International Journal of Cardiology 288 (2019) 140–147

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

Predictors of in-hospital length of stay among cardiac patients: A machine learning approach

Tahani A. Daghistani [a], Radwa Elshawi [d], Sherif Sakr [b,d,*], Amjad M. Ahmed [c], Abdullah Al-Thwayee [c], Mouaz H. Al-Mallah [a,b,c]

[a] King Abdullah International Medical Research Center, Riyadh, Saudi Arabia
[b] King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
[c] King Abdulaziz Cardiac Center, King Abdulaziz Medical city for National Guard, Riyadh, Saudi Arabia
[d] University of Tartu, Tartu, Estonia

ARTICLE INFO

Article history:
Received 24 July 2018
Received in revised form 20 December 2018
Accepted 14 January 2019
Available online 19 January 2019

Keywords:
Cardiac science
In-hospital length of stay
Prediction model
Machine learning
Classification techniques

ABSTRACT

Objective: The In-hospital length of stay (LOS) is expected to increase as cardiovascular diseases complexity increases and the population ages. This will affect healthcare systems especially with the current situation of decreased bed capacity and increasing costs. Therefore, accurately predicting LOS would have a positive impact on healthcare metrics. The aim of this study is to develop a machine learning-based model approach for predicting in-hospital LOS for cardiac patients.

Design: Using electronic medical records, we retrospectively extracted all records of patients' visits that were admitted under adult cardiology service. Admission diagnosis and primary treating physician were reviewed to verify selection criteria. A predictive machine learning-based model approach was applied to incorporate simple baseline health data at admission time to predict LOS. Patients were divided into three groups based on their LOS: short (<3 days), intermediate (3–5 days) and long (>5 days). Information gain algorithm was utilized to select the most relevant attributes. Only attributes with information gain of more than zero were used in model building. Four different machine learning techniques were evaluated and their diagnostic accuracy measures were compared.

Setting: The dataset of this study included adult patients who were admitted between 2008 and 2016 in King Abdulaziz Cardiac Center (KACC). The center is located in King Abdulaziz Medical City Complex in Riyadh, the capital of Saudi Arabia.

Participants (dataset): A total of 16,414 consecutive inpatient visits for 12,769 unique patients (mean age of 58.8 ± 16 years of which 68.2% were males) between 2008 and 2016 were included. The study cohort had a high prevalence of cardiovascular risk factors (hypertension 56%, diabetes 56%, dyslipidemia 52%, obesity 33% and smoking 24%). The most common admitting diagnosis was acute coronary syndrome (36%).

Results: The variables with highest impact on the prediction of in-hospital LOS were on admission heart rate, on admission systolic and diastolic blood pressure, age and insurance status (eligibility). Using machine learning models; Random Forest (RF) model outperformed among all other models (sensitivity (0.80), accuracy (0.80), and AUROC (0.94)).

Conclusion: We showed that machine learning methods provide accurate prediction of LOS for cardiac patients. This is can be used in clinical bed management and resources allocation.

1. Introduction

Cardiovascular diseases (CVD) are one of the leading causes of mortality and morbidty worldwide with >17.7 million deaths recorded every year [1]. Care provided to cardiac patients is labour intensive and costly, and the financial burden caused by cardiovascular diseases is expected to exceed 20 billion US dollar by 2030 in some countries [2]. The length of stay (LOS) is an important metric in healthcare access which is expected to increase in the coming years as the population ages

Abbreviations: LOS, length of stay; ML, machine learning; CVD, cardiovascular disease; NMHRS, National Medical and Health Research Strategy; RF, Random Forest; ANN, Artificial Neural Network; SVM, Support Vector Machine; BN, Bayesian Network; RMSE, root mean squared error; ROC, receiver operating characteristic; ALM, Automatic Linear Modeling; RBF, radial basis function.

* Corresponding author at: College of Public Health and Health Informatics, King Saud bin Abdulaziz University for Health Sciences, Saudi Arabia.
E-mail address: sakrs@ksau-hs.edu.sa (S. Sakr).

https://doi.org/10.1016/j.ijcard.2019.01.046
0167-5273/

[3]. Patients receiving cardiac care have a wide variation in LOS which is influenced by several factors including patient and hospital characteristics [4,5]. The major risk factors of CVD include medical complexity as well as frailty [6]. This is coupled with the limited number of beds that results in cardiac departments having to operate at full capacity in certain admission seasons. Thus, unavailability of beds is a major concern and may significantly impact upon other services [4].

It is well documented that LOS is a metric of quality, efficiency and hospital performance. Predicting LOS accurately is one of the effective solutions to tackle capacity management, recourse planning and staffing levels [4]. In addition, it significantly impacts the fluency of the institutional workflow, the optimization of resources consumption and the reduction of healthcare costs. Recently, machine learning (ML) techniques are gaining increasing momentum and a lot of attention [7]. In practice, accurate LOS estimation has a positive impact on different health care outcomes such as increasing the number of patient receiving treatment, fluent institutional workflow, optimizing consumption of resources, improving patients' safety and the reduction in healthcare costs [8–10]. The aim of this study is to develop a robust prediction model for in-hospital length of stay for cardiac patients using ML techniques.

2. Material and methods

2.1. Study setting and population

The dataset of this study included adult patients who were admitted between 2008 and 2016 in King Abdulaziz Cardiac Center (KACC). The center is located in King Abdulaziz Medical City Complex in Riyadh, the capital of Saudi Arabia. Chronologically, KACC started to operate in 1999 with limited number of staff, patients and beds. It has gone through several expansional initiatives since that time. Recently, the total bed capacity has reached 101 beds, 45 of which are for intensive care [11]. In 2008, electronic medical records were provided using structured reporting. These data elements were entered by trained physicians and each record is audited by dedicated trained technical and medical teams [12]. The KACC provide its service to a wide range of population in the central region of the kingdom with both governmental and private coverage.

2.2. Data extraction and inclusion

A total of 16,917 cardiac patients records and 70 attributes were extracted from electronic medical records. These attributes were collected prospectively by experienced physicians and trained nurses. The attributes of the dataset included demographic information, cardiovascular risk factors, vital signs on admission, frequent laboratory test and admission criteria. Admission diagnosis and the primary treating physician were reviewed to verify selection criteria. We included all patients who were admitted to the cardiac wards under the cardiology clinical service. Patients who were admitted under other services in the cardiac wards (due to bed availability issues) for non-cardiac-related admitting diagnosis were excluded (<100 patients in the study period).

2.3. Data definition

The top twenty attributes that were selected in the final model based on their information gain (the amount of information gained about the predicted class given a particular attribute). The attributes included are demographic, cardiovascular risk factors, admission and discharge diagnosis, vital signs and laboratory tests on admission. Obesity was defined by body mass index (BMI) when the value is >30 kg/m^2. Diabetes, hypertension and dyslipidemia were defined based on medications utilization or previous diagnosis of patients. The final diagnosis was made by the clinical team based on the American College of Cardiology and American Heart Association guidelines. Given the circadian variation in the admission rate, a new variable was made based on whether the admission occurred in the hot summer time. Physician experience was determined by numbers of years of practice for each physician.

2.4. Data preprocessing and discretization

The length of stay was determined by subtracting the discharge date and time from the admission date and time. For the 262 patients who died during the hospital admission, the death date/time was considered to be the discharge time. Patients were divided into three groups based on the tertiles of LOS. Then, we applied the feature selection process where we rank the attributes of the dataset according to their information gain and select the subset with the highest gain. Information gain measures how much information an attribute gives us about the class to be predicted. Attributes that perfectly partition that data such that the data in each class is homogenous are given high information gain [35]. Formally the information gain of an attribute A is defined as:

$$IG(A) = H(S) - \sum_i \frac{S_i}{S} H(S_i)$$

where $H(S)$ is the entropy of the dataset and $H(S_i)$ is the entropy of the i_{th} subset generated by partitioning dataset S based on a specific attribute A. The machine learning models were evaluated using 10-fold cross-validation. Fig. 1 shows the steps of our modeling process and the following subsections describe the details of each step.

2.4.1. Discretization

This step reduces the number of values for continuous attributes. That is done by splitting the range of the continuous attribute into nominal attributes. In general, discretization reduces the time needed to build the prediction model and improves the prediction results [13]. The length of stay attribute has been discretized into three major groups: short (0–2 days), medium (3–5 days), and long (>5 days) using equal-frequency binning.

2.4.2. Feature selection

Feature selection is an essential part of building a good model for many reasons. It implies some degree of cardinality reduction by reducing the number of attributes used to build the model. That can be done by only choosing the most important attributes that improves the prediction accuracy. Another advantage of the feature selection process is reducing the resources (time and space) needed to build the model. In this study, as there were many attributes, the InfoGain module of Weka software has been used. The information gain for each attribute was calculated using the ranker search method [14].

2.5. Machine learning classification techniques

In practice, the performance of ML models can vary from one dataset to another based on the characteristics of the attributes and the outcome. Therefore, in this study, several models and techniques were evaluated and compared to select the model that achieves the best performance for our purpose. In particular, four different classification techniques have been evaluated: Random Forest (RF), Artificial Neural Network (ANN), Support Vector Machine (SVM) and Bayesian Network (BN).

***Random Forest* (*RF*)** is a classification algorithm that works by forming multitude decision trees at training and at testing where it outputs the class that is the mode of the classes (classification). Decision tree works by learning simple decision rules extracted from the data features. The deeper the tree, the more complex the decision rules and the fitter the model. Random decision forests overcome the problem of over fitting of the decision trees. To classify a new record from an input vector, put the input vector down each of the trees in the forest. Each tree votes for a specific class label. The final class label will be the one that have the most votes over all the trees in the forest. In practice, RF is quite a fast classification technique and is able to deal with the unbalanced large datasets [15].

Artificial neural networks (ANNs). ANN tries to mimic the human brain in order to model complicated task. ANN model consists of a set of interconnected nodes, called neuron that is connected in the same way like neurons in the human brain. Every neuron receives signals from the input or other neurons through weighted connections. The signals are then weighted and added together before being applied to a function to produce the output. The output signals are then propagated to other neurons until the output of the network is reached. We train the Neural Networks with gradient descent backpropagation. We varied the number of hidden units {1, 2, 4} and the momentum {0, 0.2, 0.5} [16].

Support Vector Machine (SVM) is a classifier that aims to output an optimal hyperplane in the N-dimensional space, where N is the number of the features used to classify the data. There are many possible ways to choose a hyperplane that separates two classes. Our objective is to find a plane that maximizes the maximum margin which is defined to be the maximum distance between data points belonging to different classes so that any test point can be classified confidently to a particular class. We tested SVM using polynomial, normalized polynomial, puk kernels and vary the complexity parameter {0.1, 10, and 30} [17].

Bayesian Network (BN) is a simple probabilistic classifier that is considered a generalization of the Naive Bayes [18]. BN belong to the category of probabilistic graphical models in which knowledge about uncertain domain is represented as graphical structures. Specifically, each node in the graph represents a random variable and edges represent the probabilistic dependencies among the corresponding random variables. BN combined principles from different domains including probability theory, statistics and graph theory.

2.6. Model evaluation and validation

In order to evaluate our models, we used the 10-fold cross-validation method. Cross-validation is a resampling technique used to evaluate a machine learning technique. 10-fold cross validation refers to the number of groups in which the dataset is partitioned which is 10. In the following we summarize the steps of the 10-fold cross validation as shown in Fig. 3. First, shuffle the dataset randomly. Second, partition the dataset into 10 groups. For each unique group, take this group as a testing dataset to evaluate the model on and the other nine groups are used to train the model. All results of the different

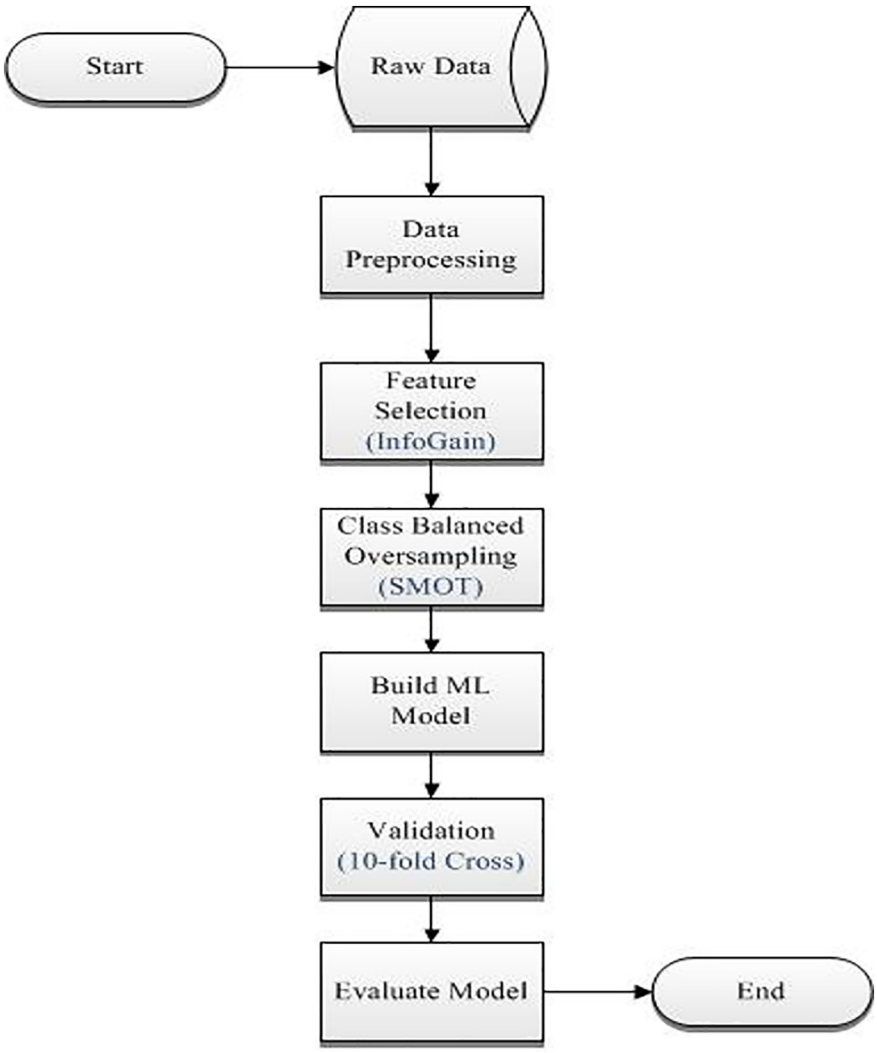


Fig. 1. Machine learning method to predict length of stay of cardiac patients.

metrics used for model evaluation are then averaged to get the final result. In general, a main advantage of the 10-fold cross-validation evaluation method is that it has a lower variance than a single hold-out set evaluator. In particular, it reduces this variance by averaging over 10 different partitions, therefore, it is less sensitive to any partitioning bias on the training or testing data and it is considered as a more robust validation method rather than the hold-out method which randomly splits the data into a single round of training and validation datasets [19].

3. Results

A total of (16,414) cardiac visits for 12,769 unique patients (mean age of 58.8 ± 16 years of which 68.2% were males) who were admitted to the cardiac center from 2008 and 2016 were analyzed. Of these, LOS classes included 30.85%, 33.45% and 35.71% for short, intermediate and long, respectively. The baseline characteristics are shown in Table 1.

As an outcome of the feature selection process, the ML models were developed using only 21 attributes (The top twenty attributes based on their information gain are; patients demographics (age and gender), vital sign on admission (heart rate, systolic and diastolic blood pressure, and body surface area), cardiovascular risk factors (hypertension, diabetes, dyslipidemia, smoking, and obesity), labs on admission (serum creatinine, high density lipoprotein, and ejection fraction), clinical diagnosis (congestive heart failure, acute coronary syndrome, acute myocardial infarction), and admission criteria (eligibility, seasonal admission, and physician experience)). Heart rate – Systolic blood pressure – Diastolic blood pressure – Age – insurance status (Eligibility) were ranked as the top significant five attributes. The full list of the outcome attributes is presented in Fig. 2.

Table 2 summarizes the performance of the different machine learning techniques using 10 fold cross validation. The Random Forest (RF) model significantly outperformed all other models for the Accuracy (80%), Sensitivity (80%), Precision (80%), F-score (80%), ROC (0.94) and RMSE (0.32) metrics. The SVM classifier is the second best performance achieving Accuracy (67%), F-score, (67%) and ROC (0.78) metrics. The performance of the BN and the ANN were very comparable as they achieved Area under Curve (AUROC) of 0.46 and 0.45 respectively, and with accuracies of 51% and 50% respectively.

To provide a comprehensive evaluation, we examined the diagnostic accuracy measures to assess the predictive ability for the prediction model. Table 3 shows the performance based on LOS class for all performance metrics (in percentage).

The performance of the SVM is reported in Table 4. Different kernels (polynomial kernel, normalized polynomial kernel and puk kernel) and complexity parameters (C) (0.1, 10 and 30) are tested. The results show that the AUROC increased as the complexity parameter increased up to 30 achieving AUROC of 0.78. Table 5 shows the performance of Neural Networks with gradient descent back propagation using hidden units H = {1, 2, 4, 8} and the momentum M = {0, 0.2, 0.5}. The number of hidden units and momentum rate that gives better AUROC value is considered here. Choosing H = 4 and M = 0.5 achieves AUROC = 0.67.

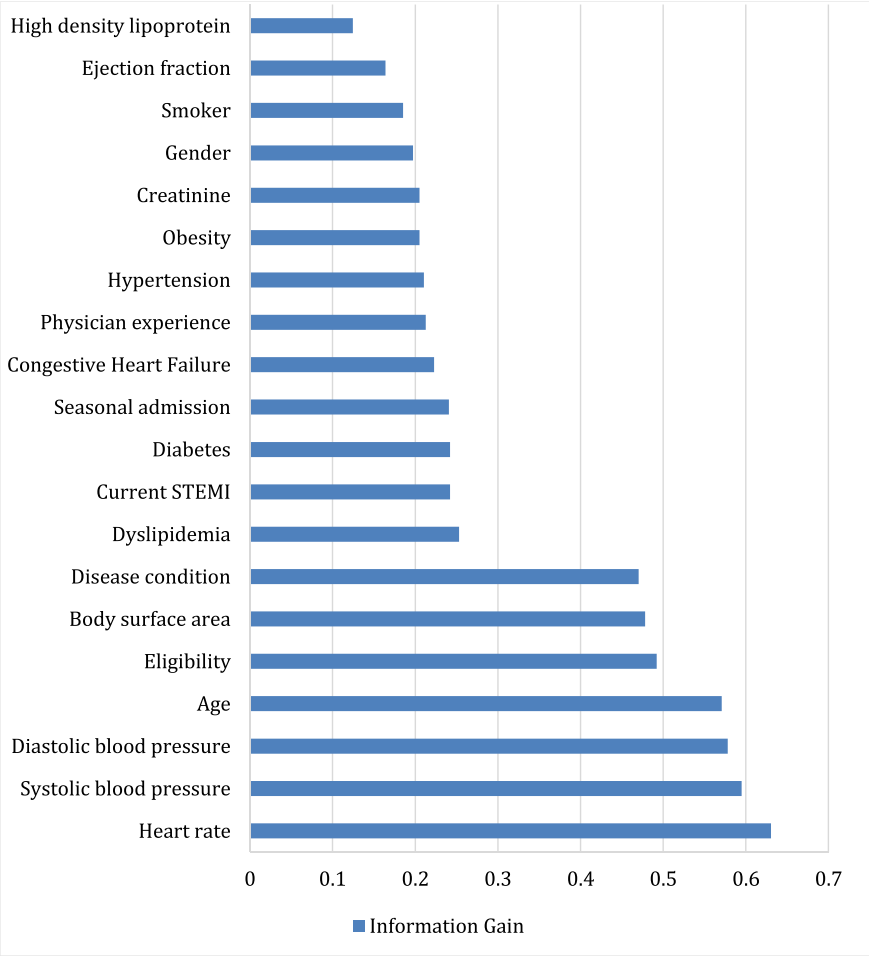


Fig. 2. Selected features for building the model for predicting the length of stay. The Information gain method has been used to identify the most important features that can significantly influence the accuracy of the model. In particular, the techniques are used to identify and remove irrelevant features by ranking all attributes based on their importance. The top twenty attributes were selected and included in the final model.

4. Discussion

This study aimed to develop a machine learning-based model for predicting LOS at admission time for cardiac patients. We analyzed the historical and vital signs of 16,917 patients' by reviewing their admission records at a tertiary care center. The results indicated that the Random Forest prediction model was the most accurate. Our analysis showed that the most important attributes that have the highest influence on the LOS are: heart rate, systolic blood pressure, diastolic blood pressure, age and insurance status. The results have shown that it is critical to carefully explore and evaluate the performance of the machine learning models using various model evaluation methods as the prediction accuracy can significantly differ. These results confirm the explorative nature of the machine learning process that requires iterative and explorative

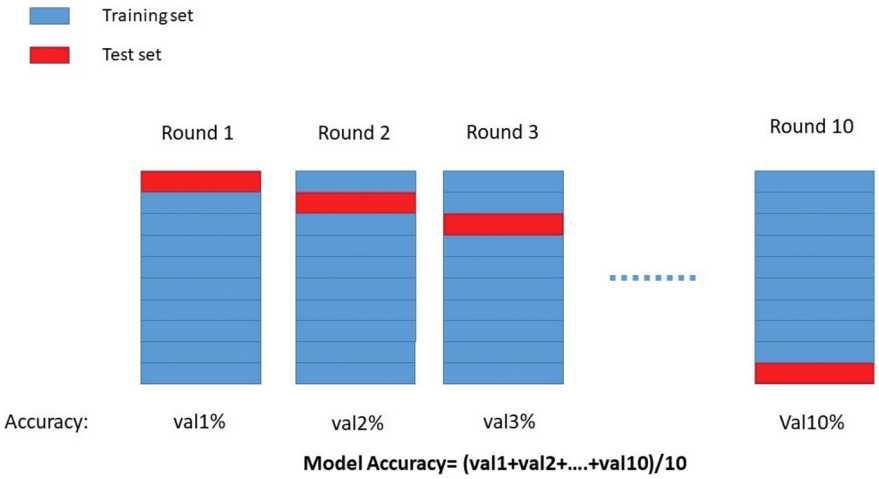


Fig. 3. Explaining the 10-fold cross validation process.

Table 1
Baseline characteristics.

Variables	Total n = 16,414	Length of stay (LOS)			p-Value
		Short (<3 days) n = 5063	Intermediate (3–5 days) n = 5490	Long (>5 days) n = 5861	
Age	58.8 ± 15.6	56.1 ± 15.8	58 ± 15.3	62.1 ± 15.1	<0.001
Gender	11,200 (68.2%)	3501(69.1%)	3955 (72%)	3744 (63.9%)	<0.001
Body mass index	105.1 ± 8663.8	36.9 ± 27.5	241.1 ± 14,980.6	36.7 ± 30.5	0.865
Body surface area	1.8 ± 0.4	1.9 ± 0.7	1.8 ± 0.3	1.8 ± 0.2	0.21
Deceased	262 (1.6%)	85 (1.7%)	42 (0.8)	135 (2.3%)	0.009
Cardiovascular risk factors					
Diabetes	9132 (55.6%)	2597 (51.3%)	2881 (52.5%)	3654 (62.3%)	0.001
Hypertension	9160 (55.8%)	2712 (53.6%)	2876 (52.4%)	3572 (60.9%)	0.138
Dyslipidemia	8456 (51.5%)	2604 (51.4%)	2712 (49.4%)	3140 (53.6%)	<0.001
Smoker	3901 (23.8%)	1217 (24.0%)	1527 (27.8%)	1157 (19.7%)	0.04
Obesity	5359 (32.6%)	1677 (33.1%)	1725 (31.4%)	1957 (33.4%)	0.541
Clinical diagnosis					
Acute coronary syndrome	5848 (35.6%)	1161 (22.9%)	2500 (45.5%)	2187 (37.3%)	<0.001
Congestive heart failure	4144 (25.2%)	750 (14.8%)	1107 (20.2%)	2287 (39.0%)	<0.001
Renal failure	1275 (7.8%)	192 (3.8%)	329 (6%)	754 (12.9%)	<0.001
Unstable angina	1205 (7.3%)	377 (7.4%)	470 (8.6%)	358 (6.1%)	0.387
Disease condition	0: 5848 (35.6%)	0: 1161 (22.9%)	0: 2500 (45.5%)	0: 2187 (37.3%)	<0.001
	1: 3143 (19.1%)	1: 620 (12.2%)	1: 818 (14.9%)	1: 1705 (29.1%)	
	2: 2134 (13.0%)	2: 897 (17.7%)	2: 647 (11.8%)	2: 590 (10.1%)	
	3: 5289 (32.2%)	3: 2385 (47.1%)	3: 1525 (27.8%)	3: 1379 (23.5%)	
Vital signs on admission					
Heart rate	80.24 ± 64	77.1 ± 18.5	78.9 ± 22.3	84.2 ± 103.4	<0.001
Systolic blood pressure	135.2 ± 269.1	136.8 ± 289.2	134.9 ± 244.5	134.1 ± 273.1	0.083
Diastolic blood pressure	71.2 ± 54.3	70.9 ± 22.9	72.7 ± 88.1	70.2 ± 22.7	0.033
Admission criteria					
Eligibility	0: 324 (2%)	0: 130 (2.6)	0: 96 (1.7%)	0: 98 (1.7%)	0.002
	1: 59 (0.4%)	1: 6 (0.1/%)	1: 22 (0.4%)	1: 31 (0.5%)	
	2: 1694 (10.3%)	2: 537 (10.6%)	2: 499 (9.1%)	2: 658 (11.2%)	
	3: 2997 (18.3%)	3: 845 (16.7%)	3: 921 (16.8%)	3: 1231 (21%)	
	4: 11,340 (69.1%)	4: 3545 (70.0%)	4: 3952 (72%)	4: 3843 (65.6%)	
Seasonal admission	4862 (29.6%)	1452 (28.7%)	1677 (30.5%)	1733 (29.6%)	0.127
Physician experience					0.011
Short experience	1016 (6.2%)	502 (9.9%)	300 (5.5%)	214 (3.7%)	
Medium experience	3889 (23.7%)	1106 (21.8)	1308 (23.8%)	1475 (25.2%)	
Long experience	11,509 (70.1%)	3455 (68.2)	3882 (70.7%)	4172 (71.2%)	
Lab results on admission					
HgbA1C	7.3 ± 11.7	7.3 ± 13.0	7.3 ± 13.2	7.2 ± 8.8	0.62
Low density lipoprotein	3.30 ± 21.9	3.7 ± 33.1	3.3 ± 17.0	3.0 ± 11.5	0.334
High density lipoprotein	1.6 ± 16.7	1.8 ± 21.1	1.7 ± 19.9	1.3 ± 4.7	0.411
Creatinine clearance	92.4 ± 77.5	1,03.3 ± 122.1	95.1 ± 46.8	80.6 ± 40.7	<0.001
Peak troponin	1790.9 ± 8846.8	1012.9 ± 3340.7	2666.5 ± 12,950.7	1642.7 ± 7155.8	0.49
Peak creatinine kinase	5,62.5 ± 3496.8	396 ± 930.2	6,44.6 ± 1248.7	6,29.3 ± 5657.3	0.282
Creatinine	106.84 ± 118.1	96.3 ± 126.3	101.0 ± 83.8	121.4 ± 135.5	<0.001

Values are mean ± SD, n (%) and *p* value.
p value is statistically highly significant as $p < 0.001$.

experiments in order to discover the model design that can achieve the target accuracy for a specific problem. In particular, there is no one-size-fits-all machine learning model, i.e., no single model can outperform all other models for all datasets for a specific domain.

In practice, the available admission information about the current state and previous hospitalization information of the cardiac patients represents an important source for predicting LOS which is crucial factor for avoiding unnecessary long-term care [21–23]. Machine learning has been successfully implemented in the medical field using structured data and images and has been shown to provide unique potential benefits [16].

Table 2
The performance of the different machine learning models evaluated using 10-fold cross validation method.

	Random Forest (RF)	BN	SVM	Artificial neural networks (ANN)
Sensitivity	80%	50%	67%	45%
Accuracy	80%	51%	67%	50%
Precision	80%	50%	67%	50%
F-score	80%	50%	67%	49%
RMSE	0.31	0.5	0.32	0.5
ROC	0.94	0.7	0.78	0.7

In the cardiac field, discovering data and relationships among a set of factors enables constructing a model that is able to predict LOS accurately. A precise estimation of the patients' length of stay would be of a significant value for systematically managing both hospital unit

Table 3
Diagnostic accuracy measures for ML prediction model using Random Forest.

	<3 days	3 to 5 days	>5 days
Sensitivity	82.64% (81.57% to 83.68%)	75.26% (74.10% to 76.40%)	83.16% (82.18% to 84.11%)
Specificity	90.51% (89.96% to 91.04%)	91.21% (90.67% to 91.74%)	88.68% (88.06% to 89.27%)
Positive likelihood ratio	8.71 (8.22 to 9.23)	8.56 (8.05 to 9.11)	7.34 (6.95 to 7.76)
Negative likelihood ratio	0.19 (0.18 to 0.20)	0.27 (0.26 to 0.28)	0.19 (0.18 to 0.20)
Positive predictive value	79.50% (78.53% to 80.43%)	81.15% (80.18% to 82.08%)	80.31% (79.43% to 81.16%)
Negative predictive value	92.14% (91.69% to 92.56%)	88.01% (87.51% to 88.49%)	90.46% (89.95% to 90.94%)
Accuracy	88.09% (87.58% to 88.58%)	85.88% (85.34% to 86.41%)	86.71% (86.18% to 87.22%)

Table 4
Comparison of the performance of Support Vector Machine (SVM) classifier using polynomial, normalized polynomial and puk kernels using complexity parameters 0.1, 10 and 30.

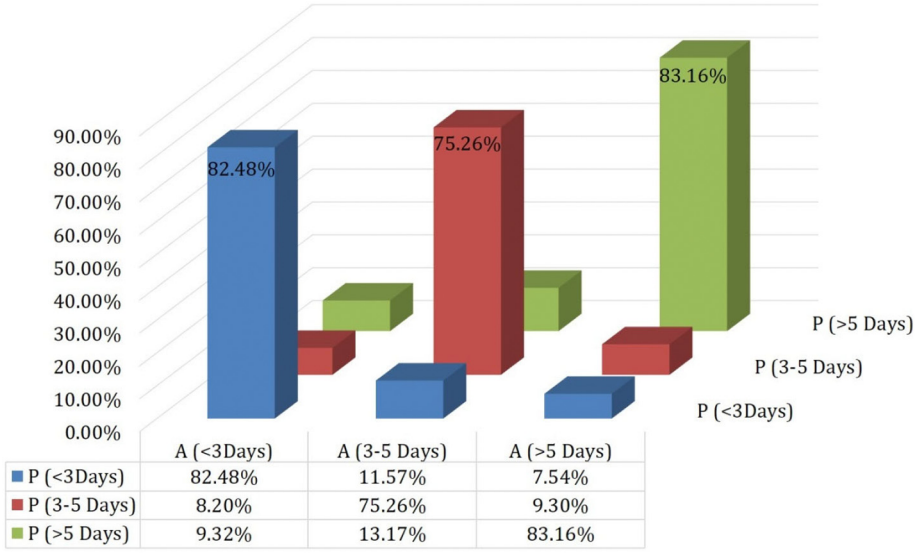
	Polynomial			Normalized polynomial			Puk		
	C = 0.1	C = 10	C = 30	C = 0.1	C = 10	C = 30	C = 0.1	C = 10	C = 30
Sensitivity	47%	48%	48%	48%	49%	50%	50%	64%	67%
Accuracy	47%	48%	48%	48%	49%	50%	50%	64%	67%
Precision	48%	49%	48%	49%	50%	50%	50%	65%	67%
F-score	50%	47%	47%	47%	49%	48%	25%	64%	67%
RMSE	0.49	0.39	0.48	0.49	0.48	0.48	0.48	0.42	0.32
ROC	0.62	0.63	0.63	0.62	0.64	0.64	0.64	0.76	0.78

Table 5
Comparison of the performance of Artificial neural networks (ANN) classifier with gradient descent backpropagation using hidden units {1, 2, 4} and the momentum {0, 0.2, 0.5}.

	H = 1			H = 2			H = 4		
	M = 0	M = 0.2	M = 0.5	M = 0	M = 0.2	M = 0.5	M = 0	M = 0.2	M = 0.5
Sensitivity	46%	46%	66%	50%	50%	50%	50%	48%	45%
Accuracy	46%	46%	46%	50%	50%	50%	50%	48%	50%
Precision	30%	30%	43%	50%	50%	50%	50%	50%	50%
F-score	37%	37%	41%	49%	49%	49%	49%	47%	49%
RMSE	0.5	0.46	0.46	0.5	0.45	0.45	0.5	0.47	0.45
ROC	0.6	0.63	0.63	0.7	0.66	0.66	0.7	0.66	0.67

resources (medication, equipment, beds) and the distribution of personnel. It also plays an important role for performance evaluation purposes which represents important aspects in the new trend of performance-based budgeting [24,25].

Several previous studies have explored the key risk factors for LOS. For example, a study was conducted to identify and quantify the influence of different patient-related or diagnostic activities-related factors on the emergency department LOS of discharged patients [27]. The study was a retrospective electronic data analysis where all patients who were discharged from the emergency department of a tertiary teaching hospital in 2013 were included. A multivariate accelerated failure time model was used to analyze the influence of the collected covariates on patient length of stay. The study developed a model for predicting the length of stay in an emergency that may provide useful information for physicians or patients to better anticipate an individual's length of stay and to help the administrative department to better plan its staffing policy [27]. Another study was conducted to determine the best predictors of hospital length of stay and discharge destination in patients admitted to a neuroscience service. The study collected data for 170 patients with a set of attributes that included age, gender, location prior to admission, principle diagnosis, various physiological measurements upon admission, comorbidity, independence in various activities of daily living prior to admission, length of stay, and disposition upon discharge. The study design was a correlational descriptive study performed through the analysis of data and the development and validation of statistically significant factors in determining the length of stay. The study showed that functional status at admission, specifically, a higher modified Rankin score and a lower systolic blood pressure along with the acquisition of deep vein thrombosis, catheter associated urinary tract infections, intubation, and admission to an intensive care unit all have a statistically significant effect on the hospital length of stay [28]. Moreover, a study used Automatic Linear Modeling (ALM) to identify the factors associated with LOS in the intensive care unit following cardiac surgery. The pre-incision (i.e. prior to initiation of surgery) factors were used as input to develop the Artificial Neural Network (ANN) model to detect patients with the highest risk of prolonged stay. The accuracy of ANN showed the highest accuracy and demonstrated that the neural network model's has higher capacity



	A (<3Days)	A (3-5 Days)	A (>5 Days)
P (<3Days)	82.48%	11.57%	7.54%
P (3-5 Days)	8.20%	75.26%	9.30%
P (>5 Days)	9.32%	13.17%	83.16%

Fig. 4. Comparison matrix of result of prediction model for each class (in percentage). Identify unpredicted values in each class and determine the assigned class for these values. In addition to determine the percentage of correct prediction.

and strength over the linear model [29]. Another study used a Cubist model for predicting LOS at admission time; this would provide a more accurate prediction than other techniques. The length of stay has been predicted using regression tree based on static features that are known at the time of admission [30]. The model has been trained and validated using Veterans Health Administration (VHA) hospitals in n Pittsburgh and achieved accuracy of 83%. Another study aimed to develop a predictive model to predict whether patients' LOS is with the standard LOS after surgery [31]. The predictive model used Random Forest classifier on complete historical medical records and lab data of 896 clinical cases which are divided into two groups urgent operation group and non-urgent group. For the urgent-group and non-urgent group the model achieves 85.7% and 89.4% respectively. The length of stay and mortality have been predicted using 14,480 patient dataset using different machine learning techniques. For predicting mortality, support vector machine was the best performing technique achieves AUROC of 0.77. For predicting the length of stay, support vector regression was the best performing model achieves a mean absolute error of 1.79 days for those patients surviving a non-prolonged stay.

Actually, it has been clearly shown that no model fits all the requirements particularly that of unexpected LOS due to the complexities of the various conditions [33]. More attention to outlier cases based on the result of prediction at the preadmission stage results in the reduction of bed transfers during hospitalization [5]. Further, understanding the characteristics of long stay patients and comparing it to other patients with similar cases and shorter length of stay is important to address the variation in processes that lead to higher LOS. Having a better understanding of the characteristics enables the development of more effective action plans to eliminate unnecessary procedures that will result in reduced length of stay duration. Accordingly, it will reduce cost, decrease unnecessary burden on patients and healthcare providers [34]. As a result of providing accurate prediction, the management of hospital and health care system will be more efficient. Here, we formed a comparison matrix (Fig. 4) to present the percentage where one class was

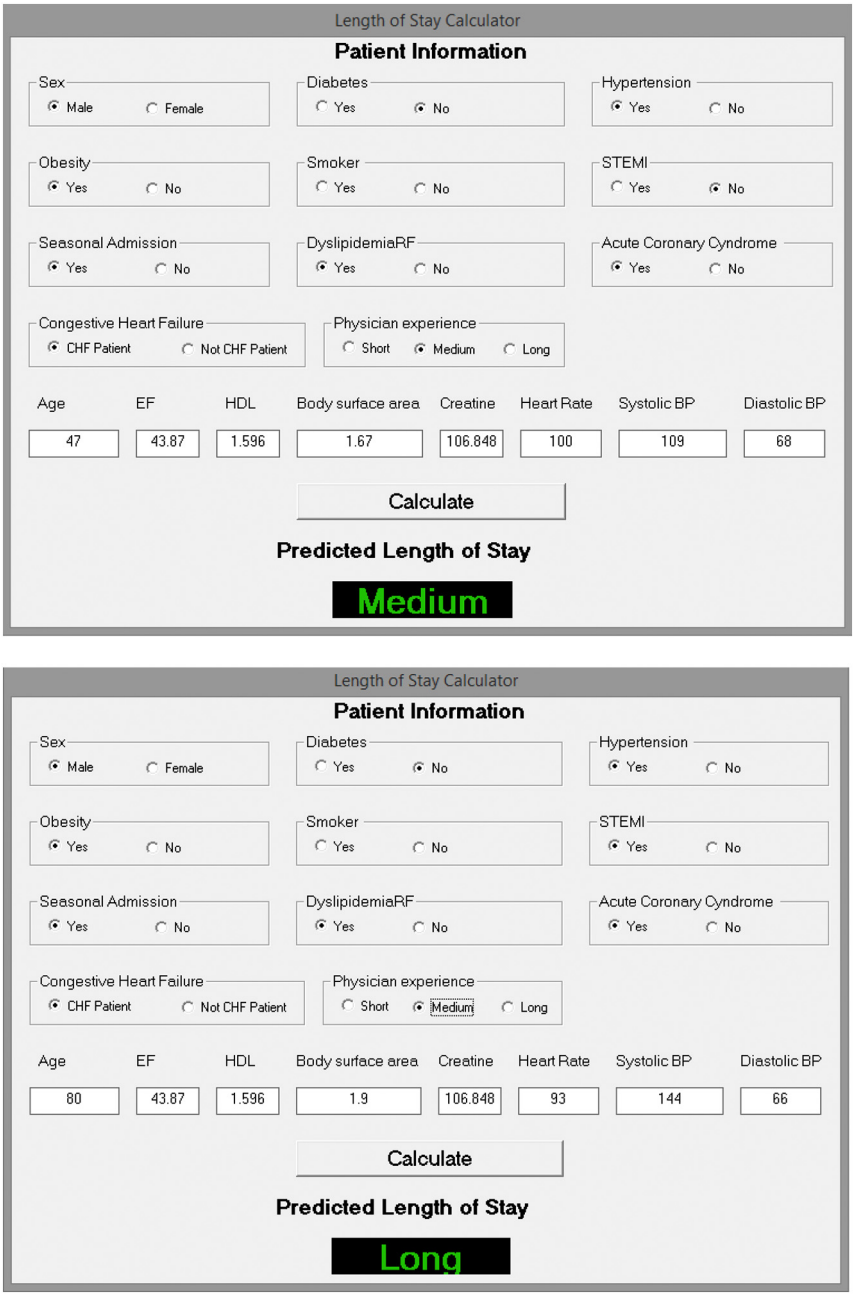


Fig. 5. Length of stay application built base on prediction model. This desktop application was built using IKVM framework to help the physician for timely predicting the length of stay.

incorrectly assigned to another class. Such matrix can help to identify opportunities to improve the accuracy of the model. This is in contrast to prior works that concentrated on identifying the results of predicted values (Fig. 5).

There are, however, several limitations of the current study that need to be recognized. As indicated in this study, we deal with a single health organization instead of several organizations, which may be not reflective and may also restrict generalizability. However, the population included in this study covers the entire kingdom of Saudi Arabia. We would like to highlight that the data came from a sick population which may have a negative impact on the performance of the model. In contrast, the use of data from other populations or applying the model in other health organization might have a different result. We also did not have a validation cohort which is an important step in our model assessment.

The developed model in this study has several advantages. In particular, the model is characterized by small number of attributes (20 attributes), yet provide an accurate assessment of the LOS. In addition, we are able to effectively predict the LOS using prospective data collected at admission time. Thus, the model is simple, accurate and can be easily adopted in practice. It achieved a high prediction accuracy which is superior to the results of models that were adopted in previous studies. To conclude, the prediction of LOS and determining the associated risks is a challenging undertaking. Our proposed prediction model supports physicians to improve the accuracy of their decisions at admission time.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The work of Radwa Elshawi and Sherif Sakr is funded by the European Regional Development Funds via the Mobilitas Plus programme (grant MOBTT75).

References

[1] (WHO) WHO, Cardiovascular Disease, 2017.
[2] (KACST) KACfSaT, Strategic Priorities for Advanced Medical and Health Research, 2017.
[3] A. Cedars, L. Benjamin, S.V. Burns, E. Novak, A. Amin, Clinical predictors of length of stay in adults with congenital heart disease, Heart (2017) (heartjnl-2016-310841).
[4] A. Almashrafi, M. Elmontsri, P. Aylin, Systematic review of factors influencing length of stay in ICU after adult cardiac surgery, BMC Health Serv. Res. 16 (2016) 318.
[5] P.-F.J. Tsai, P.-C. Chen, Y.-Y. Chen, H.-Y. Song, H.-M. Lin, F.-M. Lin, Q.-P. Huang, Length of hospital stay prediction at the admission stage for cardiology patients using artificial neural network, J. Healthc. Eng. 2016 (2016).
[6] K. Fatema, N.A. Zwar, A.H. Milton, L. Ali, B. Rahman, Prevalence of risk factors for cardiovascular diseases in Bangladesh: a systematic review and meta-analysis, PLoS One 11 (2016), e0160180.
[7] N. Caetano, P. Cortez, R.M. Laureano, Using data mining for prediction of hospital length of stay: an application of the CRISP-DM methodology, International Conference on Enterprise Information Systems 2014, pp. 149–166.
[8] P.R. Hachesu, M. Ahmadi, S. Alizadeh, F. Sadoughi, Use of data mining techniques to determine and predict length of stay of cardiac patients, Healthc. Inform. Res. 19 (2013) 121–129.
[9] F. Barili, N. Barzaghi, F.H. Cheema, A. Capo, J. Jiang, E. Ardemagni, M. Argenziano, C. Grossi, An original model to predict Intensive Care Unit length-of stay after cardiac surgery in a competing risk framework, Int. J. Cardiol. 168 (2013) 219–225.
[10] D. Zikos, K. Tsiakas, F. Qudah, V. Athitsos, F. Makedon, Evaluation of classification methods for the prediction of hospital length of stay using medicare claims data, Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments, 54, , 2014.
[11] Affairs MoNGfH. King Abdulaziz Cardiac Center. 2014.
[12] M. Al Zaibag, J.G. Franke, King Abdulaziz cardiac informatics program: an overview, Eur. Heart J. Suppl. 16 (2014) B3–B6.
[13] U. Fayyad, K. Irani, Multi-interval discretization of continuous-valued attributes for classification, Learning (1993), https://www.ijcai.org/Proceedings/93-2/Papers/022.pdf.
[14] J.T. Kent, Information gain and a general measure of correlation, Biometrika 70 (1983) 163–173.
[15] A. Liaw, M. Wiener, Classification and regression by randomForest, R news 2 (2002) 18–22.
[16] R.J. Schalkoff, Artificial Neural Networks, McGraw-Hill, New York, 1997.
[17] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297.
[18] N. Friedman, D. Geiger, M. Goldszmidt, Bayesian network classifiers, Mach. Learn. 29 (1997) 131–163.
[19] P. Refaeilzadeh, L. Tang, H. Liu, Cross-validation, Encyclopedia of Database Systems, Springer 2009, pp. 532–538.
[21] E. Bruballa, A. Wong, F. Epelde, D. Rexachs, E. Luque, A model to predict length of stay in a hospital emergency department and enable planning for non-critical patients admission, Int. J. Integr. Care 16 (2016).
[22] J.P. Curtain, M. Yu, A.B. Clark, N.D. Gollop, J.H. Bettencourt-Silva, A.K. Metcalf, K.M. Bowles, M.D. Flather, J.F. Potter, P.K. Myint, Determinants of length of stay following total anterior circulatory stroke, Geriatrics 2 (2017) 26.
[23] S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, S. Levin, Real-time prediction of inpatient length of stay for discharge prioritization, J. Am. Med. Inform. Assoc. 23 (2015) e2–e10.
[24] E.M. Carter, H.W. Potts, Predicting length of stay from an electronic patient record system: a primary total knee replacement example, BMC Med. Inform. Decis. Mak. 14 (2014) 26.
[25] A.R. Al Taleb, M. Hoque, A. Hasanat, M.B. Khan, Application of data mining techniques to predict length of stay of stroke patients, Informatics, Health & Technology (ICIHT), International Conference on 2017, pp. 1–5.
[27] C.-H. Chaou, H.-H. Chen, S.-H. Chang, P. Tang, S.-L. Pan, A.M.-F. Yen, T.-F. Chiu, Predicting length of stay among patients discharged from the emergency department—using an accelerated failure time model, PLoS One 12 (2017), e0165756.
[28] M. Stecker, M. Stecker, J. Falotico, Predictive model of length of stay and discharge destination in neuroscience admissions, Surg. Neurol. Int. 8 (2017).
[29] R.J. LaFaro, S. Pothula, K.P. Kubal, M.E. Inchiosa, V.M. Pothula, S.C. Yuan, D.A. Maerz, L. Montes, S.M. Oleszkiewicz, A. Yusupov, Neural network prediction of ICU length of stay following cardiac surgery based on pre-incision variables, PLoS One 10 (2015), e0145395.
[30] L. Turgeman, J.H. May, R. Sciulli, Insights from a machine learning model for predicting the hospital Length of Stay (LOS) at the time of admission, Expert Syst. Appl. 78 (2017) 376–385.
[31] M.T. Chuang, Yh. Hu, C.L. Lo, Predicting the prolonged length of stay of general surgery patients: a supervised learning approach, Int. Trans. Oper. Res. 25 (2018) 75–90.
[33] I.W.M. Verburg, A. Atashi, S. Eslami, R. Holman, A. Abu-Hanna, E. de Jonge, N. Peek, N.F. de Keizer, Which models can I use to predict adult ICU length of stay? A systematic review, Crit. Care Med. 45 (2017) e222–e231.
[34] D. Gartner, R. Padman, Mathematical Modelling and Cluster Analysis in Healthcare Analytics-The Case of Length of Stay Management, 2016.
[35] M. Hall, G. Holmes, Benchmarking attribute selection techniques for discrete class data mining, IEEE Trans. Knowl. Data Eng. 15 (2003) 1437–1447.