Pattern Recognition Applications and Methods (ICPRAM), Rome, Italy, 2016. [Online]. Available: <https://doi.org/10.5220/0005688804230430>.
 - [42] T. G. Nick and K. M. Campbell, "Logistic Regression," in *Topics in Biostatistics*, W. T. Ambrosius Ed. Totowa, NJ: Humana Press, 2007, pp. 273-301.
 - [43] R. Kumari and S. K. Srivastava, "Machine learning: A review on binary classification," *International Journal of Computer Applications*, vol. 160, no. 7, 2017.
 - [44] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011. [Online]. Available: <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://>.
 - [45] J. Fan, S. Upadhye, and A. Worster, "Understanding receiver operating characteristic (ROC) curves," *Canadian Journal of Emergency Medicine*, vol. 8, no. 1, pp. 19-20, 2006, doi: 10.1017/s1481803500013336.
 - [46] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017, doi: 10.48550/arXiv.1705.07874.
 - [47] K. Mahalakshmi and P. Sujatha, "Critical Analysis of Feature Selection Methods for Data Preprocessing with Heart Disease Dataset," in *Data Intelligence and Cognitive Informatics*: Springer, 2022, pp. 667-682.
 - [48] N. Bharathi, B. Rishiikeshwer, T. A. Shriram, B. Santhi, and G. Brindha, "The Significance of Feature Selection Techniques in Machine Learning," *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools and Applications*, pp. 121-134, 2022.
 - [49] S. Yue, P. Li, and P. Hao, "SVM classification: Its contents and challenges," *Applied Mathematics-A Journal of Chinese Universities*, vol. 18, no. 3, pp. 332-342, 2003.

- [50] Y. Yang, J. Li, and Y. Yang, "The research of the fast SVM classifier method," presented at the 2015 12th international computer conference on wavelet active media technology and information processing (ICCWAMTIP), Chengdu, China, 2015. [Online]. Available: <https://doi.org/10.1109/ICCWAMTIP.2015.7493959>.
- [51] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," *The journal of machine learning research*, vol. 15, no. 1, pp. 3133-3181, 2014.
- [52] P. G. Asteris *et al.*, "Early prediction of COVID-19 outcome using artificial intelligence techniques and only five laboratory indices," *Clinical Immunology*, vol. 246, p. 109218, 2023.
- [53] J. Mushatq *et al.*, "Initial chest radiographs and artificial intelligence (AI) predict clinical outcomes in COVID-19 patients: analysis of 697 Italian patients," *European radiology*, vol. 31, no. 3, pp. 1770-1779, 2021.
- [54] R. Rigo-Bonnin *et al.*, "Individual outcome prediction models for patients with COVID-19 based on their first day of admission to the intensive care unit," *Clinical Biochemistry*, vol. 100, pp. 13-21, 2022.
- [55] M. Ayed, A. A. Borahmah, A. Yazdani, A. Sultan, A. Mossad, and H. Rawdhan, "Assessment of clinical characteristics and mortality-associated factors in COVID-19 critical cases in Kuwait," *Medical Principles and Practice*, vol. 30, no. 2, pp. 185-192, 2021.
- [56] Z. Zeng *et al.*, "Simple nomogram based on initial laboratory data for predicting the probability of ICU transfer of COVID-19 patients: multicenter retrospective study," *Journal of medical virology*, vol. 93, no. 1, pp. 434-440, 2021.
- [57] N. D. Parbate, A. P. Palande, S. K. Wagh, A. Kamble, and S. Isave, "ICU Admission Prediction Using Machine Learning for Covid-19 Patients," presented at the 2021 International Conference on Communication information and Computing Technology (ICCICT), Mumbai, India, 2021. [Online]. Available: <https://doi.org/10.1109/ICCICT50803.2021.9510117>.
- [58] M. Kheir, F. Saleem, C. Wang, A. Mann, and J. Chua, "Higher albumin levels on admission predict better prognosis in patients with confirmed COVID-19," *Plos one*, vol. 16, no. 3, p. e0248358, 2021.
- [59] R. Gunduz *et al.*, "CHA2DS2-VASc score and modified CHA2DS2-VASc score can predict mortality and intensive care unit hospitalization in COVID-19 patients," *Journal of thrombosis and thrombolysis*, vol. 52, no. 3, pp. 914-924, 2021.
- [60] Z. Zhao *et al.*, "Prediction model and risk scores of ICU admission and mortality in COVID-19," *PloS one*, vol. 15, no. 7, p. e0236618, 2020.
- [61] X. Li *et al.*, "Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables," *PeerJ*, vol. 8, p. e10337, 2020.
- [62] M. Covino *et al.*, "Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores," *Resuscitation*, vol. 156, pp. 84-91, 2020.
- [63] B. Alahmad *et al.*, "Extreme temperatures and mortality in Kuwait: Who is vulnerable?," *Science of The Total Environment*, vol. 732, p. 139289, 2020/08/25/ 2020, doi: 10.1016/j.scitotenv.2020.139289.
- [64] T. I. Lat, M. K. McGraw, and H. D. White, "Gender Differences in Critical Illness and Critical Care Research," (in eng), *Clin Chest Med*, vol. 42, no. 3, pp. 543-555, Sep 2021, doi: 10.1016/j.ccm.2021.04.012.

College of Engineering and Petroleum, Kuwait University, (2017- 2019), Editorial Board of PLOS-ONE a Q1 ISI Journal (2022-Present). A board member of the Arab Robotic Association from 2016-present, Founder and President of ROBOTEC consultations in IT and Technologies 2016-present, Assistant vice president of IT (2020-present), and acting Vice President for Academic Support Services 2020-present. He specializes in Artificial intelligence, and his current research interests are in Robotics, Expert Systems, Artificial Intelligence, Signal Processing, Pattern Recognition, Deep Learning, and e-learning.

Dr. Mutawa (Member, IEEE), received his Ph.D. from Syracuse University, New York, USA, in 1999 in Artificial Intelligence and joined the Computer Engineering department at Kuwait University (1999-present). International Academic Consultant for MILSET (2022-present) and UNESCO-UN for AI-Ethics (2021-present), Awarded Shk Salim Ali Award for Excellence in AI (2021). Mutawa also served as a Director of the Office of Engineering Education at the College of Engineering and Petroleum Kuwait University (2004-2012), Assistant Vice President for academic services for computer systems and distance learning (2012-2014), Director of E-Services and Engineering Education at