Healthcare Analytics 4 (2023) 100245

Contents lists available at ScienceDirect

Healthcare Analytics

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

A forecasting approach for hospital bed capacity planning using machine learning and deep learning with application to public hospitals

Younes Mahmoudian, Arash Nemati *, Abdul Sattar Safaei

Department of Industrial Engineering, Babol Noshirvani University of Technology, Babol, Iran

ARTICLE INFO

Keywords:
Machine learning
Deep learning
Hospital bed capacity forecasting
Data analytics
Classification
Statistical inference

ABSTRACT

Hospital Bed Capacity (HBC) planning affects economic and social sustainability in healthcare through bed capacity efficiency and medical treatment accessibility. Conventionally, this problem is solved using programming or simulation models with assumptions and limits. Forecasting the HBC using time series data on bed occupancy has been considered but not with factors such as the Number of Hospitalized Patients (NHP) and patient's length of stay (LOS). This study proposes a data-driven methodology to forecast the HBC using Machine Learning (ML) and Deep Learning (DL). The LOS classification is performed using several ML techniques, including the Bayesian network, K-nearest neighbor, support vector machine, decision tree, and Linear regression. Also, the seasonal autoregressive integrated moving average, linear regression and Long short-term memory neural network are applied for the NHP forecasting. The forecasting and descriptive analysis outputs based on LOS classes are directly applied to a simple mathematical model to predict the required bed capacity. This methodology is applied in a case study in a heart ward at a public hospital. The data set includes 51231 records, and ML and DL algorithms are developed in Python. Results show that the heart ward's bed capacity must be raised from 45 to 137 by 2026. In addition, several managerial recommendations are formulated.

1. Introduction

The Hospital Bed Capacity (HBC) forecasting problem has taken significant attention because of its effects on the sustainability of hospitals, particularly in terms of hospital economic efficiency, and patient satisfaction [1–7]. The traditional approach to this problem is using simulation or programming models involving several issues, such as the need for some assumptions on attributes of some quantities, for example, the probability distribution of some factors [8–11]. In addition, reaching optimum or reasonable solutions is a big challenge of the bed capacity programming models, particularly in the case of large-scale, multi-objective, and integer models. Consequently, using model-free methods for HBC forecasting seems to be a facilitator. Nowadays, business analytics, including Data Analysis (DA), Machine Learning (ML), and Deep Learning (DL) techniques, are widely used as model-free methods in different businesses of services and manufacturing to reach insights on market trends, such as customers' behavior, costs, and technology by no reliance on the simulation or mathematical models [12–14]. Also, a reasonable number of their successful applications in the healthcare sector, such as Length of Stay (LOS) forecasting, patient classification, healthcare resources forecasting, and disease diagnosis have been illustrated [15–18]. In recent years, a few researchers applied some forecaster tools of ML to predict the required bed capacity based on time-series data sets of bed occupancy [19,20]. But they fail to consider the direct impacts of LOS and the Number of Hospitalized Patients (NHP) as the most significant factors affecting bed occupancy [21].

This paper deals with the mentioned-above gap by considering LOS and NHP in the data-driven approach of HBC. A hybrid approach, including DA, ML, DL, statistical inference, and mathematical model, is proposed to reach a reasonable forecast of bed occupancy in the future. In this methodology, several algorithms, such as Decision Tree (DT), Long Short-Term Memory (LSTM) neural network, Support Vector Machine (SVM), Bayesian network (BN), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Linear Regression (LR), are applied in the LOS classification and NHP forecasting. Also, DA techniques are used for investigating the LOS, patients' ages, and NHP in detail, singularly and simultaneously, to provide beneficial managerial insights into the HBC forecasting problem. In brief, using LOS classification and proposing a DA-ML-DL-based framework for HBC forecasting are the main contributions of this paper. This hybrid approach provides many useful descriptive, diagnostics, predictive, and to some extent, prescriptive analytics in the HBC forecasting problem. This methodology is applied using a real data set of a Heart ward of a public Hospital in Babol City.

* Corresponding author.
E-mail addresses: younesmahmoudian@nit.ac.ir (Y. Mahmoudian), r.nemati@nit.ac.ir (A. Nemati), s.safaei@nit.ac.ir (A.S. Safaei).

https://doi.org/10.1016/j.health.2023.100245
Received 22 May 2023; Received in revised form 30 July 2023; Accepted 10 August 2023

The rest of the paper is as follows. Section 2 reviews the corresponding literature, and Section 3 introduces the proposed data-driven methodology. Section 4 contains the computational results of applying the proposed hybrid data-driven approach in the case study, and finally, Section 5 involves the conclusions, limitations, and recommendations for future works.

2. Literature review

The literature on healthcare centers' bed capacity forecasting is reviewed to clear the paper's contributions and novelty aspects.

Bachouch et al. (2012) considered some limitations, such as budget, shared resources, and beds needed by acute and emergency patients, in the problem of hospital bed planning. Using constraints such as incompatibility between pathologies and continuity of care, an integer linear programming model was proposed, and the results were compared using several solvers such as the CPLEX [22]. Ben Abdelaziz and Masmoudi (2012) studied the bed capacity of 157 public hospitals in Tunisia when the demand for beds is random using a multi-objective stochastic programming model to allocate the beds to the hospital's departments. In this problem, the three objectives are the total cost of creating new beds, the number of nurses and doctors, and the optimal number of beds [11]. The healthcare capacities assessment is a matter of great importance in the same vein. Devapriya et al. (2015) proposed a discrete-event simulation model to determine the required bed capacity using forecasted patients' volume and LOS [23]. Also, Keegan et al. (2018) used an expanded HIPPOCRATES macro-simulation method to predict the number of beds needed for public and private hospitals in Ireland by 2030. Baseline year and projected demand and capacity profiles in this analysis are generated through the HIPPOCRATES macro-simulation model of demand and cost of health care developed by the ESRI.

The emergency department plays a vital role in the hospital because it is the first deal for patients. Therefore, the effective management of the emergency department is significant in improving the quality of service and treatment. Nas and Koyuncu (2019) used ten ML tools to predict the patients' arrival rates. They applied a simulation model to determine the required number of beds in the emergency department by minimizing the LOS in a case study in Turkey [24]. Kutafina et al. (2019) predicted hospital bed occupancy using recurrent neural networks based on 353520-time series records and four features from a hospital in Germany. The results show that the proposed ML model is a powerful tool for automatic planning and decision-making on existing bed capacity planning [25]. Zhu et al. (2020) developed three mathematical models considering the uncertainty of demand for inpatient beds and the number of periods [26]. They solved and analyzed these models in a public hospital in China.

Predicting the required capacity of hospital wards during pandemics, such as Covid-19s, is essential for the hospital. Deschepper et al. (2021) used an incremental Poisson model on the previous patients' statistics to establish a Mont-Carlo simulation model for a 10-day prediction of needed beds in various wards of A hospital in Ghent. Haghshenas et al. (2021) proposed a bi-objective mixed-integer optimization model for Cancer hospitals' location–allocation problem under bed capacity and efficiency considerations [27]. They forecasted the cancer incidences in the provinces of Iran in 2040 using LR models to provide a plan for cancer hospitals established in each province in some predefined efficient bed capacities. Also, they applied this forecasting method to determine the demand parameter values in a single objective mathematical model for Cancer-curing supply chain planning [28]. The healthcare system of a country is very complex and vital, that is why Kabir et al. (2021) used Recurrent Neural Network (RNN) to predict population growth and applied the Markov decision process to simulate the number of required beds by the next 30 years in a case study of Bangladesh [29]. Ordu et al. (2021) proposed a linear programming model for bed capacity and staff requirement planning in a case study in England. They also forecasted the demand for specialists with ML tools and simulated the patients' treatment pathway using a discrete-event simulation model.

It is crucial to determine the operating room bed capacity according to the critical condition of the patients. Schiele et al. (2021) predicted the bed occupancy in the Intensive Care Unit (ICU) using an Artificial Neural Network (ANN) algorithm as an ML forecaster and 6000 patients' data from 2010 to 2016 in an ICU scheduling problem in Germany. High traffic of patients in hospitals' wards, in particular the emergency department, affects the capability to provide the optimal level of care. Tello et al. (2022) predicted weekly hospital bed capacity using LR as a forecaster per cluster of beds/day. They applied the K-means clustering method in SVM as the classifier of beds/day factor in a case study in the USA [19]. Latruwe et al. (2023) proposed a simulation model called the ProMoBed, for inpatient hospitals' bed capacity forecasting. They used LOS, seasonality, and admission data to simulate a stochastic pattern of demand for beds in a case study in Belgium [8]. Garcia-Vincuna et al. (2023) initially predicted the LOS using linear and non-linear programming and then simulated the bed prediction for the ICU department [30].

The Covid epidemic sent about 19 million people to hospitals worldwide and showed the importance of forecasting during epidemics. Redondo et al. (2023) used a discrete simulation model to predict hospital discharges for Covid patients [31]. Bekker et al. (2023) proposed a linear programming model for predicting hospital beds in the short term. The model predicts admissions and uses a queue-based model to occupy beds, providing accurate results for three days [32]. Widyasari et al. (2023) predicted the bed capacity of Malahayati Hospital in Indonesia using SVM and linear programming [33]. Covid has put a lot of pressure on healthcare resources worldwide, often challenging hospital capacity and stressing hospital staff, Johnson et al. (2023) predicted hospital resources using time series. Results showed that statistical forecasting and ML methods can provide valuable predictions to help make resource planning decisions during epidemics [34].

The above-reviewed literature, summarized in Table 1, shows that the commonly used techniques in HBC forecasting have been simulation and programming models. To our best knowledge, only three researchers, including Kutafina et al. (2019), Schiele et al. (2021), and Tello et al. (2022), applied ML tools in the healthcare centers' bed capacity planning. These papers used time series data on occupied beds and forecast the required beds in the short-run future, and ignore the direct effects of LOS and NHP.

As illustrated in Table 1, this paper uses a wide variety of techniques from DA, ML, DL, and Statistics to investigate the nature of significant factors, such as patients' LOS, Patients' Age, and NHP, to discover their historical manner and provide valuable information to forecast HBC. In addition, conducting LOS classification to use in HBC forecasting is a unique attribute of this paper.

3. A hybrid data-driven approach to HBC forecasting

In this section, the framework of the proposed hybrid data-driven approach to HBC forecasting is provided, and some classification and forecasting algorithms are opted according to the corresponding literature.

3.1. New framework for HBC forecasting

As mentioned in the previous sections, factors of LOS and NHP play the most significant role in the bed capacity forecasting problem. Therefore, conducting descriptive analyses on these factors and other affecting features, for example, patient's age and bed occupancy, provides great initial insights into the problem. Consequently, the LOS classification, NHP distribution across LOS classes, and NHP forecasting are the heart of the proposed bed capacity forecasting framework. The results of these analyses are applied in a simple assumption-free

Table 1
Hospital bed capacity forecasting literature.

Attribute								
Research	Programming model	Simulation	Mathematical model	Data analysis	Machine learning	Deep learning	Statistical inference	LOS classification
Bachouch et al. (2012)	✓							
Ben Abdelaziz and Masmoudi (2012)	✓							
Devapriya et al. (2015)		✓						
Keegan et al. (2018)		✓						
Nas and Koyuncu (2019)		✓			✓			
Kutafina et al. (2019)					✓			
Zhu et al. (2020)	✓							
Deschepper et al. (2021)		✓						
Haghshenas et al. (2021)	✓				✓			
Kabir et al. (2021)		✓				✓		
Ordu et al. (2021)	✓	✓			✓			
Schiele et al. (2021)					✓			
Tello et al. (2022)					✓			
Latruwe et al. (2023)		✓						
Garcia-Vincuna et al. (2023)	✓	✓						
Redondo et al. (2023)		✓						
Bekker et al. (2023)	✓							
Widyasari et al. (2023)	✓				✓			
Johnson et al. (2023)	✓							
This paper			✓	✓	✓	✓	✓	✓

mathematical model to forecast the required beds in the future. In addition, patients' age analysis could provide beneficial insight into the scatter of NHP according to Age, particularly in hospitals without specialized wards for children. Fig. 1 represents the structure and steps of the proposed framework.

According to Fig. 1, like all data-oriented analyses, data collection and pre-processing activities, such as data set cleaning and feature transformations, are the initial conventional steps. The application of DA techniques to the three major features, including patients' Age, NHP, and patients' LOS, can provide massive great information to use in managerial prescriptive notes providing, LOS classification, and NHP forecasting. Through data analysis, some statistical inferences, such as the Hypothesis tests, might be used to verify some descriptive findings. The results of NHP forecasting and LOS-based NHP data analysis are applied in a mathematical model to forecast the hospital bed capacity. There are several algorithms for LOS classification and NHP forecasting. The following sub-sections opt for a set of algorithms that seems much more appropriate according to the literature.

3.2. LOS classification algorithms

The LOS literature review showed that although there is a long list of research on patients' LOS forecasting, no history exists on the LOS classification [35–38]. But Patients' classification based on various features using BN, K-Nearest Neighbor (KNN), SVM, DT, ANN, AdaBoost (AB), LR, Random Forest (RF), Recency-Frequency-Monetary analysis (RFM), Gradient Boosting (GB), and Ensembles algorithms has been taken great attention, as illustrated in Table 2.

Mentioning Table 2, the five most popular ML algorithms, including BN, KNN, SVM, DT, and LR, are considered for LOS classification.

3.3. NHP forecasting algorithms

Table 3 shows the conventional ML and DL forecasters used in the number of patients forecasting. Considering Table 3, Time series methods are the most convenient forecaster on NHP. In addition, LR and Neural networks were the well-accurate forecasters in most literature. Therefore, SARIMA, SLR, and LSTM neural networks have opted for NHP forecasting.

4. Case study

Rouhani Hospital is a public hospital in Babol City of Mazandaran province located in the northern band of Iran. This hospital has 508 beds, nine wards, and 391 specialists. In this hospital, the Heart ward is the busiest ward possessing 30 beds in the main hall of the Heart ward and 15 beds in the Emergency section of Heart diseases.

4.1. Data set

The collected data set contains 51231 records between 2011 and 2018. Some features, including age and LOS, are considered to be used in the Heart ward's bed capacity forecasting problem. After removing the outliers and noisy data, 47605 records remained which 70% data are categorized as training data and others as test data.

4.2. Data analysis

The NHP in the Heard ward, patients' age distribution, and their LOS play significant roles in the required beds forecasting. In this section, some Data analysis techniques are applied to the NHP, LOS, and Age to provide initial insights on these significant factors affecting the number of required beds equal to 25.

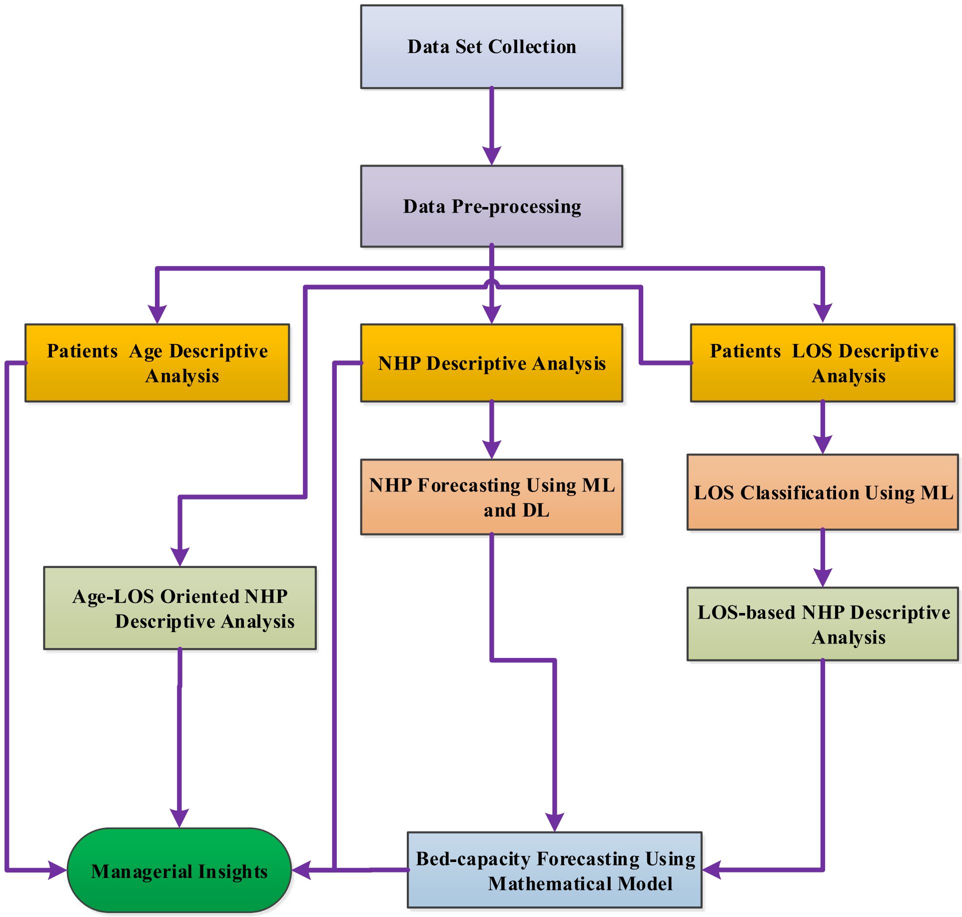

Fig. 1. The schematic representation of the proposed hybrid data-driven HBC forecasting.

Table 2
The popular ML algorithms for patients classification.

Research	Patient classification										
	BN	KNN	SVM	DT	ANN	AB	LR	RF	RFM	GB	Ensemble
Mohammadzadeh et al. (2017) [39]				✓					✓		
Sariyer et al. (2019) [40]				✓	✓		✓	✓			
Nas and Koyuncu (2019)		✓	✓	✓	✓	✓	✓	✓		✓	
Laghmati et al. (2020) [41]		✓	✓	✓		✓					
Fathi et al. (2020) [42]			✓								
Maniruzzaman et al (2020) [43]	✓			✓		✓					
Rahman et al. (2020) [38]				✓							
Badriyah et al. (2020) [44]	✓	✓	✓	✓	✓		✓	✓			
Jiande and Chindo (2021) [45]	✓	✓	✓	✓				✓			
Gurazada et al. (2022) [46]	✓			✓			✓	✓			
Alabbad et al. (2022) [47]								✓		✓	✓
Sultan et al. (2023) [48]		✓	✓	✓	✓						
Amin et al (2023) [49]		✓			✓		✓	✓			✓
This paper	✓	✓	✓	✓			✓				

4.2.1. NHP descriptive analysis

Fig. 2 represents the daily NHP in the Heart ward. In this case study, some Heart patients might be hospitalized temporarily in the Emergency section of the Heart ward or other disease wards, waiting for the Heart ward beds evacuation. The reason is that NHP is over the existing bed capacity of the Heart ward on some days.

In Fig. 2, some drastic fallings occur on NHP around the 19 March–4 April each year. Fig. 3 shows these considerable fluctuations in NHP in more detail from 2013 to 2018. All illustrated fluctuations in Fig. 3 occurred during the annual Nowruz holidays, 19 March–4 April, in Iran as the Persian new year ceremony. In this period, there is no accessibility to Heart specialists because all of them are on their private trips. Therefore, there is no new admittance during this period, and this cause to misuse of some beds and risks for patients. On the other hand, a considerable increase occurs in admittance before and after this holiday, especially from 5 April onwards. Three Hypothesis tests are performed on the mean of hospitalized patients during three 20-day periods, as presented in Table 4, to investigate the significance of falling in admittance during this period.

The sample corresponding to each period is formed by pooling the patients' statistics from 2013 to 2018 to carry out the Hypothesis tests. Table 5 depicts the outputs of the conducted test using the MINITAB 21.1.0.

The probability values from Table 5 show considerable differences between the mean of patients' admittance during BNH and NH, like

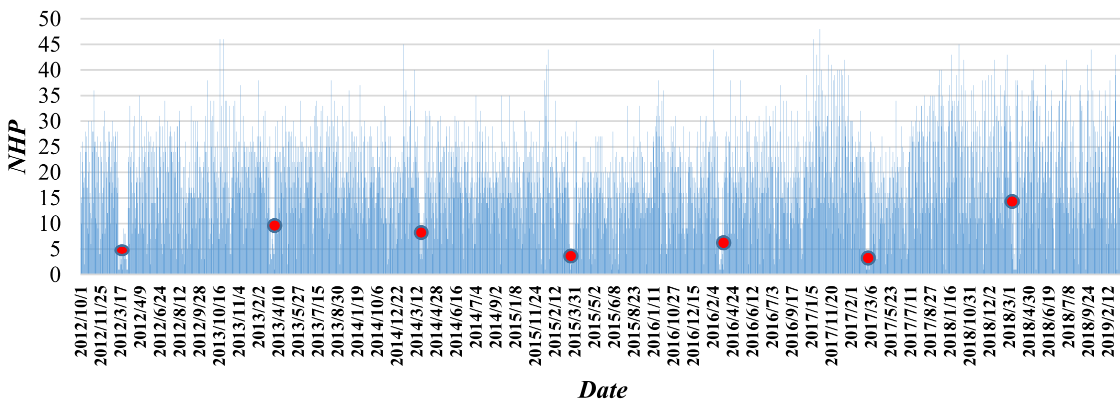

Fig. 2. The historical daily NHP in the Heart ward.

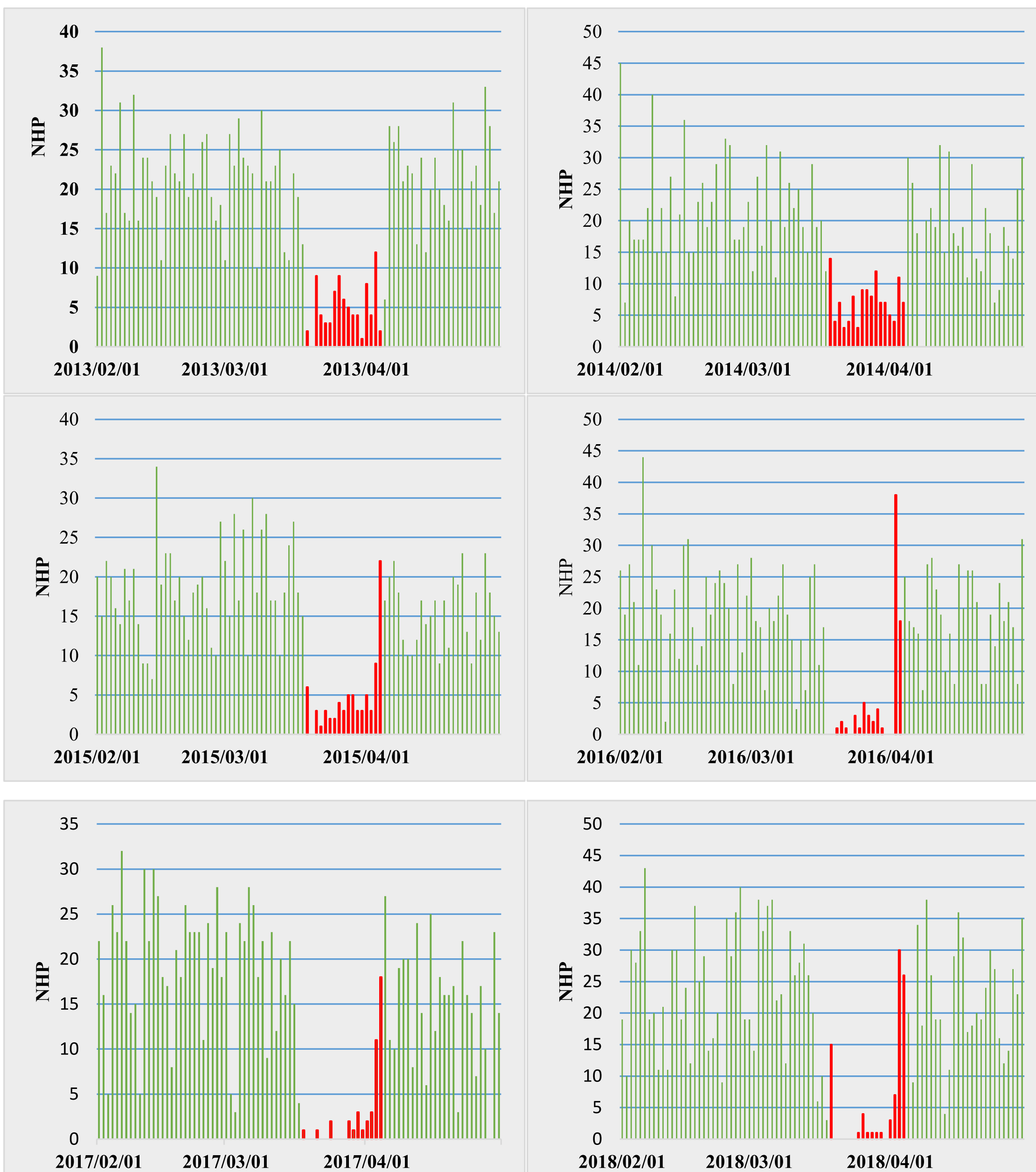

Fig. 3. The details of significant decreases in NHP in March per year.

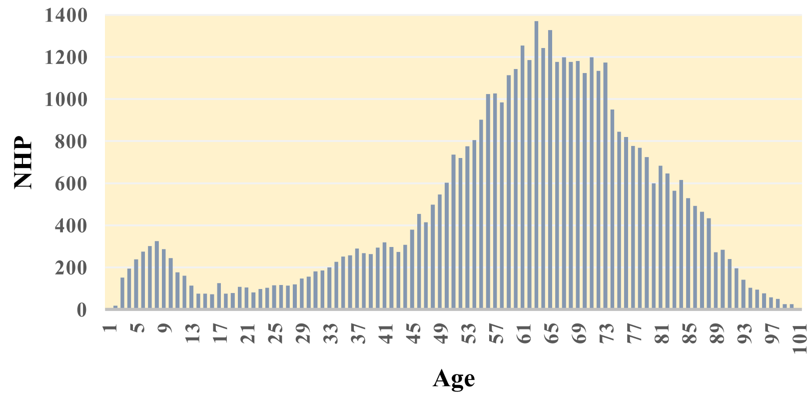

Fig. 4. The age-based distribution of NHP.

Table 3
The popular ML/DL algorithms for the number of patients forecasting.

Research	Forecasters								
	Time series	Neural networks	DT	Regression logistics	GB	LR	SVM	RF	KNN
Bergs et al. (2014) [50]	✓								
Mai et al. (2015) [51]	✓								
Ekström et al. (2015) [52]						✓			
Luo et al. (2017) [53]	✓								
Byron et al. (2018) [54]			✓	✓	✓				
Khalid et al. (2019) [55]	✓	✓							
Sahai et al. (2020) [56]	✓								
Rustam et al. (2020) [57]	✓					✓	✓		
Agrawal et al. (2021) [58]	✓								
Vollmer et al. (2021) [59]	✓				✓	✓		✓	✓
Sudarshan et al. (2021) [60]		✓						✓	
Famiglini et al. (2022) [61]		✓	✓		✓		✓		
Petsis et al. (2022) [62]					✓				
Vieira et al. (2023) [63]	✓								
Alvarez-Chaves et al. (2023) [64]	✓				✓				
This paper	✓	✓				✓			

Table 4
The considered periods for testing the significance of NHP fluctuations.

No	Period	Abbreviation	Time duration	Mean of NHP
1	Before Nowruz holidays	BNH	23 February–14 March	μ_{BNH}
2	Nowruz holidays	NH	15 March–4 April	μ_{NH}
3	Post Nowruz holidays	PNH	5 April–23 April	μ_{PNH}

Table 5
The results of the hypothesis tests.

No	Hypothesis test	Sample size	P-Value
1	$H_0 : \mu_{BNH} = \mu_{NH}$ $H_1 : \mu_{BNH} \neq \mu_{NH}$	120	0
2	$H_0 : \mu_{PNH} = \mu_{NH}$ $H_1 : \mu_{PNH} \neq \mu_{NH}$	120	0
3	$H_0 : \mu_{BNH} = \mu_{PNH}$ $H_1 : \mu_{BNH} \neq \mu_{PNH}$	120	0.175

Table 6
The proportion of NHP in each age class.

No of Class	1	2	3	4	5	6
Age Range	0–15	15–30	30–45	45–60	60–85	85–99
NHP	2641	1617	3994	11745	24266	3342
Proportion	0/06	0/03	0/08	0/25	0/51	0/07
Cumulative	0/06	0/09	0/17	0/42	0/93	1

PNH and NH. Moreover, the p-value=0.175 shows no significant difference between the mean of patients' admittance along BNH and PNH according to corresponding collected data and conventional significance level 0.05. These results verify the conducted data-driven visual inferences from Fig. 2.

4.2.2. Age-based NHP descriptive analysis

Fig. 4 represents that the NHP follows a bimodal pattern according to the patient's age. In addition, the most significant portion of NHP belongs to patients aged between 45 and 85.

Using information represented in Fig. 4, six classes of ages are considered, including (0,15], (15,30], (30,45], (45,60], (60,85], and (85,99], as depicted in Table 6.

Although Table 6 shows that children's class only involves six percent of total NHP, its share is considerable enough to provide specialized proper room in the Heart ward or even a particular children's Heart ward regarding their special emotional considerations. Another interesting note in Fig. 4 is that the mean age of children hospitalized with Heart issues is about 7. By conducting some diagnostic analyses on the modal nature of children's NHP, the roots of this manner would be investigated and might be beneficial in children's Heart disease treatment management and even prevention in the future. Fig. 5 illustrates

Table 7
The share of each LOS in terms of NHP.

LOS	NHP	Proportion	Cumulative	LOS	NHP	Proportion	Cumulative
1	22787	0.478668	*0.478668*	11	539	0.011322	*0.901754*
2	6589	0.13841	*0.617078*	12	462	0.009705	0.911459
3	4187	0.087953	*0.705031*	13	400	0.008402	0.919861
4	2526	0.053062	0.758093	14	397	0.008339	0.928201
5	1762	0.037013	0.795106	15	282	0.005924	0.934125
6	1319	0.027707	*0.822813*	16	232	0.004873	0.938998
7	1109	0.023296	0.846109	17	214	0.004495	0.943493
8	805	0.01691	0.863019	18	219	0.0046	0.948094
9	707	0.014851	0.87787	19	194	0.004075	0.952169
10	598	0.012562	0.890432	20	140	0.002941	0.95511

Table 8
The accuracy of LOS classifiers.

Classifier	Number of classes				
	4	5	6	7	8
SVM	63%	55%	**71%**	48%	51%
DT	64%	49%	69%	46%	49%
LR	52%	43%	48%	40%	42%
KNN	46%	43%	47%	42%	42%
BN	40%	36%	39%	35%	40%

Table 9
The proportion of NHP per LOS classes.

No	LOS range	NHP	Proportion	Cumulative
1	1	22787	0/48	0/48
2	2	6589	0/14	0/62
3	3	4187	0/09	0/71
4	[4,6]	5607	0/12	0/83
5	[7,15]	5299	0/11	0/94
6	[16,730]	3136	0/06	1

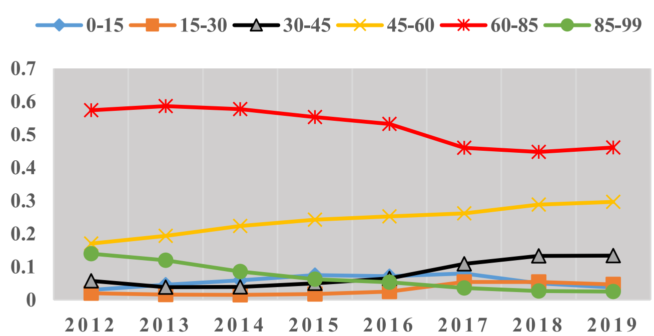

Fig. 5. The NHP-share trend per Age class.

the fact that the rate of admittance is increasing only among people aged between 30 and 60.

Fig. 5 illustrates that the admittance of people of age class (30,45] has taken more and more acceleration. In other words, Heart diseases are becoming more common over the past between younger individuals. This fact allocates more credit to promoting preventive actions on Heart disease in the public.

4.2.3. LOS descriptive analysis

Fig. 6 represents the patients' LOS distribution.

In Fig. 6, only a few LOSs are too long, and most seem to be less than ten days. This claim is verified in Figs. 7 and 8, representing the scatter of NHP in terms of LOS.

Table 7 represents the share of the first twenty values of LOS in terms of NHP to investigate the frequency of LOS values in detail.

Table 7 shows that LOS=1 makes the role of the first to fourth deciles lonely as the first quartile, LOS=2 is the fifth and sixth deciles as the second quartile, LOS=4 is the third quartile, LOS=3 is the seventh decile, LOS=6 is the eighth decile, and LOS=11 is the ninth decile. Also, more than 95.5 percent of patients stay less than 20 days. This information helps to form appropriate classes of LOS.

4.3. LOS classification

Several sets of LOS classes are made, and their accuracy is examined using several classification tools coded using the Sklearn library of Python, as presented in Table 8.

In Table 8, the classification results show that SVM provides the best accuracy and introduces the case of six classes of LOS as the best alternative. Table 9 depicts the information on NHP and the time interval of LOS classes.

Table 9 illustrates only 6% of patients experienced LOS for more than 15 days, and most would stay in the Heart ward for only one day. In addition, Table 10 presents detailed information on the NHP distribution in each Age class per LOS class.

In Table 10, the lowest proportion of NHP from the first class of LOS, LOS=1, comes from the eldest people's class which seems reasonable because of the more Heart risk and need for continual health monitoring in this class. But the following place is allocated to children means the admittance of this type of Heart patient tends to be longer over other Age classes. A more diagnostic analysis could provide helpful information to make better decisions on the specialized Heart beds for children because the children's class possesses the first rank in the proportion of NHP allocated to the long LOS class. This situation assigns crucial importance to establishing specialized children's Heart wards in the future.

4.4. NHP forecasting

Table 11 summarizes the results of applying three opted forecasters in Section 3.3 using Python libraries.

As illustrated in Table 11, LR opted for forecasting NHP in the future due to providing the least error index. The mentioned hospital was the main center for COVID-19 patients' treatment in Babol City between 2019 and 2021, and hence, the daily NHP is forecasted for the five years from 2022 to 2026.

4.5. Heart ward bed capacity forecasting

According to Fig. 1, the forecasted NHP and LOS-based distribution of NHP should be applied to forecast the Heart ward bed capacity in the future. The proportion of NHP in LOS classes, from Table 9, is applied to forecast the NHP belonging to each LOS class in the future. It is supposed that the expected nights/beds corresponding to LOS classes are equal to 1, 2, 3, 5, 11, and 22, respectively. Eq. (1) uses the forecasted NHP of LOS classes to predict the required daily beds

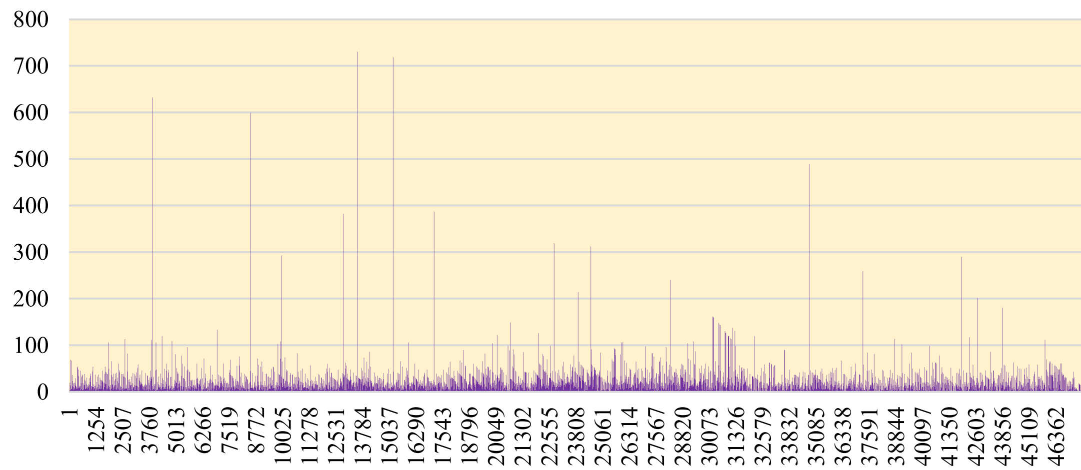

Fig. 6. The LOS distribution of Heart patients.

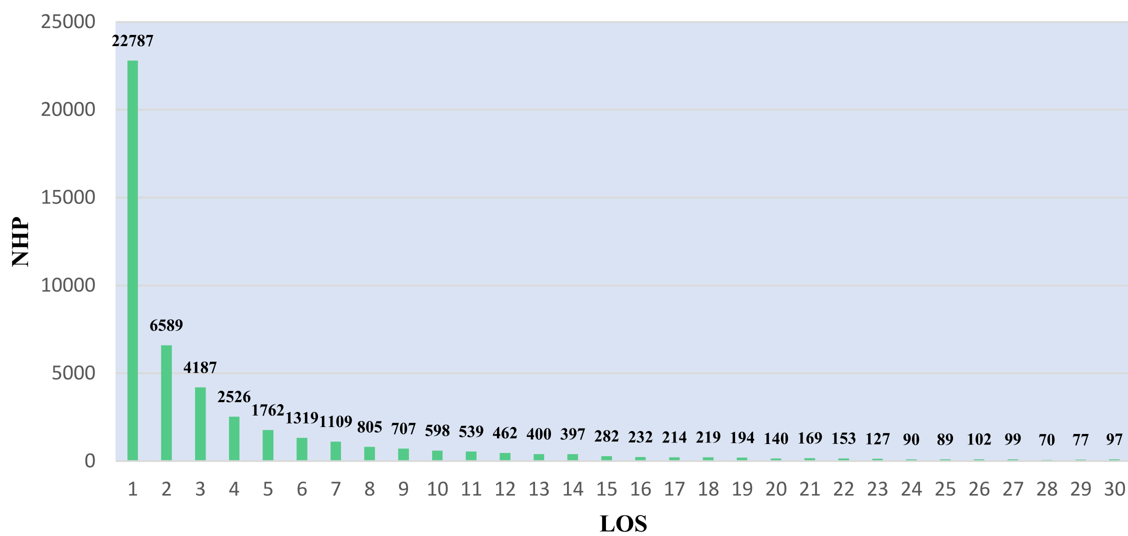

Fig. 7. The NHP with a LOS of less than 31 days.

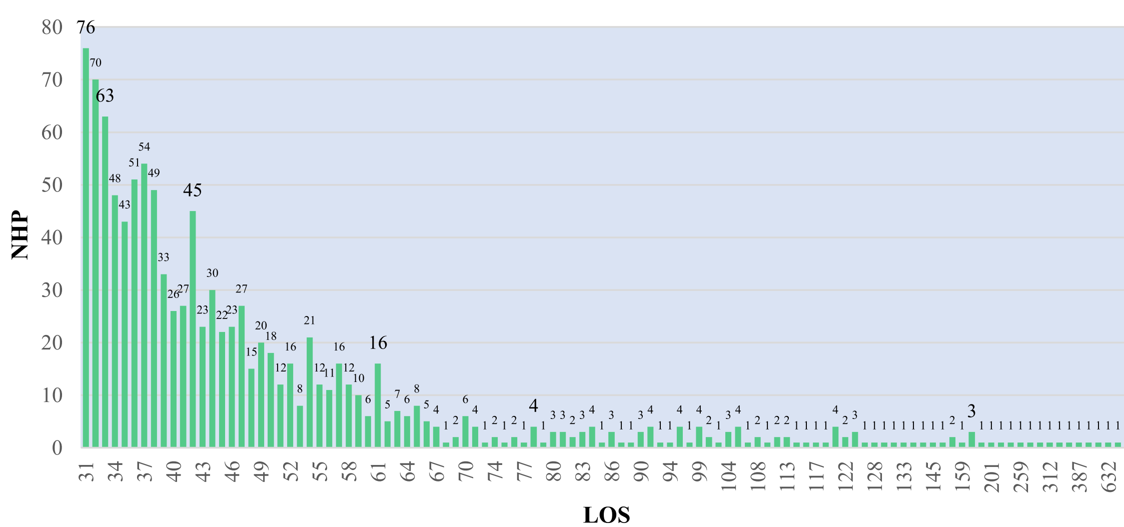

Fig. 8. The NHP with LOS greater than 30 days.

Table 10
The NHP partitioning according to Age and LOS classes.

LOS range	Age range					
	(0,15]	(15,30]	(30,45]	(45,60]	(60,85]	(85,99]
1	891	765	2082	6841	11365	843
2	498	314	667	1551	3185	374
3	462	161	288	736	2134	406
[4,6]	209	133	358	1072	2963	872
[7,15]	302	122	228	1097	3101	449
[16,730]	223	38	145	592	1895	243

Table 11
The results of NHP forecasting.

Forecasting algorithm	Python library	Fitted model specifications	RMSE
SARIMA	Statsmodels	$P = 0,\ D = 1,\ Q = 1$	9.5
LR	Sklearn	Slope=-0.0027 , Intercept=16.71	**8.6**
LSTM Neural Network	Keras	Three hidden layers with 300, 150, and 100 neurons, respectively, and epoch=1000	9.1

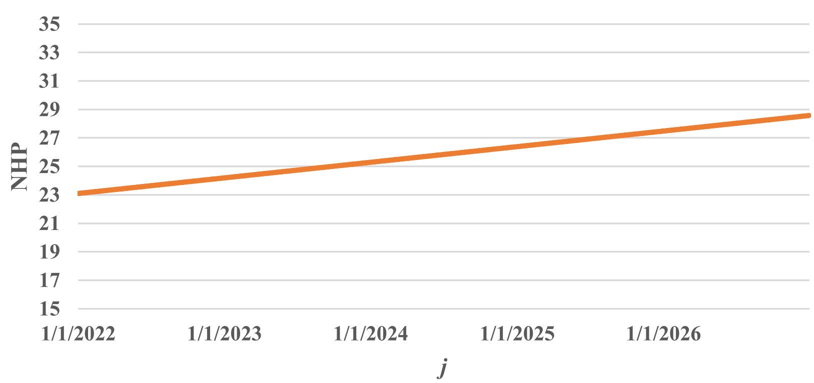

Fig. 9. The forecasted daily NHP between 2022 and 2027.

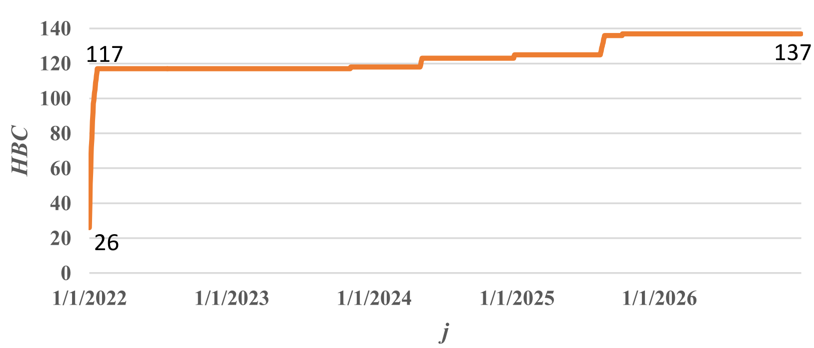

Fig. 10. The forecasted HBC in Heart ward.

between 2022 and 2026 (see Fig. 9).

$$HBC_j = \sum_i NHP_{ij} + \sum_{i\geq 2} NHP_{i,j-1} + \sum_{i\geq 3} NHP_{i,j-2} + \sum_{i\geq 4}\sum_{k=j-4}^{j-3} NHP_{ik} + \sum_{i\geq 5}\sum_{k=j-10}^{j-5} NHP_{ik} + \sum_{k=j-20}^{j-11} NHP_{6,k} \quad i = 1,2,\ldots,6; \quad j = 1,2,\ldots,1825 \tag{1}$$

In Eq. (1), i is the index of LOS classes, and j denotes the order of days between 2022 and 2027. Also, NHP_{ij} represents the NHP belonging to the ith class of LOS that will be admitted in the Heart ward during day j, and HBC_j denotes the needed beds in day j. Fig. 10 illustrates the forecasted HBC in Heart ward from 1 January 2022 to 31 December 2026.

Fig. 10 represents a drastic increase in the number of required beds because the older hospitalized patients are not considered and the calculation only is dependent on the forthcoming NHPs. In this Figure, the average needed bed capacity in the future is forecasted at 120 beds which is much more than the existing 45 beds. Therefore, this hospital's administration should make strategic decisions to increase the bed capacity of the Heart ward from 45 beds to 137 beds by 2026.

4.6. Managerial insights

In section 4.3.1, the NHP descriptive analysis showed a type of mismanagement in the use of the existing bed capacity in the Heart ward during about 20 days per year related to the Persian new year holidays. Therefore, some managerial tasks and cultural efforts are needed to provide an equitable shift working for Heart specialists during the new year holidays avoiding empty beds. This could result in sustainable Heart bed capacity in terms of productivity and reliable treatment for high risks Heart patients. In addition, Tables 6 and 10 showed that children are a considerable ratio of Heart patients, and a noticeable proportion of them require too long LOS. Therefore, establishing an appropriate specialized Heart room or Heart ward for children must be a significant goal of the hospital's authorities. Also, Fig. 10 represented the need for 117 beds in the Heart ward from 2022 upward. Providing proper facilities to establish per specialized Heart bed requires a considerable budget. Hence, conducting financial analysis to estimate the needed resources such as money is inevitable, particularly in a developing country such as Iran with tremendous limits on monitorial resources.

5. Conclusions, limitations, and recommendations

This paper considered the hospital bed capacity forecasting problem and proposed a Data-driven methodology to avoid the limits and difficulties of using traditional programming or simulation models. A hybrid approach involving DA, ML, DL, and statistical inference was applied in a Heart ward bed capacity forecasting. Results showed the need for establishing more beds in the Heart ward from 26 to 137 in 2027. In addition, more considerations on the child Heart patients' treatment and the increasing trend of the occurring Heart diseases among younger people are recommended by results.

This paper only focused on HBC forecasting and paid no attention to other required resources such as Specialists, Nurses, Equipment, and Budget. Forecasting these significant factors could provide beneficial information for decision-makers. In addition, some other factors can affect the NHP of a hospital in the future, such as insurance service level, the accessibility to other hospitals, the service quality level, limits on the number of available corresponding specialists, the number of specialized equipment, the economic conditions, and the people's accessibility to periodic checkup. This needs to use multivariate time series data sets and multivariate ML tools in future works. Also, uncertainty is an inevitable fact in decision-making that could be mentioned in the future use of DA, ML, and DL for HBC forecasting.

Ethical statement

The Authors declare that no real clinical data is applied in this paper.

Data and code availability statement

The Authors would submit the related data set and codes whenever a request is received.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] R.P. Jones, A model to compare international hospital bed numbers, including a case study on the role of indigenous people on acute occupied bed demand in australian states, Int. J. Environ. Res. Public Health 19 (18) (2022) 11239, http://dx.doi.org/10.3390/ijerph191811239.

[2] H. Ravaghi, S. Alidoost, et al., Models and methods for determining the optimal number of beds in hospitals and regions: a systematic scoping review, BMC Health Serv. Res. 20 (2020) 1–13, http://dx.doi.org/10.1787/health_glance-2017-en.

[3] J.O. Ferstad, A. Gu, et al., A model to forecast regional demand for COVID-19 related hospital beds, 2020, http://dx.doi.org/10.1101/2020.03.26.20044842, med-archive.

[4] R.P. Jones, How many medical beds does a country need?, Br. J. Health Care Manag. 26 (9) 248–259, http://dx.doi.org/10.12968/bjhc.2020.0028.

[5] C. Keegan, A. Brick, et al., How many beds? Capacity implications of hospital care demand projections in the irish hospital system, 2015–2030, Br. J. Health Care Manag. 34 (1) (2019) e569–e582, http://dx.doi.org/10.1002/hpm.2673.

[6] M. McKee, Reducing Hospital Beds: What are the Lessons To Be Learned?, W. H. Organization, 2004, https://apps.who.int/iris/handle/10665/107615.

[7] L.V. Green, How many hospital beds? Inq. J. Health Care Org. Pro. Fin. 39 (4) (2002) 400–412, http://dx.doi.org/10.5034/inquiryjrnl_39.4.400.

[8] T. Latruwe, M. Van der Wee, et al., A long-term forecasting and simulation model for strategic planning of hospital bed capacity, Oper. Res. Health Care. 36 (2023) 100375, http://dx.doi.org/10.1016/j.orhc.2022.100375.

[9] M. Ordu, E. Demir, et al., A novel healthcare resource allocation decision support tool: A forecasting-simulation–optimization approach, J. Oper. Res. Soc. 72 (3) (2021) 485–500, http://dx.doi.org/10.1080/01605682.2019.1700186.

[10] M. Deschepper, K. Eeckloo, et al., Prediction of hospital bed capacity during the COVID- 19 pandemic, BMC Health Serv. Res. 21 (1) (2021) 1–10, http://dx.doi.org/10.1186/s12913-021-06492-3.

[11] F. Ben Abdelaziz, M. Masmoudi, A multiobjective stochastic program for hospital bed planning, J. Oper. Res. Soc. 63 (4) (2012) 530–538, http://dx.doi.org/10.1057/jors.2011.39.

[12] S.M.C. Loureiro, J. Guerreiro, I. Tussyadiah, Artificial intelligence in business: State of the art and future research agenda, J. Bus. Res. 129 (2021) 911–926, http://dx.doi.org/10.1016/j.jbusres.2020.11.001.

[13] R. Akerkar, Artificial Intelligence for Business, Springer, 2019.

[14] W.R. Paczkowski, Business Analytics: Data Science for Business Problems, Springer Nature, 2022.

[15] A. Nguyen, S. Lamouri, et al., Data analytics in pharmaceutical supply chains: state of the art, opportunities, and challenges, Int. J. Prod. Res. 60 (22) (2022) 6888–6907, http://dx.doi.org/10.1080/00207543.2021.1950937.

[16] L.P. Provost, S.K. Murray, The Health Care Data Guide: Learning from Data for Improvement, John Wiley & Sons, 2022.

[17] S.U.D. Wani, N.A. Khan, et al., Utilization of artificial intelligence in disease prevention: Diagnosis, treatment, and implications for the healthcare workforce, Healthcare 10 (4) (2022) 608, http://dx.doi.org/10.3390/healthcare10040608.

[18] K. Santosh, L. Gaur, Artificial Intelligence and Machine Learning in Public Healthcare: Opportunities and Societal Impact, Springer Nature, 2022.

[19] M. Tello, E.S. Reich, et al., Machine learning based forecast for the prediction of inpatient bed demand, BMC Medical Inform. Decis. Mak. 22 (1) (2022) 1–13, http://dx.doi.org/10.1186/s12911-022-01787-9.

[20] J. Schiele, T. Koperna, J.O. Brunner, Predicting intensive care unit bed occupancy for integrated operating room scheduling via neural networks, Nav. Res. Logist. 68 (1).

[21] 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, http://dx.doi.org/10.1016/j.eswa.2017.02.023.

[22] R.B. Bachouch, A. Guinet, S. Hajri-Gabouj, An integer linear model for hospital bed planning, Int. J. Prod. Econ. 140 (2) (2012) 833–843, http://dx.doi.org/10.1016/j.ijpe.2012.07.023.

[23] P. Devapriya, C.T.B. Strömblad, et al., Stratbam: a discrete-event simulation model to support strategic hospital bed capacity decisions, J. Med. Syst. 39 (10) (2015) 1–13, http://dx.doi.org/10.1007/s10916-015-0325-0.

[24] S. Nas, M. Koyuncu, Emergency department capacity planning: a recurrent neural network and simulation approach, Comput. Math. Methods Med. (2019) http://dx.doi.org/10.1155/2019/4359719.

[25] E. Kutafina, I. Bechtold, et al., Recursive neural networks in hospital bed occupancy forecasting, BMC Medical Inform. Decis. Mak. 19 (1) (2019) 1–10, http://dx.doi.org/10.1186/s12911-019-0776-1.

[26] T. Zhu, P. Liao, et al., Data-driven models for capacity allocation of inpatient beds in a Chinese public hospital, Comput. Math. Methods Med. (2020) http://dx.doi.org/10.1155/2020/8740457.

[27] M. Haghshenas, A. Nemati, E. Asadi-Gangraj, A bi-objective model for cancer hospitals' location and cancer patients' allocation in, Iran, Int. J. Hos. Res 10 (4) (2021) ijhr.iums.ac.ir..

[28] M. Haghshenas, A. Nemati, E. Asadi-Gangraj, Cancer-curing supply chain planning with regard to hospital bed-capacity efficiency: a plan for Iran in 2040, Int. J. Serv. Oper. Manag. 45 (2) (2023) 170–186, http://dx.doi.org/10.1504/IJSOM.2023.131495.

[29] S.B. Kabir, S.S. Shuvo, H.U. Ahmed, Use of machine learning for long term planning and cost minimization in healthcare management, 2021, http://dx.doi.org/10.1101/2021.10.06.21264654, med-archive.

[30] D. Garcia-Vicuña, A. López-Cheda, et al., Estimation of patient flow in hospitals using up-to-date data. Application to bed demand prediction during pandemic waves, PLoS One 18 (2) (2023) e0282331, http://dx.doi.org/10.1371/journal.pone.0282331.

[31] E. Redondo, V. Nicoletta, et al., A simulation model for predicting hospital occupancy for Covid-19 using archetype analysis, HealthCare Anal. 3 (2023) 100197, http://dx.doi.org/10.1016/j.health.2023.100197.

[32] R. Bekker, M. uit het Broek, G. Koole, Modeling COVID-19 hospital admissions and occupancy in the netherlands, European J. Oper. Res. 304 (1) (2023) 207–218, http://dx.doi.org/10.1016/j.ejor.2021.12.044.

[33] R. Widyasari, H. Cipta, M. Nasution, An optimization model for hospital emergency room based on patient growth and capacity management, Sinkron: J. Dan Penel. Tekn. Inf. 8 (1) (2023) 188–194, http://dx.doi.org/10.33395/sinkron.v8i1.11890.

[34] M.R. Johnson, H. Naik, et al., Forecasting ward-level bed requirements to aid pandemic resource planning: Lessons learned and future directions, Health Care Manag. Sci. (2023) 1–24, http://dx.doi.org/10.1007/s10729-023-09639-2.

[35] M.M. Rahman, D. Kundu, et al., Hospital patients' length of stay prediction: A federated learning approach, J. King Saud Univ. Comput. Inf. Sci. 34 (10) (2022) 7874–7884, http://dx.doi.org/10.1007/s10729-023-09639-2.

[36] W.K.D. Jayamini, F. Mirza, et al., State of asthma-related hospital admissions in New Zealand and predicting length of stay using machine learning, Appl. Sci. 12 (19) (2022) 9890, http://dx.doi.org/10.3390/app12199890.

[37] J. Ebinger, M. Wells, et al., A machine learning algorithm predicts duration of hospitalization in COVID-19 patients, Intell. Based. Med. 5 (2021) 100035, http://dx.doi.org/10.1016/j.ibmed.2021.100035.

[38] M.A. Rahman, B. Honan, et al., Using data mining to predict emergency department length of stay greater than 4 hours: Derivation and single-site validation of a decision tree algorithm, Emerg. Med. Australas. 32 (3) (2020) 416–421, http://dx.doi.org/10.1111/1742-6723.13421.

[39] M. Mohammadzadeh, Z.Z. Hoseini, H. Derafshi, A data mining approach for modeling churn behavior via RFM model in specialized clinics case study: A public sector hospital in tehran, Procedia Comput. Sci. 120 (2017) 23–30, http://dx.doi.org/10.1016/j.procs.2017.11.206.

[40] G. Sariyer, C.Ö. Taşar, G.E. Cepe, Use of data mining techniques to classify length of stay of emergency department patients, Bio-Algorithms Med-Syst. 15 (1) (2019) 20180044, http://dx.doi.org/10.1515/bams-2018-0044.

[41] S. Laghmati, B. Cherradi, et al., Classification of patients with breast cancer using neighbourhood component analysis and supervised machine learning techniques, in: Int. Con. Com. Tech. (CommNet), 2020, pp. 1–6, http://dx.doi.org/10.1109/CommNet49926.2020.9199633.

[42] M. Fathi, M. Nemati, et al., A machine learning approach based on SVM for classification of liver diseases, Biomed. Eng. - Appl. 32 (3) (2020) 2050018, http://dx.doi.org/10.4015/S1016237220500180.

[43] M. Maniruzzaman, M. Rahman, et al., Classification and prediction of diabetes disease using machine learning paradigm, Health Inf. Sci. Syst. 8 (1) (2020) 1–14, http://dx.doi.org/10.1007/s13755-019-0095-z.

[44] T. Badriyah, N. Sakinah, et al., Machine learning algorithm for stroke disease classification, Int. Con. Elec. Comn. Com. Eng. (ICECCE) (2020) 1–5, http://dx.doi.org/10.1109/ICECCE49384.2020.9179307.

[45] J. Wu, C. Hicks, Breast cancer type classification using machine learning, J. Pers. Med. 11 (2) (2021) 61, http://dx.doi.org/10.3390/s22134968.

[46] S.G. Gurazada, S. Gao, et al., Predicting patient length of stay in Australian emergency departments using data mining, Sensors 22 (13) (2022) 4968.

[47] D.A. Alabbad, A.M. Almuhaideb, et al., Machine learning model for predicting the length of stay in the intensive care unit for Covid-19 patients in the eastern province of Saudi Arabia, Inform. Med. Unlocked. 30 (2022) 100937, http://dx.doi.org/10.1016/j.imu.2022.100937.

[48] A. Sultana, F. Ahmed, M.S. Alam, A systematic review on surface electromyography-based classification system for identifying hand and finger movements, Healthcare Anal. (2022) 100126, http://dx.doi.org/10.1016/j.health.2022.100126.

[49] R. Amin, R. Yasmin, et al., Prediction of chronic liver disease patients using integrated projection based statistical feature extraction with machine learning algorithms, Inform. Med. Unlocked 36 (2023) 101155, http://dx.doi.org/10.1016/j.imu.2022.101155.

[50] J. Bergs, P. Heerinckx, S. Verelst, Knowing.what.to. expect, Forecasting monthly emergency department visits: A time-series analysis, Int. Emerg. Nurs. 22 (2) (2014) 112–115, http://dx.doi.org/10.1016/j.ienj.2013.08.001.

[51] Q. Mai, P. Aboagye-Sarfo, et al., Predicting the number of emergency department presentations in western Australia: A population-based time series analysis, Emerg. Med. Australas. 27 (1) (2015) 16–21, http://dx.doi.org/10.1111/1742-6723.12344.

[52] A. Ekström, L. Kurland, et al., Forecasting emergency department visits using internet data, Ann. Emerg. Med. 65 (4) (2015) 436–442, http://dx.doi.org/10.1016/j.annemergmed.2014.10.008.

[53] L. Luo, L. Luo, et al., Hospital daily outpatient visits forecasting using a combinatorial model based on ARIMA and SES models, BMC Health. Serv. Res. 17 (1) (2017) 1–13, http://dx.doi.org/10.1186/s12913-017-2407-9.

[54] B. Graham, R. Bond, et al., Using data mining to predict hospital admissions from the emergency department, IEEE Access 6 (2018) 10458–10469, http://dx.doi.org/10.1109/ACCESS.2018.2808843.

[55] R. Khaldi, A. El Afia, R. Chiheb, Forecasting of weekly patient visits to emergency department: real case study, Proc. comput. sci. 148 (2019) 532–541, http://dx.doi.org/10.1016/j.procs.2019.01.026.

[56] A.K. Sahai, N. Rath, et al., ARIMA modelling and forecasting of COVID-19 in top five affected countries, Diabetes Metab. Syndr. 14 (5) 1419–1427, http://dx.doi.org/10.1016/j.dsx.2020.07.042.

[57] F. Rustam, A.A. Reshi, et al., COVID-19 future forecasting using supervised machine learning models, IEEE Access 8 (2020) 101489-101499, http://dx.doi.org/10.1109/ACCESS.2020.2997311.

[58] P. Agrawal, V. Madaan, et al., FOCOMO: Forecasting and monitoring the worldwide spread of COVID-19 using machine learning methods, J. Interdiscip. Math. 24 (2) (2021) 443–466, http://dx.doi.org/10.1080/09720502.2021.1885812.

[59] M.A.C. Vollmer, B. Glampson, et al., A unified machine learning approach to time series forecasting applied to demand at emergency departments, BMC Emerg. Med. 21 (1) (2021) 1–14, http://dx.doi.org/10.1186/s12873-020-00395-y.

[60] V.K. Sudarshan, M. Brabrand, et al., Performance evaluation of emergency department patient arrivals forecasting models by including meteorological and calendar information: A comparative study, Comput. Biol. Med. 135 (2021) 104541, http://dx.doi.org/10.1016/j.compbiomed.2021.104541.

[61] L. Famiglini, A. Campagner, et al., A robust and parsimonious machine learning method to predict ICU admission of COVID-19 patients, Med. Biol. Eng. Comput. (2022) 1–13, http://dx.doi.org/10.1016/j.compbiomed.2021.104541.

[62] S. Petsis, A. Karamanou, et al., Forecasting and explaining emergency department visits in a public hospital, J. Intell. Inf. Syst. (2022) 1–22, http://dx.doi.org/10.1007/s10844-022-00716-6.

[63] A. Vieira, I. Sousa, S. Dória-Nóbrega, Forecasting daily admissions to an emergency department considering single and multiple seasonal patterns, Healthcare Anal. 3 (2023) 100146, http://dx.doi.org/10.1007/s10844-022-00716-6.

[64] H. Álvarez-Chaves, P. Muñoz, Machine learning methods for predicting the admissions and hospitalisations in the emergency department of a civil and military hospital, J. Intell. Inf. Syst. 1–20, http://dx.doi.org/10.1007/s10844-023-00790-4.