The prediction of mortality influential variables in an intensive care unit: a case study

Naghmeh Khajehali¹ · Zohreh Khajehali² · Mohammad Jafar Tarokh¹

Received: 21 February 2020 / Accepted: 10 February 2021 / Published online: 26 February 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd. part of Springer Nature 2021

Abstract

The intensive care units (ICUs) are among the most expensive and essential parts of all hospitals for extremely ill patients. This study aims to predict mortality and explore the crucial factors affecting it. Generally, in the health care systems, having a fast and precise ICU mortality prediction for patients plays a key role in care quality, resulting in reduced costs and improved survival chances of the patients. In this study, we used a medical dataset, including patients' demographic details, underlying diseases, laboratory disorder, and LOS. Since accurate estimates are required to have optimal results, various data pre-processings as the initial steps are used here. Besides, machine learning models are employed to predict the risk of mortality ICU discharge. For AdaBoost model, these measures are considered AUC= 0.966, sensitivity (recall) = 87.88%, Kappa=0.859, F-measure = 89.23% making it, AdaBoost, accounts for the highest rate. Our model outperforms other comparison models by using various scenarios of data processing. The obtained results demonstrate that the high mortality can be caused by underlying diseases such as diabetes mellitus and high blood pressure, moderate Pulmonary Embolism Wells Score risk, platelet blood count less than 100000 (mcl), hypertension (HTN), high level of Bilirubin, smoking, and GCS level between 6 and 9.

Keywords Mortality prediction · Intensive care unit · AdaBoost · Machine learning methods

1 Introduction

The intensive care units (ICUs) are considered for the cure of the patients who require extensive monitoring status and parameters of their health and physiological [1], so it demands expert personnel, state-of-the-art facilities, emergency equipment for treatment, monitoring patients admitted to this unit, and it is much more specialized than other sections [2]. As a result, this hospital unit tends to play an important role and is considered one of the most high-priced parts of all hospitals [3–5]. ICU mortality prediction plays a significant part in the patient's health care and allocation of hospital resources,

leading to the improvement of patient survival [1]. Furthermore, it supplies a method to assess the performance of various medical facilities, services, and differences among them. It not only contributes in declining discrepancies of health care but also reduces the hospital costs due to discharging patients from the ICU as soon as possible [6, 7]. A growing number of researchers allocated a great deal of time to study to improve the accuracy in identifying ICU patient mortality since a huge amount of data, and a variety of complex attributes are produced every day in ICUs. This can be challenging to develop new algorithms in ICU mortality prediction. New technologies based on machine learning are considered one of the well-known modeling methods, including artificial neural network (ANN), support vector machine (SVM), and random forest (RF). These algorithms are capable of predicting the results more precisely and correctly for patients than other methods. Many researchers use machine learning to develop models for predicting complex clinical scenarios like ICU settings because clinical attributes and outcomes have nonlinear relationships, so the machine learning algorithms can demonstrate these nonlinear relationships [8–10]. In this work, machine learning methods such as decision tree (DT), neural network (NN), K-nearest network (K-

✉ Naghmeh Khajehali
nkhajehali@mail.kntu.ac.ir

Zohreh Khajehali
zohrehkhajehali@gmail.com

Mohammad Jafar Tarokh
mjtarokh@kntu.ac.ir

¹ K. N. Toosi University of Technology, Tehran, Iran

² Isfahan University of Medical Science, Isfahan, Iran

NN), random forest (RF), linear regression(LR), AdaBoost, Bayesian Boosting, Vote (DT+K-NN), and Vote (DT+K-NN+LR) are employed to evaluate the rate of patients mortality who are admitted in ICUs. The prior studies used a rich set of features that is available at discharge. In contrast, we used commonly available data at admission time. The data includes the following (see Table 1): demographic information (age, gender, and skin type). The body mass index (BMI) is a value in which weight in kilograms (kg) relates to height in meters squared. It is divided into four groups by the National Institutes of Health (NIH): the normal weight which is 18.5 to 25 kg/m², obesity, overweight, and underweight which have the BMI of more than 30, 25, to 30 and less than 18.5

kg/m² respectively [11]. The Pulmonary Embolism Wells Score is a score of risk classification and clinical decision rule to predict the critical pulmonary embolism (PE) probable in patients who had the possibility of acute PE in their history and examination. A score is classified into three categories: a score greater than 6, a score between 2 and 6, and a score less than 2, considered “high risk,” “intermediate risk,” and “low risk” of PE, respectively [12]. The Wells score is a number that shows the risk of developing deep vein thrombosis (DVT) in the body. DVT occurs when blood clot forms in a vein that is deep inside the body. The Wells score is calculated based on several factors. Using this score, the doctor can determine the existence of DVT. This helps doctors in deciding whether to

Table 1 Medical variable along with its definition which is used in this work

Variable	Definition
Hemiplegia/paraplegia	Paraplegia is an impairment of the legs and lower body resulting from injury to nerves in the lumbar or thoracic vertebrae areas of the body. Hemiplegia is the impairment of one vertical half of the body.
Anemia	It is a condition when the body lacks enough healthy red blood cells to transfer sufficient oxygen to the body's tissues. The normal hemoglobin range for men, in general, is defined between 13.2–16.6 grams (g) of hemoglobin per deciliter (dL) of blood and between 11.6–15. G/dL for women.
Motor sensory disorder	It is a condition when the brain has trouble in the process of sensory information
Bilirubin	It is an orange-yellow substance in the body that is generally formed during the normal breakdown of red blood cells. Normal Bilirubin levels are less than 1.2 (mg/dL), and higher Bilirubin levels are an indicator of different types of liver problems.
Serum Creatinine	It is a waste substance in the blood that arises from the activity of muscle. Its normal level in the blood in adults is from 0.5 to 1.2 milligrams (mg) per deciliter (dL). Severe kidney impairment is a sign when Creatinine levels reach 2.0 or more in adults
Platelet blood count	Platelet blood test measures the number of platelets in the blood. High or low platelet levels are indicators of severe conditions. Its normal range is between 150,000 and 450,000 platelets per microliter (mcL) of blood.
Body mass index (BMI)	It is a value in which weight in kilograms (kg) relates to height in meters squared. The normal BMI range is from 18.5 to 25 kg/m ² . Obesity, overweight, and underweight have a BMI of more than 30, 25, to 30, and less than 18.5 kg/m ² .
Pulmonary Embolism Wells Score	It is a score of risk classification and clinical decision rule to predict the acute pulmonary embolism (PE) probable in patients who had the possibility of acute PE in their history and examination. A score is classified into three categories: a score greater than 6, a score of 2 to 6, and a score less than 2 considering as “high risk,” “intermediate risk,” and “low risk” of PE respectively.
Deep vein thrombosis (DVT)	A normal complication in trauma patients. It accounts for three risk classes: high with the range of 3 points or more, intermediate ranging from 1 to 2 points, and low including less than 1 point.
Hypertension (HTN)	A medical condition in which the blood pressure is continuously raised in the arteries. The range of 100–130 millimeters mercury (mmHg) systolic, and 60–80 mmHg diastolic is normal blood pressure at rest for most adults.
Glasgow Coma Scale (GCS)	It is a neurological scaling system describing the brain's consciousness level and its changes to traumatic brain injury patients and further assessment. The criteria of the scale for assessing a person's consciousness are ranging from 3 (a sign of deep unconsciousness) to 15 (normal scale of consciousness)
Intubation	It is a procedure that is used when one cannot breathe on their own
Diabetes mellitus	It is considered a metabolic disease that can result in high blood sugar. Normal blood sugar is 70–99 (mg)/(dl) and high blood sugar is between 80 and 130 (mg) and (dl) when fasting.
Multiple sclerosis (MS)	It is a disease that can potentially disable the brain and central nervous system (called the spinal cord).
Nutrition	It is the overall food that a person or other organism uses to maintain, grow, reproduce, health, and disease of an organism. It is divided into low in calories, NPO, and lose weight
Mobility	Is an average movement of patients during hospitalization. It includes agitation, slightly limited, completely immobile.
Bone fracture	It is a medical condition where the continuity of the bone is broken.
Cerebrovascular accident (CVA)	It is a medical condition in which blood cells' flow in the brain suddenly deteriorates and stops suddenly.

proceed with further diagnostic tests, such as CT scans. Deep Vein Thrombosis (DVT) is a normal complication in trauma patients, although there is a controversy concerning the appropriate use of venous duplex surveillance. It is mainly used for DVT identification. The Wells criterion is a clinically scoring system in an outpatient setting that offers the possible range between 2 and 8. It accounts for three risk classes: high with the range of 3 points or more, intermediate ranging from 1 to 2 point, and low, including less than 1 point. This is also sometimes mentioned a Hamilton score [13–15]. Paraplegia is an impairment of the legs and lower body resulting from injury to nerves in the lumbar or thoracic vertebrae areas of the body. Hemiplegia is an impairment of one vertical half of the body. Stroke is the most common cause of it. Anemia is a condition that the body lacks enough healthy red blood cells to transfer sufficient oxygen into the body's tissues. Sensory processing disorder is also a condition when the brain has trouble in the process of sensory information (includes touch, taste, hear, smell, or see things), both in receiving the information and responding to. It means people who are too sensitive in their typical environment. Such things can be painful or overwhelming. The normal hemoglobin range for men, in general, is defined between 13.2–16.6 grams (g) of hemoglobin per deciliter (dL) of blood and between 11.6 and 15. G/dL for women. Anemia is a situation in which the hemoglobin level is lower than normal [16]. Bilirubin is an orange-yellow substance in the body that is generally formed during the normal breakdown of red blood cells. At first, it goes through the liver and then excretes out of the body. Normal Bilirubin level is less than 1.2 (mg/dL), and the higher level of Bilirubin is an indicator of different types of liver problems. A bilirubin test evaluates the amount of Bilirubin in the blood. It contributes to discovering the root cause of health conditions such as jaundice, anemia, and liver disease [16]. Serum Creatinine is a waste substance in the blood that arises from the activity of the muscle. The kidneys normally remove it from the blood; the creatinine level rises when the kidney function slows down. Its normal level in the blood in adults is from 0.5 to 1.2 milligrams (mg) per deciliter (dL). Severe kidney impairment is a sign when the Creatinine level reaches 2.0 or more in adults. Platelet blood test measures the number of platelets in the blood. The average platelet count blood test is commonly considered a part of a complete blood count (CBC) test. Platelets assist in healing blood wounds and preventing excessive bleeding. High or low platelet levels are indicators of severe conditions. Its normal range is between 150,000 and 450,000 platelets per microliter (mcL) of blood. Having more than 450,000 platelets is a state which is known as thrombocytosis; having less than 150,000 is called thrombocytopenia.

Hypertension This is a medical condition in which the blood pressure is continuously raised in the arteries [17]. It is also

named high blood pressure (HBP) and normally does not cause symptoms [18]. However, in the long-term, it can be a major risk factor for disease including atrial fibrillation, vision loss, stroke, coronary artery, peripheral arterial, chronic kidney, heart failure, and dementia [19–22]. The systolic pressure and diastolic pressure are two measurements expressing the maximum and minimum of Blood pressure [18]. The range of 100–130 millimeters mercury (mmHg) systolic and 60–80 mmHg diastolic are normal blood pressure for most adults [23, 24]. For most adults, the blood pressure at or above 130/80 or 140/90 mmHg is a sign of high blood pressure [23, 25]. The Glasgow Coma Scale (GCS) is a neurological scaling system describing the brain's consciousness level and its changes to traumatic brain injury patients. It is a tool that needs nurses who equipped themselves with the latest information on its function. The criteria of the scale for assessing a person's consciousness are ranging from 3 (a sign of deep unconsciousness) to 15 (normal scale of consciousness) [26]. Diabetes mellitus is considered a metabolic disease that can result in high blood sugar. Sugar is carried by hormone insulin from the blood into the body cells either to be stored or produced. Without diabetes, not only the body can make enough insulin, but also it can use insulin effectively. It is commonly known as diabetes; its normal range is 70–99 (mg)/(dL), and high blood sugar is between 80 and 130 (mg)/(dL) when fasting [27].

The doctor puts a tube down the throat and into the windpipe to make it easier to get air into the lungs. The ventilator pumps the air with extra oxygen, and then, it helps breathe out air full of carbon dioxide (CO₂). Multiple sclerosis (MS) is a disease that can potentially disable the brain and central nervous system (called spinal cord). In MS, the protective sheath (myelin) covers nerve fibers. The protective sheath is in danger of communication problems. Eventually, the disease can cause permanent damage or nerve deterioration [28]. Nutrition or diet is the overall food that a person or other organisms use to maintain, grow, reproduce, and healthy. Diet usually refers to the consumption of exact nutrition for some reasons, including health- or weight-management. It is divided into “low in calories,” “nothing by mouth” is a medical instruction that withholds food and fluids (known as nil per os (npo or NPO)) and loses weight. Mobility is an average movement of patients during hospitalization, which can be included: agitation, slightly limited, completely immobile. A bone fracture is a medical condition where the continuity of the bone is broken. A cerebrovascular accident (CVA) is a medical condition in which the flow of blood cells in the brain impairs and stops suddenly. It is also mentioned as a stroke. Stroke can weak one part of the body with partial or complete loss of voluntary movement or sensation in a leg or an arm such as numbness, speech, vision and balance problems, weak face muscles, and even unconsciousness(see Table 1). The remainder of this

work is structured as follows: related work is represented in Section 2, followed by Section 3, which describes the proposed method, along with the data selection and the strategy of model evaluation. We also discussed the results and the underlying analyses in Section 4, and in Section 5, the conclusions and future works are drawn.

2 Literature review

The ICU commonly admits the most severely ill patients who need a particular cure or continuous monitoring. The most advanced monitoring equipment and device of the hospital are gathered inside the ICU. It plays a vital role in improving treatment and declining mortality. Mortality prediction in ICU can reflect the severity of disease or the prognosis of patients, get the reasonable allocation of clinical resources, and help clinicians make the right decision. Hence, the prediction of mortality for ICU patients is always one of the most remarkable topics in medical and healthcare research, which has attracted many scholars these days [29]. A growing number of them assess patient mortality with their organizations' data source [30–34].

In contrast, public datasets are used by the rest [35–37]. MIMIC-II (The Multi-parameter Intelligent Monitoring in Intensive Care) database can be cited as an excellent example for such studies [38, 39]. It is one of the most available databases, including physiological signals and comprehensive clinical data for a cohort of ICU patients. Researchers carried out a considerable amount of research to predict the patient's mortality in this section. In this section, we will review recent studies and methods in this area proposed by the researchers. Fika et al. (2018) [40] tried to develop and validate a new multivariate logistic regression model in order to predict mortality in an intensive care unit (ICU). Their model improved both the performance and the predictive capability regarding ICU mortality. Shukeri et al. (2017) [30] suggested a logistic regression model to demonstrate the biomarker's potential. They aimed to use logistic regression to obtain a prediction model for sepsis mortality score (SMS). They made use of a multimarker approach including leukocytes count, PCT, IL-6, and PON and ARE activities of PON-1 for 30-day mortality in sepsis and compared its performance to the Sequential Organ Failure Assessment (SOFA) score. The mortality rate that they reached was 28.9%. Liu et al. (2018) [41], proposed ICU mortality prediction model. Their model includes pre-processing, feature extraction, predictive model, and parameter optimization. The stress was on data of high dimensionality and imbalance distribution that in feature extraction stage, a cost-sensitive principal component (CSPCA) strategy was suggested to handle these problems. The best AUC performance (0.77) belongs to support the vector machine (SVM). Davoodi et al. (2018) [35] used a modified deep rule-based fuzzy model

(DRBFS) in forecasting the patient's in-hospital mortality rate by MIMIC-III dataset. A supervised fuzzy clustering technique for training was employed. Their results demonstrated the superiority of DRBFS among various classifiers such as naïve Bayes (NB), decision trees, gradient boosting (GB), deep belief networks (DBN), and D-TSK-FC. Ding et al. (2018) [31] showed a procedure of two-step JITL-ELM to mortality prediction for patients in ICU in which applied just-in-time learning (JITL) and extreme learning machine (ELM). The first step is clustering and the second one is for mortality prediction. Their model achieved 0.8568 AUC, which outperformed the existing traditional model such as ELM, LR, and BP neural network. Darabi et al. (2018) [32] extracted all the necessary information by using a subset of MIMIC III and combined gradient boosted trees and deep neural networks to estimate the mortality risk of patients who are admitted into an ICU. They used the AUC index to validate the proposed algorithm, and it was around 87.30 AUC on the 10% test set. Meadows et al. (2018) [42] have researched patients admitted to the ICU after cardiac surgery over a period of 10-year to predict LOS (length of stay) at the hospital. Logistic regression analysis was carried out to formulate. By using this approach, patients separate into two groups: less than 48 h and more than 48 h named short stay and long stay; the overall accuracy was 79.77%. Kutyrev et al. (2019) [43] used machine learning and neural network methods to predict patients mortality. Models such as CHF, RF, NB, DT classifier, and multilayer perceptron were used. The performance was improved by using machine learning methods (97% ROC curve). Five algorithms such as SVM, elastic-net penalized logistic regression, RF, NN, and stochastic gradient boosting were assessed by Karhade et al. (2019) [44] for prediction of 90-day mortality in a spinal epidural abscess (SEA). The first digital clinical prediction tool was established for mortality spinal epidural abscess. A logistic regression model was presented to predict the risk of ICU discharge based on snapshot measurements by MIMIC-II dataset by Xuea et al. (2019) [19]. They compared various widely used imputation methods and their impact on predictive models. The best AUC of their methodology was 0.661. Lin et al. (2019) [37] compared RF with other machine learning algorithms, including ANN and SVM models, and customized SAPS II model for acute kidney injury (AKI) patients in the ICU to obtain the best model for mortality prediction. They showed the RF model has the lowest Brier score (0.085, 95%, CI: 0.084–0.086) and the largest AUROC (0.866, 95%CI: 0.862–0.870). They achieved an acceptable accuracy and F1 score for RF among the models. Todd et al. (2019) [33] assessed high-frequency data and the performance of models based on model accuracy, calibration, and discrimination to improve ICU scoring systems. Three logistic regressions (baseline baseline + pulse baseline + MAP) were used. As a result, they have proved that the scoring system (APACHE

III) can be raised by better frequency. Deliberato et al. (2019) [34] proposed a mortality prediction model for ICU by taking advantage of logistic regression. The dataset included two group patients: the former, namely the HIAE group, in which patients were elderly and had high BMI and lower rates of smoking and alcoholism compared to the latter group (HMVSC group). The threat of hospital mortality was higher in the latter group (18.7%) compared to the former group (8.4%). The use of the least absolute shrinkage and selection operator (LASSO), random forest (RF), gradient boosting machine (GBM), and the traditional logistic regression (LR) algorithms has been proposed by Kong et al. (2020) [45], to predict the risk of in-hospital death for sepsis patients. Data were collected from MIMIC III as an ICU database including the diagnostic codes, vital signs, laboratory tests, demographics, and some other clinical characteristics of each patient between 18 and 90 years old. Gradient boosting machine (GBM) model showed the best performance, and the AUROCs of GBM model was 0.845. They reached the decedents of patients who were older and they had higher SOFA and SAPS II scores compared to the survivors of patients. The results lead to more prolonged ICU stays and hospital stay reduction. Kang et al. (2020) [46], by using k -nearest neighbor (KNN), support vector machine (SVM), multivariate adaptive regression splines (MARS), random forest (RF), extreme gradient boost (XGB), and artificial neural network (ANN) algorithms (10-fold cross-validation) for patients undergoing CRRT for AKI, have shown the power of machine learning algorithms. The used parameters include age, sex, application of mechanical ventilation, and co-morbidities including diabetes mellitus, hypertension, myocardial infarction, chronic heart failure, stroke, peripheral vascular disease, dementia, chronic obstructive pulmonary disease, connective tissue disease, peptic ulcer disease, cancer, ischemic heart disease, chronic kidney disease, atrial fibrillation, and Vital signs, such as mean arterial pressure, heart rate, respiratory rate, and body temperature were collected. As a result, random forest (RF) by Gini impurity to explore the necessary variables demonstrated better results than other machine learning models and also was the best among APACHE II, SOFA, and MOSAIC. Assaf et al. (2020) [47] aimed to explore the risk scale of factors affecting critical COVID (envisage patients with bacteria *P. aeruginosa* in the ICU) by making use of neural network, random forest, and classification and regression decision tree (CRT) models. They assessed clinical, hematological, and biochemical factors at admission. Different features including medical history, current complaints, measurement of vital signs, baseline testing for blood count, kidney and liver function tests, and inflammatory markers were selected in this work and SVM achieved the best performance and the best accuracy rates of 0.92%. Random forest regression, one of the best machine learning algorithms, is applied by Ripoli et al. (2020) [48], to envisage the candidemia possibility among

patients admitted in IMWs. They collected demographic and clinical data of patients and reached C-statistics = 0.874 ± 0.003 , sensitivity = $84.24\% \pm 0.67\%$, and specificity = $91\% \pm 2.63\%$. In their work, in-hospital MHIA therapy has been proven a frequent risk factor for candidemia as well as other factors such as TPN, previous hospitalization, previous antibiotic therapy, and CVC or PICC.5. Rodríguez et al. (2020) [49] suggested supervised classification methods for mortality prediction among patients with sepsis. In their work, the best discrimination is provided by the supervised classification method via physiological variables. Some techniques were used in their work: decision tree, random forest, artificial neural networks (ANN), and support vector machine (SVM) models. The data was included in patients >18 years. They used two databases: the former includes clinical care variables in which random forest demonstrated the best performance; the latter is direct physiological variables. The best accuracy was achieved by the neural network (ANN) model with an AUC-ROC of 0.69 (95%CI: 0.62; 0.76) and AUC-ROC of 0.69 (95% CI: 0.61; 0.76), respectively. Liao et al. (2020) [50] tried to establish a prediction method for patients with pneumonia. They applied machine learning (K-nearest neighbors, Naive Bayes, decision tree, neural network, support vector matching including the linear kernel, polynomial kernel, and radial basis kernel, and random forest). The data were collected from the breath of the endotracheal tube via ventilator and then analyzed. Two groups were analyzed: first, patients with pneumonia, myocardial infarction, diabetes, aspiration pneumonia, hepatitis, endocarditis, heart failure, lung cancer, chronic obstructive pulmonary disease, hepatocellular carcinoma, idiopathic pulmonary fibrosis, colon cancer, necrotizing fasciitis, kidney injuries, hyponatremia, and cardiac arrest. Second, the intracranial hemorrhage, gastric cancer, traffic accident, fracture, gastric ulcer, coronary artery disease, acute kidney injury, traumatic brain injury, aortic dissection, lung cancer, Fournier's gangrene, liver abscess, and some attributes demonstrated of low importance such as age, gender, smoking status, liver and renal function tests, and the number of comorbidities. In this case, random forest achieved the best AUC (%90). Beteck et al. (2020) [51] established various supervised classification methodologies including SVM, LR, RF, and gradient boosted decision trees (XGBoost) by using Mexican Electronic Health Records information of patients who test for SARS-CoV-2: demographics preconditions and outcomes of admission such as age, gender, diabetes, COPD, asthma, immunosuppression, hypertension, obesity, pregnancy, chronic renal failure, tobacco use, other diseases, and the SARS-CoV-2 test result to an ICU. They envisaged mortality and the importance of an ICU or ventilator. As a result, SVM and LR were better than RF and XGBoost. In their model, age, gender, immunosuppression, chronic renal insufficiency, obesity, and diabetes were significant attributes.

Our work is different from other methods due to some reasons: first, we advance an illustratable method that can envisage the outcomes. Second, it can evaluate the meaningful existence of different variables in making this prediction. Third, our method also clarifies the importance of various data processing scenarios in the pre-processing phase and strong feature selection ability, thereby differing our results from other previous studies mentioned above. Fourth, in our work, underlying diseases, age, gender, laboratory results, and nutrition are considered significant factors, so our results and analysis differ from previous reports. Fifth, although machine learning algorithms appear to be useful and have been used in different medical fields to predict critically ill mortality and outcomes, they do not adequately predict mortality because of low prediction accuracy and limited variables. Besides, this study explored whether machine learning algorithms are also applicable to predicting the mortality of ICU patients. We also compare the performance of several machine learning models and may explain the power of the AdaBoost algorithm, which is not considered in the previously mentioned works. Furthermore, the proposed method acquires high AUC, sensitivity (recall), Kappa, and F-measure for ICU patients. Our experimental results demonstrate that the proposed method is superior to state-of-the-art methodology.

We use the dataset of an ICU ward in a General Hospital of Isfahan in our study. The goal of our study is to predict the mortality rate of patients who are admitted to an intensive care unit. In this work, we applied the essential forecasting methods such as decision tree (DT), neural network (NN), K-nearest network (K-NN), random forest (RF), linear regression (LR), AdaBoost, Bayesian Boosting, Vote (DT+K-NN), and Vote (DT+K-NN+LR).

3 Methods

To evaluate the ability of the proposed model, we use a dataset of 180 patients collected from a general hospital (a ward in an ICUs) between 2017 and 2018. The dataset consists of detailed information about ICU patients' stays, including demographic information and medication variables. The attributes were extracted, and a new dataset was created. We select 28 medical attributes of patients (see Table 2). We only include patients who have been admitted in a ward of ICUs. Each patient must have at least 12 h of data since we used data from the last 12 h before discharge to train our models. However, some attributes were eliminated from data analysis since their patient data records set were not available. As a result, a total of 180 patients opted in the final data set for data analysis to satisfy our criteria.

3.1 Methodology

Around 28 attributes used as input including body mass index (BMI), Pulmonary Embolism Wells score, deep vein thrombosis (DVT) Wells score, paraplegia/hemiplegia, anemia, diabetes mellitus, Bilirubin, Serum Creatinine, platelet blood count, hypertension (HTN or HT), GCS admit, GCS discharge (indicating discharged patients state and is considered true in the case of patients betterment), age, gender, skin type, having operation during hospitalization, length of stay (LOS) at ICU which is determined as the difference between the date of admission and discharge, smoking, intubation, diabetes mellitus, MS, nutrition, mobility, fracture, CVA. The selected attributes and their types are shown in Table 2. Also, Rapid Miner is a statistical tool and software which is used for data analysis, manipulation, calculation, and graphical representation. Figure 1 shows the overall figure which we consider as an ICU mortality prediction model. Health data frequently includes high dimensionality, imbalance, and time asynchronization data. Such complex properties and

Table 2 Selected attributes for modeling

Variable	Type
Operation	Binominal
Smoking	
Hemiplegia/paraplegia	
Anemia	
Diabetes mellitus	
Motor sensory disorder	
Hypertension (HTN)	
Intubation	
Multiple sclerosis (MS)	
Bone fracture	
Cerebrovascular accident (CVA)	
Gender	
Bilirubin (mg/dL)	
Serum Creatinine (mg/dL)	
Platelet blood count (mcL)	Numerical
Body mass index (BMI) (kg/m ²)	
Pulmonary Embolism Wells Score admit	
Pulmonary Embolism Wells Score discharge	
Deep vein thrombosis (DVT) score admit/discharge	
Deep vein thrombosis (DVT) score discharge	
Glasgow Coma Scale (GCS) admit	
Glasgow Coma Scale (GCS) discharge	
LOS	
Age	
Nutrition	Polynomial
Mobility	
Skin Type	

Fig. 1 The proposed model, moving from selecting data to predicting discharge

characteristics of data involve accurate methods for processing. The proposed ICU mortality prediction model is made up of four basic parts. The first step is data preprocessing, which is the most important step. Our study includes replacing missing values by k-NN-imputation and random forest-imputation, feature selection, data sampling, data reduction, and feature transformation stages. The next step is modeling includes decision tree (DT), neural network (NN), K-nearest network (KNN), random forest (RF), linear regression (LR), AdaBoost, Bayesian Boosting, Vote (DT+K-NN), and Vote (DT+K-NN+LR), then model testing that contains confusion matrix, finally, an evaluation of data analysis. Four core components are effectively combined to develop a prediction model of a patient's mortality at the hospital.

3.2 Data preprocessing

Four steps for preprocessing are considered in the phase of data preparation to prepare clean data for modeling.

3.2.1 Replacing the missing value

In the beginning, we imputed the missing value with k-NN-imputation and random forest-imputation. Missing value usually happens for medical data set due to some reasons, and patients with incomplete data may bias our study. The k-nearest neighbor algorithm can be used for imputing missing data by discovering the k-closest neighbors. It not only predicts the most frequent value among the k-nearest neighbor as discrete attributes, but also it can predict the mean among the k-nearest neighbor as continuous attributes [52]. RF also puts a stop to data over-fitting. Moreover, it efficiently runs on many variables (except for dimensionality problems), and it assesses variable importance. It includes methods which are

effective for missing data estimation [53–56] such properties make the k-NN and RF algorithm logical candidates for missing data in our study, and as a result, the k-NN-imputation method proved to be more effective than others.

3.2.2 Sampling of the existing data set

After data imputation, in the next step of pre-processing, sampling was conducted to achieve a more accurate model. Data sampling is a statistical analysis technique. It is mainly used to select, manipulate, and analyze a representative subset of data points to identify patterns and trends in the examined data set.

3.2.3 Feature generation

As the next step of preprocessing, the standardized features after feature generation were obtained. Table 3 provides the following information.

3.2.4 Feature selection

Feature selection (FS) procedure can be employed to maintain the feature subset that contributes to the classification accuracy and facilitates the classification task. Moreover, the removal of irrelevant and redundant features could significantly reduce the risk of complexity. It aims to be discarded in the case of imbalanced classification problems to achieve a subset of relevant features in compact models that are easier to interpret, so improve knowledge extraction [28, 57, 58]. In our work, by using FS, some features including skin type, BMI, MS, and CVA have been removed since they were of low importance.

Table 3 Selected attributes and related ranges in our work

Variable	Range
Bilirubin (mg/dL)	A: Bilirubin ≤ 1.2 B : $1.2 < \text{Bilirubin} \leq 1.9$ C : $1.9 < \text{Bilirubin} \leq 5.9$ D : $5.9 < \text{Bilirubin} \leq 11.9$ E : $11.9 < \text{Bilirubin}$
Serum Creatinine (mg/dL)	A: Cr ≤ 1.2 B : $1.2 < \text{Cr} \leq 1.9$ C : $1.9 < \text{Cr} \leq 3.4$ D : $3.4 < \text{Cr} \leq 4.9$ E : $4.9 < \text{Cr}$
Platelet blood count (mcL)	A: Plt ≤ 20000 B : $20000 < \text{Plt} \leq 50000$ C : $50000 < \text{Plt} \leq 100000$ D : $100000 < \text{Plt} \leq 150000$ E : $150000 < \text{Plt}$
Body mass index (BMI) (kg/m ²)	Normal : $18.5 \leq \text{BM} < 25$ Obesity : $25 \leq \text{BM} < 30$ Overweight : $\text{BM} < 30$ Underweight : $\text{BM} < 18.5$
Pulmonary Embolism Wells Score admit/discharge	High probability: score > 6 Moderate probability: score > 2 and ≤ 6 Low probability : score < 2
Deep vein thrombosis (DVT) score admit/discharge low: 0–2 normal: 2–3 high: 3–8 Glasgow Coma Scale (GCS) admit	Deep coma/Death: GCS < 3 Severe: $3 \leq \text{GCS} \leq 5$ Low : $5 < \text{GCS} \leq 9$ Moderate : $9 < \text{GCS} \leq 12$ Mild: $12 < \text{GCS} \leq 14$ Fully alert : $14 < \text{GCS}$
Glasgow Coma Scale (GCS) discharge	False: $15 \leq \text{GCS} \leq 9$ True : $9 < \text{GCS} < 2$
Mobility	Agitation, slightly limited, completely immobile
Skin type	Dry, normal, moist

3.3 Model evaluation

Eight different machine learning techniques based on data type are selected in this study. Decision tree (DT), K-NN, logistic regression (LR), random forest (RF), and ensemble methods such as Bayesian Boosting, Vote (DT, K-NN), Vote (DT, K-NN, and RF), and AdaBoost are mainly used among various machine learning techniques.

- Decision tree (DT) classifies instances or examples into branchlike segments by taking paths from the root node through internal nodes to leaf nodes. In this method, the root node implies the predictor that gives the best split of the target class values. In this model, a set of training examples is broken down into smaller and smaller subsets. In contrast, an associated decision tree gets incrementally developed (in our case, the maximum depth of a tree is considered 20). At the end of the learning process, a decision tree that covers the training set is returned [59].

- K-NN is a type of instance-based learning in situations that the function is only approximated locally and all computation is deferred until the classification phase. In this method, data is divided into training and test; then, K is selected (we have selected training set and test data 80% and 20% of total dataset respectively, and $K=5$). Determine which distance function is to be used. After that, a sample from the test data is selected to classify and it is computed the distance to its n training samples. Then, the obtained distances and the K-nearest data samples are sorted. In the end, based on the majority vote of its k neighbors, the test class is assigned to the class [60].
- Logistic regression (LR) is used exclusively for two-class problems, so we convert our data into binary variables. Input values are combined by using weights or coefficient values to predict an output value. The first predicted output is mainly considered the bias or intercepts term; the next one is the coefficient for the single input value. Each column in input data has a constant real value that must be learned from training data. In all sections, we considered 80% and 20% of the total dataset for training and testing, respectively. This model uses maximum likelihood estimation to evaluate the probability of class membership [61].
- Artificial neural network (ANN) is mainly based on the biological and processing nervous system, which includes a multilayer perceptron and a single hidden layer. It would be an appropriate choice for survival prediction. At first, some weights are assigned to input, and then their predicted output values are calculated (we selected the learning rate=0.3). After finding the value of the predicted output, it is compared with the target output. The aim is to reach a low error level and high accuracy for the model [61].
- Random forest (RF) is considered an ensemble classifier that contains a multitude of decision trees. This method is taking advantage of the principle of “voting” to reach an accurate output. Multiple trees (in our study trees = 10) grows to the largest extent possible, and there is no pruning; then, in order to classify a new object and predict new data according to features, a classification is taken from each tree to vote for that class. As a result, the class with the most votes is selected [61, 62]. Voting is the simplest form of combining individual classification algorithms in which different multiple models and rules are selected. Then, they combined to build a strong classifier. We have selected the combination of DT and K-NN as a model, and DT, K-NN, and RF as another one.
- AdaBoost is one of the most well-known algorithms of boosting, which combines the result of a series of the classifier algorithms. These algorithms are often decision trees in a single robust classifier; in each model, the data set is trained and assigned an amount of weight to achieve an acceptable accuracy score for the overall classifier. The

AdaBoost algorithm is one of the best boosting techniques. It is a linear model that is mainly used as an ensemble method in ML. It is used to decrease both bias and variance in supervised learning. Firstly, it creates n number of decision trees during the training period of data, when the first tree is constructed. In this model, the errors of the first model are recorded by the algorithm. Then, the nodes which are incorrectly classified are considered the input for the next model. Each model aims to solve the errors of prior model. This process is repeated (models 1, 2, N) to meet the specified condition (see Fig. 2). The algorithm creates proper trees that include a start node with several leaves nodes, in our case, 4 number of trees [62, 63].

- Boosting is an ensemble method based on Bayes' theorem which improves the predictive performance of base learning algorithms. Various classifiers are applied and trained. After many iterations, they are combined by boosting algorithm to obtain a strict rule. It implements a meta-algorithm that can be used in conjunction with many other learning algorithms to improve their performance. In our study, we have considered 1 as a subset for training [54, 64, 65]. The advantages and disadvantages of these methods are compared in Table 4.

3.4 Performance: evaluation of metrics

To define how a classifier is practical, various metrics can be evaluated. Although accuracy has been proven to be effective overall, it does not work for some classes. For instance, when there is imbalanced data with small amount of death patients and the majority of survivals like our dataset, so it can have a detrimental effect on the accuracy of the model. As a result, we used the area under receiver operating characteristics (AUC) curve as the most widely-used metric for imbalanced data [66]. Moreover, it refers to the area under the ROC curve. Furthermore, there are lots of other indices: Sensitivity in statistical analysis measures the proportion of true positives that the classifier detected correctly. It is also called recall, or true positive rate, (see Eq. (1.1)); it means correctly detected patients who have at least a condition.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1.1)$$

The Kappa statistic is an agreement indication between the predicted outcomes based on Eq. (1.2). P_o and P_e are

the observed agreement rate between the predictions and the actual values. These show percentage of chance agreement between the predictions and actual values, respectively.

$$\text{Kappa} = \frac{PO - Pe}{1 - Pe} \quad (1.2)$$

F-measure, or F score according to Eq. (1.3), is a metric which is used in the accuracy of a test. The precision and the recall of the test are mainly taken into consideration to calculate F score. F-measure is the harmonic average of precision and recall. In F-measure, 1 is the best score (perfect precision and recall) and 0 is the worst [66–69].

$$\text{F-measure} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recal}}} \quad (1.3)$$

In this case, such measures have provided a proper model evaluation. They also are utilized for model comparison. Comparison of the performance of various machine learning methods in our model is demonstrated in Table 5. In the next section, we analyze the findings. These measures for the AdaBoost model are AUC= 0.966, sensitivity (recall) = 87.88%, Kappa=0.859, F-measure = 89.23%, so AdaBoost accounts for the highest rate. We used 5-fold cross-validation ($K=5$) to determine the model performance, which means that the data is split into 5 folds. In one round of iteration, 1 fold is used to test, and the rest of the folds are served to its train. After five rounds in which a fold has been used as the testing set, the validation results are combined. In the end, the accuracy of the model is acquired. The confusion matrix is mainly used to show the performance of a model on two various categories (see Table 6). Classification problems are considered by using two classes. Generally, each instance is mapped into an element of the set p, n , which are labeled as positive and negative class. A classifier is a mapping from instances to predicted classes. The labels Y, N are used for distinguishing the actual class from the predicted class. There are four possible outcomes as follow: the instance is positive and it is classified as positive, the instance is positive and it is classified as negative, the instance is negative and it is classified as negative, and the instance is negative and it is classified as positive which are counted as true positive (TP), false negative (FN), true negative (TN), false positive (FP), respectively. The results of the confusion matrix eight models are presented in Table 7. Figure 3 is a bar chart of different factors affecting the rate of mortality in the present work. Overall, it is evident from the information provided that the highest amount of mortality allocated to the age group of more than 59 years old; in contrast, mortality risk for people between 2 and 59 years old is low (see figure (a)). Figure (b) shows that the mortality among men is less than women. By looking at figure (c), we can find the

Table 4 The advantages and disadvantages of our model with other models used in this study

Model	Advantages	Disadvantages
Bayesian Boosting	Is moderately robust to outliers Can be used on a small dataset Can learn nonlinear relationships	Vulnerable to overfitting Vulnerable to uniform noise
AdaBoost	Act remarkably in reality Is a type of reinforcement learning algorithm that can realize high precision classification by training several weak classifiers and assembling them into one strong classifier Can learn nonlinear relationships It is scalable Is moderately robust to outliers Its algorithm is fast, simple to implement, and easy to program Implicit feature selection Can remove overfitting (individual trees are inclined to overfitting since they can support branching till they memorize the training data)	Must adjust for cost-sensitive or imbalanced class problems Sensitive to noisy data & outliers Needs a termination condition Vulnerable to uniform noise
Vote (DT+ K-NN)	<ul style="list-style-type: none"> • Can learn nonlinear relationships • It is scalable • Can be used on a small dataset 	<ul style="list-style-type: none"> • Vulnerable to overfitting
K-nearest neighbors	<ul style="list-style-type: none"> • Not assume underlying data distribution • Is easy to implement 	<ul style="list-style-type: none"> • Unable to construct a model • Unable to explore the relationship between a feature and the class • Computationally intensive recall • Require huge storage • With the existence of one class sample in excessive compared to other class, the control class will dominance the classification leading to an incorrect result
Decision tree	<ul style="list-style-type: none"> • Can be used on a small dataset • Can learn nonlinear relationships • Simple to analyze 	<ul style="list-style-type: none"> • Over fitting or under fitting of the model is effortless • Negligible changes in the training data can generate huge changes in the result
Neural network	<ul style="list-style-type: none"> • Able to model complicated samples • Able to detect all possible interactions between predictor variables • Has the availability of multiple training algorithms 	<ul style="list-style-type: none"> • Has a “black box” nature, • Has a great computational burden • Proneness to the overfitting • Inclined to an overfitting training dataset
Random forest	<ul style="list-style-type: none"> • Able to estimate noisy or missing data • Able to keep accuracy when a huge portion of the data is missing • Able to handle thousands of input variables • Can be Suitable for class imbalance issues 	<ul style="list-style-type: none"> • Is not simply understandable model • It is like a black box approach for statistical modelers, the user has a low control level on what the model does • A large number of trees can make the algorithm too slow and ineffective for real-time predictions
Logistic regression(-LR)	<ul style="list-style-type: none"> • Is easy to recognize and describe • Can be improved with no difficulty with new data • Does not require high computation power 	<ul style="list-style-type: none"> • Acts weak when there are nonlinear relationships • Vulnerable to overfitting • Inflexible to absorb intricate patterns • Adding the right interaction terms or polynomials is risky and time-consuming.
Vote (DT+ K-NN+LR)	<ul style="list-style-type: none"> • Is moderately robust to outliers • It is scalable • Can be used on a small dataset • Can learn nonlinear relationships 	<ul style="list-style-type: none"> • Vulnerable to overfitting • Negligible changes in the training data can generate huge changes in the result • Require huge storage

mortality for completely immobile is more than other mobility types (agitation and slightly limited). In figure (d), we can see the mortality for people with moderate PE wells score discharge is far more than mortality for people with low PE wells score. Based on (e), individuals with diabetes mellitus have a high level of mortality compared to people without diabetes mellitus. As we can see in figure (f), the length of stay more than 7 days can increase the mortality rate. Figure (g) shows that people who do not smoke have a higher mortality rate than others. The

mortality of people with moderate PE wells score admit is far more likely than mortality for people with low PE wells score (h). Hemiplegia/paraplegia has an insignificant effect on mortality rate (figure (i)). Patients with GCS less than 9 are more in danger of death (j). Figure (K) illustrates Bilirubin less than 1.2 (mg/dl) leads to a higher mortality rate. Figure (l) depicts that almost a great number of patients have a normal rate of platelet blood count. As a result, it cannot play a key role in the mortality rate. Figure (m) displays a higher mortality rate

Table 5 Comparison the performance of various machine learning methods

Model	AUC	Sensitivity (recall) (%)	Kappa	F-measure
Bayesian Boosting	0.895	86.11%	0.847	88.57%
AdaBoost	0.966	87.88%	0.859	89.23%
Vote(DT+K-NN)	0.898	55.26%	0.606	68.85%
K-nearest neighbors	0.881	40.62%	0.440	53.06%
Decision tree	0.836	72.41%	0.764	82.35%
Neural network	0.888	47.62%	0.534	58.82%
Random forest	0.916	73.33%	0.614	68.75%
Logistic regression (LR)	0.923	81.25%	0.802	83.87%
Vote (DT+K-NN+LR)	0.875	70.59%	0.815	81.36%

that can be as a result of Creatinine more than 1.9. Patients with hypertension can be a key factor of high mortality rate (figure (n)). Figure (o) manifests the high risk of mortality that arise from moderate and high Pulmonary Embolism Wells Score admit. Figure (p) represents that high and moderate DVT wells score discharge puts individuals at risk of death.

4 Discussion

The data analysis of the present study is discussed as follow:

4.1 Mortality and underlying diseases

High mortality can be as a result of underlying diseases (such as diabetes mellitus which is a metabolic disease that can result in high blood sugar and high blood pressure). To cite an example, patients with the Glasgow Coma Scale (GCS) less than 9 indicate the low level of brain consciousness and laboratory disorders, including the low level of platelets can give rise to discharge. In contrast, with such conditions, the existence of underlying diseases can lead to death. There is no denying fact that the more “completely immobile” patients are, the more rate of mortality they have, even the patients are “fully alert” in admission. Underlying diseases such as Hypertension (HTN) as a medical condition in which the blood pressure is continuously raised in the arteries can be a cause of high mortality. It really depends on three factors: the age of patients, laboratory disorder, and smoking. For example, it can lead to a high mortality rate among patients who are

more than 76 years old and have a moderate Pulmonary Embolism Wells Score and low level of brain consciousness (GCS=6–9). Hypertension (HTN) among patients with a low level of brain consciousness (GCS= 9–6) and completely immobile and nutrition = NPO can cause death in comparison with patients without hypertension. As a result, hypertension can increase the mortality rate. Diabetes mellitus can be an influencing attribute in boosting mortality, especially among female patients.

4.2 Mortality and the Glasgow Coma Scale

The low level of brain consciousness (GCS= 9–6) plays a remarkable role in boosting discharge rate, regardless of underlying diseases. Underlying diseases include high blood pressure and impairment a vertical half of the body/impairment of the legs and lower body. These are as a result of the injury to nerves in the lumbar or thoracic vertebrae areas

Table 6 The performance of confusion matrix and related metrics.

Hypothesized class/true class	P	N
Y	True positive (TP)	False positive (FP)
N	False negative (FN)	True negative (TN)

Table 7 The results of the confusion matrix eight models

Model/confusion matrix	P	N	Model/confusion matrix	P	N
Bayesian Boosting			Random forest		
Y	98	5	Y	112	19
N	3	31	N	2	4
AdaBoost			Neural network		
Y	101	5	Y	63	4
N	3	31	N	6	11
Vote (DT, K-NN)			Regression		
Y	97	17	Y	65	3
N	2	21	N	2	13
K-NN			Vote (DT, K-NN, RF)		
Y	101	4	Y	102	10
N	3	13	N	1	24
Decision tree					
Y	107	8			
N	1	21

Fig. 2 Our proposed model evaluation in this study

of the body. Mortality rate also can arise from the result of the moderate Pulmonary Embolism Wells Score risk (a score of risk classification and clinical decision rule to predict the critical pulmonary embolism (PE) probable in patients who had the possibility of acute PE in their history and examination, even though patients with low level of Glasgow Coma Scale (GCS), laboratory disorder, and low Pulmonary Embolism Wells Score risk can increase pre-discharge and reduce the mortality rate (discharge is by no means fully alert)). The

Fig. 3 The bar plotter indicates different factors affecting the rate of mortality. **a** The effect of age on mortality. **b** The effect of gender on mortality. **c** The effect of mobility on mortality. **d** The effect of PE wells scores discharge on mortality. **e** The effect of diabetes mellitus on mortality. **f** The effect of LOS on mortality rate. **g** The effect of smoking on mortality. **h** The effect of PE wells scores on mortality. **i** The effect of hemiplegia/paraplegia on mortality. **j** The effect of GCS admits on mortality. **k** The effect of Bilirubin on mortality. **l** The effect of DVT wells scores discharge on mortality. **m** The effect of Creatinine on mortality. **n** The effect of hypertension on mortality. **o** The effect of Pulmonary Embolism Wells Score admits on mortality. **p** The effect of platelet blood count on mortality

higher level of “GCS admission” can lead to increasing the probability of discharge. It needs to be highlighted that the “GCS admission” level is a contributing factor in discharge.

4.3 Mortality and gender

Mortality rate increased among females who were 26–38 years old and had a low level of brain consciousness (GCS= 9–6) and moderate Pulmonary Embolism Wells Score risk with length of stay (LOS) = more than 16 days. Patients who were less than 26 years old and had a low level of brain consciousness (GCS=6–9), platelet blood count less than 100000 (mcl), and moderate Pulmonary Embolism Wells Score are more in danger of death than others. Gender is of low importance among patients with a low level of brain consciousness (GCS less than 5) and normal complication in trauma patients called moderate deep vein thrombosis (DVT) wells score; nonetheless, there is a risk of DVT wells score. Female patients who are 76 years old and have a normal rate of Serum Creatinine and are completely immobile have a higher mortality rate, whereas males with these conditions have a higher discharge rate.

4.4 Mortality and laboratory findings

Platelet blood count of less than 100000 (mcl) can cause death. Moreover, death can arise from laboratory findings, along with platelet blood count less than 70000 (mcl). The underlying disease solely cannot cause death unless it is accompanied by the blood count of less than 100000. Low level of platelet blood count among patients less than 29 years old leads to death and this factor has no effect on the death of patients more than 29 years old.

4.5 Mortality and serum creatinine

Mortality is likely to be affected by Serum Creatinine (a waste substance in the blood that arises from muscle activity) more than 3.4(mg/dl), indicating if the kidney’s function slows down, mortality rate increases.

4.6 Mortality and age

Patients who were 76 years of age and above and have creatinine 1.2–1.9 were fatal just if they had platelets less than 100000 (mcl). Otherwise, they have been discharged even with aging. In patients older than 76 years old, mortality is increased just if there is a laboratory disorder Serum Creatinine or being a smoker. In contrast, discharge is increased with the same conditions, just if the patient is a non-smoker.

4.7 Mortality and length of stay

Length of stay cannot be the sole reason for the high mortality rate. It should be along with laboratory disorder or smoking to increase mortality risk with the exception of the patients less than 76-year-old patients. Mortality also can arise from the high level of Bilirubin, low level of brain consciousness (GCS= 6–9), and length of stay for more than 12 days. In contrast, the normal level of Bilirubin along with the length of stay = 6–12 days accounts for discharge.

4.8 Mortality and smoking

Patients who are less than 72 years old with the low level of Glasgow Coma Scale (GCS= 6–9), Bilirubin levels less than 1.2 (mg/dL), low level of Serum Creatinine, and deep vein thrombosis (DVT) risk have length of stay 4–6 days. In contrast, these conditions for smoker patients who are less than 26 years old have increased length of stay up to 6–13 days. It means the smoking factor has increased the length of stay from 4–6 to 6–13 days for patients who are less than 26 years old.

4.9 Mortality and nutrition

Conditions, such as “GCS admission” between 6 and 9, slightly limited mobility and nutrition = NPO, leading to death after 6–11 days. In contrast, the level of Glasgow Coma Scale of patients who have nutrition = low in calorie, increased to 13–14, which means a normal scale of consciousness, so nutrition can be considered a significant factor in discharging alert patients.

5 Conclusion

Reaching a precise prediction of mortality risk for patients is of great importance in health care and hospital resource allocation, leading to an improvement in patients' durability. The model in our study included various scenarios of data processing. This model demonstrates the pre-processing as a significant step in data mining, and it improves the scale efficiency

of the model. In this article, several machine learning methods were used; the AdaBoost method led to better results in discovering the influential factors in mortality with AUC= 0.966, sensitivity (recall) = 87.88%, Kappa=0.859, F-measure = 89.23%. These results show that the high mortality can be a result of underlying diseases such as diabetes mellitus and high blood pressure, moderate Pulmonary Embolism Wells Score risk, platelet blood count less than 100000 (mcl), hypertension (HTN), high level of Bilirubin, smoking, and GCS level between 6 and 9. GCS level also plays a remarkable role in the rise of patient discharge rate, regardless of underlying diseases like high blood pressure and hemiplegia/paraplegia. Diabetes mellitus can be an influencing attribute in boosting the mortality rate, especially among female patients. A low level of platelet blood count among patients less than 29 years old leads to death, and this factor does not affect the death of patients who are more than 29 years old. Mortality is likely to be affected by Serum Creatinine. In patients older than 76 years old, mortality is increased just if there is a laboratory disorder Serum Creatinine or being a smoker, while discharge is increased with the same condition, just if the patient is a non-smoker. The length of stay cannot be the sole reason for the high mortality rate; it should be along with laboratory disorder or smoking to increase mortality risk except for the patients less than 76-year-old. As a factor, smoking has increased the length of stay from 4–6 to 6–13 days for patients who were less than 26 years old. Nutrition also can be considered a significant factor in discharging with fully alert. Our suggestions for future researches can be listed as follow: considering the patients of alternative hospitals during a longer period of time, utilizing visual information, and determining whether alternative metrics are able to achieve better performance. Since AdaBoost is a reliable predictor of ICU mortality, further research seems possible in developing hybrid models (machine learning and statistical and deep learning models) with the aim of its evaluation, validation, and applicability to other ICUs to improve the performance of the forecast.

References

1. Xu J, Zhang Y, Zhang P, Mahmood A, Li Y, Khatoon S (2017) Data mining on icu mortality prediction using early temporal data: a survey. *Int J Inf Technol Decis Mak* 16:117–159. <https://doi.org/10.1142/S0219622016300020>
2. Moridani MK, Setarehdan SK, Nasrabadi AM, Hajinasrollah E (2015) New algorithm of mortality risk prediction for cardiovascular patients admitted in intensive care unit. *Int J Clin Exp Med* 8(6): 8916–8926
3. Moridani MK, Farhadi H (2017) Heart rate variability as a biomarker for epilepsy seizure prediction. *Bratisl Lek Listy* 118(1):3–8. https://doi.org/10.4149/BLL_2017_001

4. Karimi Moridani M, Setarehdan SK, Motie Nasrabadi A, Hajinasrollah E (2016) Non linear feature extraction from hrv signal for mortality prediction of icu cardiovascular patient. *J Med Eng Technol* 40(3):87–98. <https://doi.org/10.3109/03091902.2016.1139201>
5. Karimi Moridani M, Setarehdan SK, Motie Nasrabadi A, Hajinasrollah E (2013) Mortality risk assessment of icu cardiovascular patients using physiological variables. *Univ J Biomed Eng* 1(1):6–9. <https://doi.org/10.13189/ujbe.2013.010102>
6. Baigelman W, Katz R, Geary G (1983) Patient readmission to critical care units during the same hospitalization at a community teaching hospital. *Intensive Care Med* 9(5):253–256. <https://doi.org/10.1007/bf01691250>
7. Durbin CG, Kopel RF Jr (1993) Patient readmission to critical care units during the same hospitalization at a community teaching hospital. *Intensive Care Med* 21(10):1547–1553. <https://doi.org/10.1097/00003246-199310000-00025>
8. Van Poucke S, Zhang Z, Schmitz M, Vukicevic M, Laenen MV, Celi LA, De Deyne C (2016) Scalable predictive analysis in critically ill patients using a visual open data analysis platform. *PLoS One* 11(1):0e145791. <https://doi.org/10.1371/journal.pone.0145791>
9. Gurm HS, Kooiman J, LaLonde T, Grines C, Share D, Seth M (2014) A random forest based risk model for reliable and accurate prediction of receipt of transfusion in patients undergoing percutaneous coronary intervention. *PLoS One* 9(5):e96385. <https://doi.org/10.1371/journal.pone.0096385>
10. Xie J, Su B, Li C, Lin K, Li H, Hu Y, Kong G (2017) A review of modeling methods for predicting in-hospital mortality of patients in intensive care unit. *J Emerg Crit Care Med* 1(8):1–10. <https://doi.org/10.21037/jecm.2017.08.03>
11. Global BMI Mortality Collaboration, Di Angelantonio E, Bhupathiraju ShN, Wormser D, Gao P, Kaptoge S, Berrington de Gonzalez A, Cairns BJ, Huxley R, Jackson ChL, Joshy G, Lewington S, Manson JE, Murphy N, Patel AV, Samet JM, Woodward M, Zheng W, Zhou M, Bansal N, Barricarte A, Carter B, Cerhan JR, Smith GD, Fang X, Franco OH, Green J, Halsey J, Hildebrand JS, Jung KJ, Korda RJ, McLerran DF, Moore SC, O’Keeffe LM, Paige E, Ramond A, Reeves GK, Rolland B, Sacerdote C, Sattar N, Sofianopoulou E, Stevens J, Thun M, Ueshima H, Yang L, Yun YD, Willeit P, Banks E, Beral V, Chen Zh, Gapstur SM, Gunter MJ, Hartge P, Jee SH, Lam TH, Peto R, Potter JD, Willett WC, Thompson SG, Danesh J, Hu FB. Body-mass index and all-cause mortality.
12. Kline JA (2017) Utility of a clinical prediction rule to exclude pulmonary embolism among lowrisk emergency department patients: Reason to perc up. *JAMA* 319(6):551–553. <https://doi.org/10.1001/jama.2017.21901>
13. Toker S, Hak DJ, Morgan SJ (2011) Deep vein thrombosis prophylaxis in trauma patients. *Thrombosis* 2011:505373. <https://doi.org/10.1155/2011/505373>
14. Modi S, Deisler R, Gozel K, Reicks P, Irwin E, Brunsvold M, Banton K, Beilman GJ (2016) Wells criteria for dvt is a reliable clinical tool to assess the risk of deep venous thrombosis in trauma patients. *World J Emerg Surg* 11:24. <https://doi.org/10.1186/s13017-016-0078-1>
15. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, Forgie M, Kovacs G, Ward J, Kovacs MJ (2001) Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. *Ann Intern Med* 135(2):98–107. <https://doi.org/10.7326/0003-4819-135-2-20010717000010>
16. Voss A, OTR. Your Essential Guide to Understanding Sensory Processing Disorder: PLUS: Travelling with a Sensory Kiddo. Create Space Independent Publishing Platform; 2 edition.
17. Naish J Medical sciences. Elsevier
18. Farley TA, Dalal MA, Mostashari F, Frieden TR (2010) Deaths preventable in the u.s. by improvements in the use of clinical preventive services. *Am J Prev Med* 38(6):600–609. <https://doi.org/10.1016/j.amepre.2010.02.016>
19. Lackland DT, Weber MA (2015) Global burden of cardiovascular disease and stroke: hypertension at the core. *Can J Cardiol* 31(5): 569–571. <https://doi.org/10.1016/j.cjca.2015.01.009>
20. Mendis S, Puska P, Norrving B Global atlas on cardiovascular disease prevention and control. World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization, Geneva
21. Hermendorena I, Duron E, Vidal JS, Hanon O (2017) Treatment options and considerations for hypertensive patients to prevent dementia. *Expert Opin Pharmacother* 18(10):989–1000. <https://doi.org/10.1080/14656566.2017.1333599>
22. Lau DH, Nattel S, Kalman JM, Sanders P (2017) Modifiable risk factors and atrial fibrillation. *Circulation* 136(6):583–596. <https://doi.org/10.1161/CIRCULATIONAHA.116.023163>
23. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, DePalma SM, Gidding S, Jamerson KA, Jones DW, MacLaughlin EJ, Muntner P, Oviagele B, Smith SC, Spencer CC, Stafford RS, Taler SJ, Thomas RJ, Williams KA, Williamson JD (2018) 2017acc/aha/aapa/abc/acpm/ags/apha/ash/aspc/nma/pcna guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: A report of the american college of cardiology/american heart association task force on clinical practice guidelines. *Hypertension* 71(6):e13–e115. <https://doi.org/10.1161/HYP.0000000000000065>
24. Mancia G, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M (2013) 2013 esh/esc guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the european society of hypertension (esh) and of the european society of cardiology (esc). *Eur Heart J* 34(28):2159–2219. <https://doi.org/10.1093/eurheartj/ehf151>
25. Poulter NR, Prabhakaran D, Caulfield M (2015) Hypertension. *Lancet* 386(9995):801–812. [https://doi.org/10.1016/s0140-6736\(14\)61468-9](https://doi.org/10.1016/s0140-6736(14)61468-9)
26. Teasdale G, Murray G, Parker L, Jennett B (1972) Adding up the glasgow coma score. *Acta Neurochir Suppl (Wien)* 28(1):13–16. https://doi.org/10.1007/978-3-7091-4088-8_2
27. Kitabchi AE, Umpierrez GE, Miles JM, Fisher JN (2009) Hyperglycemic crises in adult patients with diabetes. *Diabetes Care* 32(7):1335–1343. <https://doi.org/10.2337/dc09-9032>
28. Cancela B, Bolón-Canedo V, Alonso-Betanzos A (2019) A scalable saliency-based feature selection method with instancelevel information. *Knowl-Based Syst*:105326. <https://doi.org/10.1016/j.knosys.2019.105326>
29. Allaudeen N, Schnipper JL, Orav EJ, Wachter RM, Vidyarthi AR (2011) Inability of providers to predict unplanned readmissions. *J Gen Intern Med* 26(7):771–776. <https://doi.org/10.1007/s11606-011-1663-3>
30. Shukeri WF, Ralib A, Abdulah NZ, Mat-Nor MB (2018) Sepsis mortality score for the prediction of mortality in septic patients. *J Crit Care* 43:163–168. <https://doi.org/10.1016/j.jccr.2017.09.009>
31. Ding Y, Wang Y, Zhou D (2018) Mortality prediction for icu patients combining just-in-time learning and extreme learning machine. *Neurocomputing* 281:12–19. <https://doi.org/10.1016/j.neucom.2017.10.044>
32. Darabia HR, Tsinisa D, Zecchinia K (2018) Forecasting mortality risk for patients admitted to intensive care units using machine learning. *Procedia Comput Sci* 140:306–313. <https://doi.org/10.1016/j.procs.2018.10.313>
33. Todd J, Gepp A, Richard B, Vanstone BJ (2019) Improving mortality models in the icu with high-frequency data. *Int J Med Inform* 129:318–323. <https://doi.org/10.1016/j.ijmedinf.2019.07.002>

34. Deliberato R, Escudero G, Bulgarellia L, Ary Serpa Netob AS, Koe S, Camposb NS, Saatf B, Júnior EA, Lopesg FS, Johnsonc AE (2019) Severitas: an externally validated mortality prediction for critically ill patients in low and middle-income countries. *Int J Med Inform* 131:103959. <https://doi.org/10.1016/j.ijmedinf.2019.103959>
35. Davoodi R, Hassan Moradi M (2018) Mortality prediction in intensive care units (icub) using a deep rule-based fuzzy classifier. *J Biomed Inform* 79:48–59. <https://doi.org/10.1016/j.jbi.2018.02.008>
36. Xuea Y, Klabjanb D, Luo Y (2019) Predicting icu readmission using grouped physiological and medication trends. *Artif Intell Med* 95:27–37. <https://doi.org/10.1016/j.artmed.2018.08.004>
37. Lin K, Hu Y, Kong G (2019) Predicting in-hospital mortality of patients with acute kidney injury in the icu using random forest model. *Int J Med Inform* 125:55–61. <https://doi.org/10.1016/j.ijmedinf.2019.02.002>
38. Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LW, Moody G et al (2011) Multiparameter intelligent monitoring in intensive care ii: a public-access intensive care unit database. *Crit Care Med* 39(5):952–960. <https://doi.org/10.1097/CCM.0b013e31820a92c6>
39. Johnson AEW, Pollard TJ, Shen L, Lehman LWH, Feng ML, Ghassemi M et al (2016) Mimic-iii, a freely accessible critical care database. *Sci Data* 3:160035. <https://doi.org/10.1038/sdata.2016.35>
40. Fika S, Nanas S, Baltopoulos G, Charitidou E, Myrianthefts P (2018) A novel mortality prediction model for the current population in an adult intensive care unit. *Heart Lung* 47(1):10–15. <https://doi.org/10.1016/j.hrtlng.2017.10.009>
41. Liua J, Xiang Chena X, Fanga L, Xia Lic J, Yangd T, Zhand Q, Tongb K, Fang Z (2018) Mortality prediction based on imbalanced high-dimensional icu big data. *Comput Ind* 98:218–225. <https://doi.org/10.1016/j.compind.2018.01.017>
42. Meadows K, Gibbens R, Caroline Gerrard C, Vuylsteke A (2018) Prediction of patient length of stay on the intensive care unit following cardiac surgery: a logistic regression analysis based on the cardiac operative mortality risk calculator, euroscore. *J Cardiothorac Vasc Anesth* 32(6):2676–2682. <https://doi.org/10.1053/j.jvca.2018.03.007>
43. Kuttyrev K, Aleksey Yakovlev A, Metsker O (2018) Mortality prediction based on echocardiographic data and machine learning: Chf, chd, aneurism, acs cases. *Procedia Comput Sci* 156:97–104. <https://doi.org/10.1016/j.procs.2019.08.184>
44. Karhade AV, Shah A, Bono C, Ferrone ML, Sandra B, Nelson MD, Schoenfeld A, Harris M, Schwab J (2019) Development of machine learning algorithms for prediction of mortality in spinal epidural abscess. *Spine J* 19(12):1950–1959. <https://doi.org/10.1016/j.spinee.2019.06.024>
45. Kong G, Lin K, Hu Y (2020) Using machine learning methods to predict in-hospital mortality of sepsis patients in the icu. *BMC Med Inform Decis Mak* 251. <https://doi.org/10.1186/s12911-020-01271-2>
46. Kang MW, Kim J, Kim DK (2020) Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy. *Crit Care* 42:24. <https://doi.org/10.1186/s13054-020-2752-7>
47. Assaf D, Gutman Y, Neuman Y (2020) Utilization of machine-learning models to accurately predict the risk for critical covid-19. *Intern Emerg Med*. <https://doi.org/10.1007/s11739-020-02475-0>
48. Ripoli A, Sozio E, Sbrana F (2020) Using machine learning methods to predict in-hospital mortality of sepsis patients in the icu. *BMC Med Inform Decis Mak* 48:749–759. <https://doi.org/10.1007/s15010-020-01488-3>
49. Rodríguez A, Mendoza D, Ascuntar J (2020) Personalized machine learning approach to predict candidemia in medical wards. *Am J Emerg Med* 251. <https://doi.org/10.1016/j.ajem.2020.09.013>
50. Liao Y, Shih C, Abbod MF (2020) Development of an e-nose system using machine learning methods to predict ventilator-associated pneumonia. *Microsyst Technol*. <https://doi.org/10.1007/s00542-020-04782-0>
51. Wollenstein-Betech S, Cassandras CG, Ioannis C, Paschalidis I (2020) Personalized predictive models for symptomatic covid-19 patients using basic preconditions: hospitalizations, mortality, and the need for an icu or Ventilator. *Int J Med Inform*. <https://doi.org/10.1016/j.ijmedinf.2020.104258>
52. Jonsson P, Wohlin C (2004) An evaluation of k-nearest neighbour imputation using likert data. 10th International Symposium on Software Metrics. IEEE. <https://doi.org/10.1109/METRIC.2004.1357895>
53. Breiman L, Cutler A (1998) “Random forests” The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell* 20(8):832–844. <https://doi.org/10.1109/34.709601>
54. Biau G, Devroye L, Lugosi G (2008) Consistency of random forests and other averaging classifiers. *J Mach Learn Res* 9:2015–2033. <http://www.jmlr.org/papers/volume9/biau08a/biau08a.pdf>
55. Pantanowitz A, Marwala T (2009) Evaluating the impact of missing data imputation. *Int Conf Adv Data Min Appl* 5678:577–586. https://doi.org/10.1007/978-3-642-03348-3_59
56. Beretta L, Santaniello A (2016) Nearest neighbor imputation algorithms: a critical evaluation. *BMC Med Inform Decis Mak* 16:–74. <https://doi.org/10.1186/s12911-016-0318-z>
57. Yijing L, Haixiang G, Xiao L, Yanan L, Jinling L (2016) Adapted ensemble classification algorithm based on multiple classifier system and feature selection for classifying multi-class imbalanced data. *Knowl-Based Syst* 94:88–104. <https://doi.org/10.1016/j.knosys.2015.11.013>
58. Mirzaei A, Pourahmadi V, Soltani M, Sheikhzadeh H Deep feature selection using a teacher-student network. *Neurocomputing*, 2019. <https://doi.org/10.1016/j.neucom.2019.12.017>
59. Yu Z, Haghighat F, Fung B, Yoshino H (2010) A decision tree method for building energy demand modeling. *Energ Build* 42: 1637–1646. <https://doi.org/10.1016/j.enbuild.2010.04.006>
60. Loh W, Shih Y (1997) Split selection methods for classification trees. *Institute of Statistical Science. Academia Sinica* 7(4):815–840. <https://www.jstor.org/stable/24306157>
61. Deepnarain N, Nasr M, Kumari S, Stenström TA, Reddy P, Pillay K, Bux F (2019) Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. *Process Saf Environ Prot* 126:25–34. <https://doi.org/10.1016/j.psep.2019.02.023>
62. Lantz B Machine Learning with R: expert techniques for predictive modeling, 3rd edn. Packet Publishing Ltd

63. Gunturi S, Sarkar D (2020) Ensemble machine learning models for the detection of energy theft. *Electr Power Syst Res* <https://doi.org/10.1016/j.epsr.2020.106904>
64. Yin C, Wan M, Zhang Y, Fang Z (2020) Recognition of diseased pinus trees in uav images using deep learning and adaboost classifier. *Biosyst Eng* 194:138–151. <https://doi.org/10.1016/j.biosystemseng.2020.03.021>
65. Weng W-H (2020) Machine learning for clinical predictive analytics. Springer International Publishing, Cham, pp 199–217. https://doi.org/10.1007/978-3-030-47994-7_12
66. Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang G (2017) Deep learning for health informatics. *IEEE J Biomed Health Inf* 21(1):4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
67. Majnik M, Bosnic Z (2013) Roc analysis of classifiers in machine learning: a survey. *Intell Data Anal* 17(3):531–558. <https://doi.org/10.3233/IDA-130592>
68. Elsevier Machine learning approach for breast cancer prognosis prediction. Wiley, New York, Ch. Chapter 2, pp 41–68
69. Höftberger R, Lassmann H (2017) Multiple sclerosis and other inflammatory demyelinating diseases of the central nervous system. *Handb Clin Neurol* 145:263–283. <https://doi.org/10.1016/B978-0-12-802395-2.00019-5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.