공학석사 학위논문

병원 프로세스 개선을 위한 전자의무기록 기반 인공지능 기술 개발 및 활용

Development and utilization of artificial intelligence technology based on electronic medical records for improving hospital processes

울산대학교 대학원
의 과 학 과
서 혜 람

Development and utilization of artificial intelligence technology based on electronic medical records for improving hospital processes

지도교수 김 영 학

이 논문을 공학석사학위 논문으로 제출함

2024 년　　2 월

울산대학교 대 학 원
의 과 학 과
서 혜 람

서혜람의 공학석사학위 논문을 인준함

심사위원 이 계 화 (인)

심사위원 김 영 학 (인)

심사위원 전 태 준 (인)

울 산 대 학 교 대 학 원

2024 년 2 월

Abstract

Background and objective: As the global healthcare market expands and medical standards improve, healthcare costs are increasing. At the same time, the financial burden of healthcare costs on both nations with limited budgets and individuals is growing. To address this issue, various approaches to reducing and controlling healthcare costs have been proposed. Among these approaches, hospital process improvement is considered a significant method for providing important benefits to patients and treatment services. In particular, process optimization is regarded as an effective way to reduce inefficiencies for both hospitals and patients. Recent research has been actively applying artificial intelligence (AI) to the field of medicine, and efforts to overcome the constraints of medical resources are increasing. Among various methods, there is particular attention to predicting emergency department (ED) overcrowding and hospital bed occupancy rates (BORs). ED overcrowding can lead to issues such as increased mortality rates, longer wait times, treatment errors, diagnostic and procedural delays, and more. High BORs can also negatively affect the health of healthcare staff and increase the risk of infections. In this study, we developed AI models that utilize electronic medical records (EMR) to improve hospital processes.

Method: In the first chapter, we focus on creating models to predict the likelihood of admission within 24 hours for patients passing through the ED and forecasting expected wait times. This model was developed to support quick decision-making by ED physicians and has shown outstanding performance in predicting the likelihood of admission within 24 hours and wait times. Furthermore, by leveraging unstructured text data, we enhanced the model's performance and proved the importance of unstructured text in ED notes using explainable artificial intelligence (XAI) to confirm variable influences.

In the second chapter, we conducted research on predicting the BORs of individual wards and rooms. We combined time-series data related to bed occupancy recorded at hourly intervals with static room data to create various datasets. Using these datasets, we developed two models for predicting ward BORs and four models for predicting room BORs. These models demonstrated high performance, with the model that combined dynamic and static data and predicted BORs at weekly intervals performing the best. This emphasized the importance of static data.

Results: In chapter 1, among several evaluated models, the extreme gradient boosting model (XGB) that incorporated text data yielded the best performance. This model achieved an area under the receiver operating characteristic curve (AUROC) score of 0.922 and an area under the precision-recall curve (AUPRC) score of 0.687. The mean absolute error (MAE) revealed a difference of approximately 3 hours. Through XAI, we identified important variables affecting this classification and found that unstructured text data variables mainly had a large impact.

In chapter 2, the ward-level prediction model with an MAE of 0.057, a mean squared error (MSE) of 0.007, a root mean squared error (RMSE) of 0.082, and an R2 score of 0.582. Among the room-level prediction models, the model that combined static data exhibited superior performance with an MAE of 0.123, an MSE of 0.051, an RMSE of 0.226, and an R2 score of 0.320. Model results can be displayed on an electronic dashboard for easy access via the web.

Conclusions: Research aimed at improving hospital processes must produce practical results that can be used in healthcare institutions. Therefore, we have proposed a virtual web application that is practically applicable to significantly enhance the economic efficiency of hospital and ED operations. Applying AI models to hospital processes can simplify procedures and efficiently utilize limited medical resources. This not only enhances medical services but also offers the potential for cost savings.

Key words: EMR, Emergency department, Artificial intelligence, Natural language processing, Time series forecasting, Combining static and dynamic variables

Abbreviation

AI	artificial intelligence
AMC	Asan Medical Center
AUC	area under the curve
AUPRC	area under the precision-recall curve
AUROC	area under the receiver operating characteristic
BiLSTM	bidirectional long short-term memory
BORs	bed occupancy rates
DF	document frequency
DL	deep learning
DNN	deep neural network
DTM	document-term matrix
ED	emergency department
EHR	electronic health record
EMR	electronic medical record
FPR	false positive rate
GBM	gradient boosting machine
IRB	institutional review board
IUC	intensive care unit
KTAS	Korea Triage and Acuity Scale
LR	logistic regression
LSTM	long short-term memory
MAE	mean absolute error
MAPE	mean absolute percentage error
ML	machine learning
MLP	multi-layer perceptron
MSE	mean squared error
NARX	nonlinear autoregressive exogenous
NB	naïve bayes
NEDIS	National Emergency Department Information System
NLP	natural language processing
OECD	Organisation for Economic Co-operation and Development
PCR	polymerase chain reaction
R3D	room 3 days
R7D	room 7 days
RBORs	room bed occupancy rates
RF	random forest

RMSE	root mean squared error
RNN	recurrent neural network
ROC	receiver operating characteristic
RS3D	room static 3 days
RS7D	room static 7 days
SHAP	SHapley Additive exPlanations
TF-IDF	term frequency-inverse document frequency
TPR	true positive rate
W30D	ward 30 days
W7D	ward 7 days
WBORs	ward bed occupancy rates
XAI	eXplainable Artificial Intelligence
XGB	extreme gradient boosting

Contents

Introduction

The global healthcare market has been expanding, leading to improvements in medical standards. However, this progress has also resulted in an increase in healthcare costs. Simultaneously, both governments and individuals have limited budgets for these expenses, leading to a growing burden of healthcare costs. To address this issue, various approaches have been proposed to reduce and control healthcare cost wastage [1, 2]. Among these approaches, the improvement of hospital processes stands out as one of the key methods that provide significant benefits to patients by delivering essential services and treatments. Simplifying processes, in particular, is considered an effective way to reduce inefficiencies and waste for both hospitals and patients [3].

Recent research has seen a surge in the application of artificial intelligence (AI) in the field of healthcare, with a growing effort to overcome the limitations of medical resources. Among various approaches, methods for alleviating emergency department (ED) overcrowding and predicting hospital bed occupancy rates (BORs) have garnered particular attention [4-20]. ED overcrowding can lead to issues such as increased mortality rates, longer wait times, treatment errors, diagnostic delays, and more [21-23]. Additionally, high BORs can have adverse effects on the health of medical staff and increase the risk of infections [24-25]. The application of AI models to hospital processes can streamline these processes, enabling medical professionals and hospital administrators to efficiently utilize limited medical resources. This, in turn, holds the promise of improving healthcare services and reducing costs.

Studies aimed at improving hospital processes should yield practical, real-world applications within the healthcare setting. Therefore, this research proposes the following approach: First, the development of a model for predicting the likelihood of hospitalization within 24 hours and waiting times for patients visiting the ED. This model aims to enable ED physicians to rapidly and accurately assess a patient's likelihood of hospitalization, contributing to the alleviation of ED overcrowding. Furthermore, by utilizing bed occupancy data recorded at hourly intervals, predictions for the occupancy rates of individual wards and rooms can be made. This enables the establishment of medium and long-term bed resource management plans.

In the first chapter, we predict the likelihood of hospitalization within 24 hours and estimate waiting times for patients who pass through the ED. We refine electronic medical record (EMR) data derived from admission decision forms prescribed to patients passing through the ED, and build two versions of a Gradient Boosting Machine (GBM). The first model serves as a classification model for predicting hospitalization likelihood, while the second model offers a regression prediction model for estimated waiting times. Subsequently, we assess the influence of relevant variables through unstructured text data in the ED bed notes. By employing Natural Language Processing (NLP) to process unstructured text data recorded in the bed notes, we conduct additional training to compare its performance with the existing models. Furthermore, we utilize eXplainable Artificial Intelligence (XAI) to determine the impact of unstructured text data on the models.

In the second chapter, we predict the BORs of individual wards and rooms. We refine the bed occupancy time-series data recorded at one-hour intervals and group them into separate datasets corresponding to each ward and room. To predict the BORs of each ward and room, we build Deep Learning (DL) models. These DL-based models are trained on one-week and one-month BORs for each ward to predict the expected BORs. We employ various techniques, such as predicting short-term occupancy rates for individual rooms and combining them with the static features of the rooms. We compare the performance of six DL-based models and select the final model based on their performance.

The first study was published as " Prediction of hospitalization and waiting time within 24 hours of emergency department patients with unstructured text data" in Health Care Management Science on Nov, 2023.

Chapter 1.

Prediction of hospitalization and waiting time within 24 hours of emergency department patients with unstructured text data

1. Introduction

In addition to the ED serving as a key central access point for cases of critical emergencies, it functions as a primary health care system for patients who present with non-critical, concerning symptoms without an alternative option for outpatient care [26]. The overcrowding of EDs occurs as a result of the lack of hospital beds, distorted nurse-to-patient ratios, diagnostic errors, and ambulance diversions to the EDs, as well as delays in diagnostic and procedural practices [27-29]. ED overcrowding has negative consequences, including increased mortality rates, longer ED lengths of stay, treatment errors, lower rates of patient satisfaction, and challenges related to ambulance availability [21-23]. Moreover, as the medical industry is service-oriented, patient dissatisfaction is correlated with a decline in ED consultations and an unfavorable perception of hospitals. Consequently, the operation of EDs may be additionally adversely affected, thus impacting the hospital's financial management [4].

Under the governance of the Ministry of Health and Welfare of South Korea, there are numerous operational medical institutions and health systems. Furthermore, the emergency medical system in Korea functions through the collaboration of emergency medical technicians, centers, institutions, and rooms. The emergency medical response system comprises accident scene management, transportation, stages, and communication systems. Among these stages, the hospital care stage occurs at the level of the emergency medical center [30].

The healthcare system in South Korea aims to qualitatively improve and expand emergency medical services through the "Comprehensive Plan for Emergency Medical System Improvement" policy initiative. The government has provided policy support to achieve this goal, and the decisions arising from this initiative have considerably influenced EDs nationally. In South Korea, the "24-hour Emergency Department Restriction Act" was enacted in December 2017 for upper-level general hospitals to address the need for the quick examination and diagnosis of critically ill emergency department patients, as well as the resolution of overcrowding in the ED. According to this law, the percentage of patients with an ED length of stay exceeding 24 hours should be maintained within 5% annually.

Additionally, in 2004, South Korea introduced the National Emergency Department Information System (NEDIS) to computerize and manage the medical records of patients presenting at emergency medical facilities. This system enables real-time tracking and sharing of the emergency conditions and treatment details of patients. In 2017, by utilizing the medical records collected by NEDIS, treatment delays for critically ill patients, overcrowding in EDs, and lack of diagnostic reliability were nationally addressed by assessing this data using the Korea Triage and Acuity Scale (KTAS). KTAS is a comprehensive patient classification system that spans from the pre-hospitalization to in-hospital stages, determining patient priorities and urgency levels. These efforts have led to the emergence of specialized hospitals that evaluate various medical facilities, healthcare systems, and medical data. These hospitals serve as the foundation for research and technological development to enhance South Korea's emergency medical system [31].

In this study, our motivations included contributing to the emergency medical system and alleviating overcrowding in the ED, thus improving the ED environment. To act on these motivations, we implemented the following steps: First, we developed models that could hypothetically predict the likelihood of whether ED patients would require hospital admission and estimate the waiting time [32]. This was theorized to enable support for the decision-making process of ED physicians, allowing for the minimization of the proportion of patients with an ED length of stay exceeding 24 hours. Second, we compared the performance of a training model with NLP generated variables, from unstructured text data to models without it. The free handwritten text notes of ED doctors and nurses are a mixture of Korean, English, numerals, and symbols, conveying a substantial amount of information [33]. Consequently, the rapid reproducibility and easy maintenance of hospitalization prediction models emphasized the importance of text processing and comparing the performance of models. Finally, we utilized XAI to distinguish ineligibility factors, explaining why the patient could not be admitted. Thus, this allowed for identifying departments with waiting time delays and minimizing these waiting times by improving patient flow.

Therefore, to recapitulate, we aimed to develop a practical and cost-effective tool to assist ED physicians in making clinical decisions using EMR. Recently, Johns Hopkins Hospital introduced a command center that utilizes an electronic dashboard to facilitate systematic communication with hospital staff [34]. Figure 1 represents a visualization of a virtual dashboard to potentially be applied in a hospital, integrating the results of this study. This visualization dashboard is designed to be practical for ED physicians. Displaying real-time information to medical staff and informing patients about the estimated waiting time on the screen reduces the number of ED consultations and lessens waiting time in the operating room. Furthermore, forbearance is improved, and anxiety is reduced in patients. Providing patients with delayed information diminishes hospital revenue losses by preventing them from leaving the hospital abruptly [35-36].

In the Related Works section below, an overview of the existing literature addressing overcrowding in EDs has been performed. The Methods section covers the datasets used, pre-processing steps, training and evaluation methods, explanations of AI model algorithms, and XAI techniques. In the Results section, the findings and performance metrics of this study have been presented. The Discussion section provides an interpretation of the results of this study, highlights its limitations, and suggests scope for improvement. Finally, the Conclusion section provides an overall summary of this study, in addition to a detailed explanation of its clinical implications.

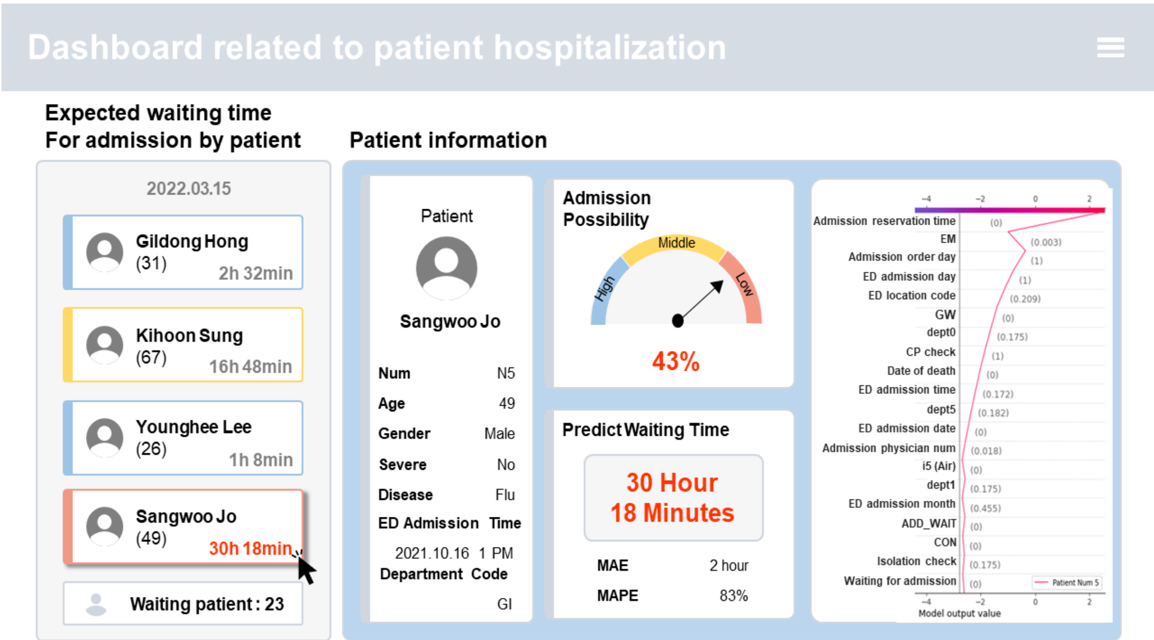

Figure 1. A synthesis of the findings of this study: A virtual visualization dashboard depicting the prediction of the likelihood of hospitalization and estimated waiting times for patients admitted to the ED. A patient's likelihood of hospitalization is indicated by blue (High), yellow (Medium), and red (Low) in order of probability. Displayed in the table on the right, in order of importance from the top, are the causes of patients not hospitalized within 24 hours of presenting to the ED (Please note that the names of the patients provided are hypothetical, as only de-identified data were extracted)

2. Related works

Patients often perceive the waiting time for admission to be longer than the actual elapsed time, and reducing this perceived waiting time has been shown to increase patient satisfaction. Various studies have suggested several methods to reduce the perceived waiting time of patients. These potential strategies include improving interpersonal interactions, providing patients with information and guidance on appropriate waiting times, and considering changes in staffing levels [36-38]. Physicians need to counsel patients regarding lengthy waiting times until admission or when hospital admission is not possible. Large data and predictive analytics from EDs improve the efficiency of emergency medical services and enhance the treatment of patients. This results in personalizing patient care, improved efficiency, and limited wasteful spending by providing practical guidance without investing in extra resources [40]. To do this, we require efficient and cost-effective decision-making tools to help ED physicians [41].

One technique to alleviate overcrowding is to admit the appropriate ED patients to the hospital promptly. Numerous studies have been conducted on the admission prediction for mitigating ED overcrowding. A high-performance logistic regression (LR) admission prediction model incorporating demographics, management, and clinical data can be routinely obtained in a hospital, informing patients about the likelihood of their admission [5]. One model uses patient information commonly acquired during ambulance transportation [6]. Based on the baseline characteristics provided on presenting at an ED, yet another predictive model divides a patient's electronic health record (EHR) dataset regarding statistics on prior medical use, past medical histories, insurance companies, and employment agencies into three categories (triage, history, full) [7]. The optimal waiting time for patients with low acuity grades using historical patient data, such as a machine learning (ML) tree-based predictive model, has been predicted [4]. Moreover, a model exists for predicting the expected waiting time from classification to consultation; therefore, predictive models of the median and 95th percentile waiting times of patients based on queuing theory have been studied [8].

The findings of some studies have revealed better performance of hospital admission prediction models by including text data. A study was conducted to train a deep neural network (DNN) using unstructured text data from EHR datasets of pediatric emergency patients. The model from this study showed that the DNN achieved a high-performance score of 0.892, which was 2% higher in the area under the curve (AUC) than that of the model not including text data [9]. Additionally, LR trained by NLP on the reason for a patient's visit demonstrated a 2% higher performance score at 0.846 compared to when text data was not included [10].

While acknowledging the considerable value of previous studies on mitigating overcrowding in EDs, limitations remain in the practical implementation of these studies. This is primarily due to a lack of intuitiveness in supporting ED physicians and effectively applying these models in real ED scenarios for patient communication [4-10]. Therefore, we have improved on the existing literature by developing our approach based on these previous studies.

3. Methods

3.1 Materials

3.1.1 Study design and setting

This retrospective, single-center, cohort study included a total of 271,143 patients who consulted at the hospital's ED between June 2018 and May 2022. Data was collected from 192,240 patients ≥ 19-years-old. Of these, 49,266 patients required hospitalization. The exclusion of 142,974 patients from the dataset was due to the difficulty in distinguishing between patients with prolonged waiting times and those who were transferred back to the ED based on the discretion of the physician. Patients who were transferred back to the ED had to wait for hospitalization and frequently experienced a total waiting time exceeding 24 hours, which made them ineligible for our study. Although this exclusion could have introduced bias to our results, the decision was made because it was challenging to distinguish the reasons for transfer in the dataset. This is depicted in Figure 2.

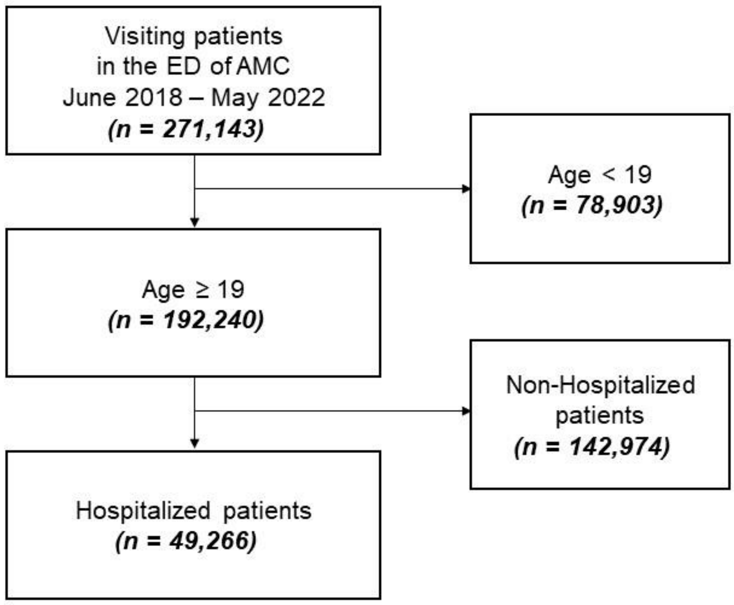

Figure 2. The number of patients who consulted at the ED between June 2018 and May 2022, requiring hospitalization for further care, is depicted. Patients < 19 years old and those who did not require hospitalization were excluded from the final data analyses

This study used ethically pre-approved data and underwent Asan Medical Center (AMC) Institutional Review Board (IRB) review (IRB 2021-0321), which was conducted by the Declaration of Helsinki (2008). De-identified data were extracted from the "data" and "clinical research data" warehouses.

3.1.2 The process from ED registration to inpatient hospitalization

As per the government guidelines, the ED length of stay should be ≤24 hours after arrival at our hospital in Seoul, South Korea. Upon arrival at the ED, patients first undergo registration and then proceed to the hospital office to obtain the ED treatment application form. Thereafter, they move to the triage area to undergo a preliminary medical examination and are subsequently assigned to a treatment room based on the severity of their condition. After the patients have submitted their medical receipts to the ED nurses in the treatment area, the ED physicians will review the recommended treatment. Depending on the severity of the condition or the appropriateness of hospitalization, patients are discharged from the ED. If the patient requires hospitalization for treatment, a hospitalization decision form is prepared. Subsequently, the patients are admitted on the designated admission date stated on the hospitalization decision form. The input variables we use for prediction are based on this setup, allowing immediate predictions afterward. The waiting time in the ED refers to the period from ED registration until receipt of the hospitalization decision form, with a discharge from the ED. The models we have developed predict the likelihood of hospitalization and estimate the waiting time, with a borderline within 24 hours of the patients presenting at the ED. Depending on the likelihood of admission, if the expected waiting time exceeds 24 hours, the patient may be advised to be transferred or discharged. This process is shown in Figure 3.

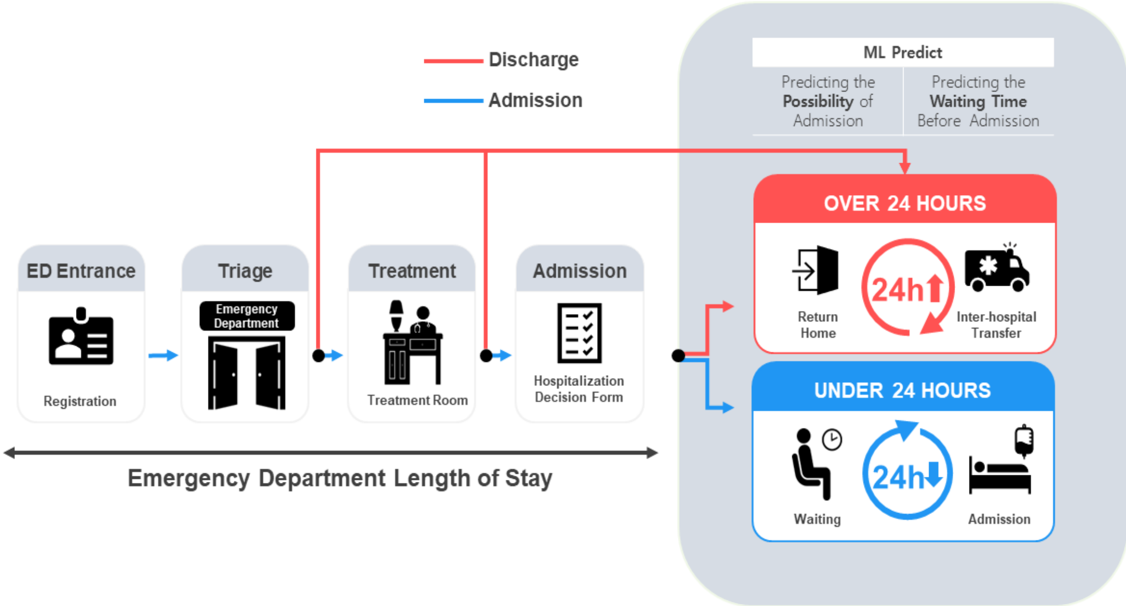

Figure 3. This model of prediction flow is based on the process from ED registration to hospitalization. The red line represents patients who have been discharged from the ED, while the blue line represents hospitalized patients. Depending on the discretion of the ED physicians, decisions regarding discharge can be made in "Triage" and "Treatment." The ED length of stay is defined as the waiting time from the arrival of the patients at the ED to when they are discharged from the ED. According to the prediction of the ML model, a patient who has received a hospitalization decision form will be classified as "discharged" if the total waiting time exceeds 24 hours or classified as "hospitalized" if it does not

3.1.3 Description of independent variables

Demographic, administrative, and clinical variables were collected from the information provided during the process of the ED visit registration to the hospitalization decision form. A detailed presentation of the variables generated for each category of data in Table 1 is depicted.

Table 1. Independent variables description

Category	Data source	Description
Administrative	Non-Hospitalized patients	A patient transferred to another hospital after visiting the ED
	When ED registration	Year, Month, Date, Day, Time, Hour, and Minute
	Hospitalization decision form	Admission order date, Expected date of admission, Scheduled surgery date, Diagnosis code in hospital, Reason for exceeded schedule, Tuberculosis ward admission possible, Organ transplant, CP application, Isolation, Single room availability, Withholding the hospitalization decision form, Suspension of patient treatment, Single room availability at the time of admission reservation, Exceeded schedule, The time from the moment patients visit the ED to a decision for admission
	In emergency department	Pregnant, Severe emergency disease, Admission decision check, Delay in treatment,

		Delay in treatment decision, Waiting for admission, Inspection, Reading, Follow-up, Cross-consultation, Hospitalization decision, Check-out, Cardio-pulmonary resuscitation (CPR), Do not resuscitate (DNR), Intensive care unit (ICU), Multiple traumas, Emergency surgery, Helicopter transport; Code of ED area, National emergency department information system (NEDIS), Classification for admission through ED, Purpose of treatment; Patient number and type, Doctor's number, Admission officer's doctor number;
	ED sickbed information note	Information that can be obtained in the ED sickbed being a brief note in Korean. The word was made into a variable according to the frequency, and the description of each variable is as follows: Unknown meaning missing value; Acute care unit (ACU); Acute stroke unit (ASU); Critical care unit (CCU); Neurological intensive care unit (NRICU); General ward meaning the rest of the wards except for special departments such as ICU/ED; Respiratory syncytial (RS); Emergency; Level; Possible; Multi-person rooms; Assignment; Preemptive; Person room; Waiting for hospitalization outside; Admission postpone; Severe; Confirm;
Demographic	Patient information	Hospital registration number, Gender, Birth, and Age
Clinical	Medical department data	Department code when making a reservation for admission (dept0), Department (dept1), Main department code in ED (dept2), Department code of admission (dept3), Department code in ED (dept4), Department code (dept5), Admission department code on the hospitalization decision form (dept6)
	Hospitalization decision form	Isolation type: Unknown, Air, Protect, Droplet, Contact, and Blood

3.2 Data pre-processing

3.2.1 Pre-processing of structured data

Table 2 describes and lists the main structured variables of this study. Duplicate or “Admission” tagged features were removed because they were classified as future variables and were unnecessary for this study. To calculate the waiting time, we created a variable that subtracted the ED visit time from the admission reservation time and converted it into hourly units. Text-type features were converted into integers. For example, ‘4.0’ indicated the ED location code was converted to the number “4.” Missing value transformations were determined based on the data type. Variables were populated with “Unknown,” “U,” or “0.” The target value, which is the time spent in the ED, was expressed as a decimal by converting minutes into hours.

Table 2. Description of the major variables that affect the outcome among all the variables

	Type	Variable Name	Description
Patient Information			
Patient Number	Number	num	Randomly
Age	Number	age	19, 20, …103, 109
Gender	2 categories	gender	Male, Female
Kinds of Department code	7 categories	dept0, dept1, dept2, dept3, dept4, dept5, dept6	GI[a], ONC[b], OBY[c], CV[d], NR[e] …

Information in ED			
ED Location Code	17 categories	ed_loc_code	1, 2, 3, 4(Severe level), 5, 7(Waiting for hospitalization level), 6(Mild level), C(Cancer), D(Delivery), R(CPR) …
ED NEDIS Location Code	7 categories	ed_nedis_code	1.0, 2.0, … 6.0, 8.0
Severe Emergency Disease	2 categories	severe	No(N), Yes(Y)
Hospitalization Decision Check	2 categories	adm_check	No(0), Yes(1)
Waiting for Admission	2 categories	waiting	No(0), Yes(1)
Via ED Classification	15 categories	via_ed_code	11.0 (Via general ward), 28.0 (Via intensive care unit) …
Hospitalization Decision Form			
Admission Physician Number	577 categories	adm_dr_num	D990010, D130578, D070498 …
Admission Order Date	883 categories	adm_day	20180614, … 20201101
Diagnosis Code in Hospital	2811 categories	dx_code	U(Unknown), D003017, D012602 …
Organ Transplant Check	3 categories	organ	No(0), Yes(1), Unknown(U)
Isolation Check	3 categories	iso	No(0), Yes(1), Unknown(U)
ED Admission Time			
ED Admission Year	3 categories	ed_y	2018, 2019, 2020
ED Admission Month	12 categories	ed_m	01, 02, 03, … 10, 11, 12
ED Admission Date	31 categories	ed_d	01, 02, 03, … 29, 30, 31
ED Admission Time	1440 categories	ed_time	0000, 0001, … 2358, 2359
ED Admission Hour	24 categories	ed_hour	00, 01, 02, … 21, 22, 23
ED Admission Day	7 categories	ed_days	Mon, Tue, Wed, … Sat, Sun
Admission Reservation Time	72 categories	adm_reserv_t	1.0, 2.0, … 98.0 ,118.0 (hour)

[a]GI: gastroenterology division

[b]ONC: oncology-hematology division

[c]OBY: obstetrics and gynecology

[d]CV: cardiology division

[e]NR: neurology

3.2.2 Pre-processing of unstructured text data

Variables with unstructured text included “Isolation types” and “ED Sickbed Information Notes.” Each text variable was pre-processed through normalization, tokenization, word frequency counting, term frequency-inverse document frequency (TF-IDF), and missing value filling.

First, after filling in the missing values with “Unknown,” unnecessary words, such as special characters and numbers, were removed. Additionally, after consolidation, all letters were changed to lowercase [42].

Second, for tokenization, the length was cropped to 2–5 words for word extraction. Korean words that were not automatically tokenized were manually spaced [42]. Table 3 shows the original and modified text through normalization and tokenization.

Table 3. Pre-processing results of unstructured text data. In case Korean and English are mixed in the text, the Korean meaning is written in parentheses for clarity

Raw data	Normalization	Tokenization
Isolation type		
접촉주의(contact precaution)	접촉주의,	접촉주의
비말주의(droplet precaution)	비말주의	비말주의
접촉주의+, 비말주의	접촉주의+, 비말주의	접촉주의 비말주의
공기(air), 비말(droplet), 접촉주의	공기, 비말, 접촉주의	공기 비말 접촉주의
ED sickbed information note		
GW)다인실가능(shared room available)	gw, 다인실가능	gw 다인실가능
gw,Lv4(다인실(shared room)),rS(N)	gw, lv, 다인실, rs	gw lv 다인실 rs
CSICU/전(pre) PS prof	csicu, ps, prof	csicu ps prof
GW) VRE	gw, vre	gw vre
GW 1O-> 타(other)FA	gw, 타 fa	gw 타 fa

Third, the tokenized words were converted into bag-of-words-encoded vectors. In the "Isolation type" column, only words with a document frequency (DF) of ≥ 10 and "Sickbed Information Notes" of ≥ 500 were extracted. The df number criterion was set based on the interval in which the frequency of tokens generated for each column sharply differed.

Fourth, we applied TF-IDF, which calculates the importance of words in a document-term matrix (DTM) [42]. This technique assigns weights to specific words in a DTM by applying a specific formula to the word and document frequency. It is simple to use and characteristically performs better in ML models.

The TF-IDF score created a DTM. Thereafter, the value was obtained by multiplying TF, which represented the relative frequency of term t in document d, and IDF measured the importance of the term t in the corpus. This can be expressed in the equation as follows:

$$TF(t, d) = \frac{Number\ of\ occurrences\ of\ term\ t\ in\ document\ d}{Total\ number\ of\ terms\ in\ the\ document\ d}$$

$$IDF(t, D) = \frac{Total\ number\ of\ documents\ in\ the\ corpus}{Number\ of\ documents\ with\ term\ t\ in\ them}$$

$$TF - IDF = TF(t, d) \times IDF(t, D)$$

Finally, the TF-IDF scores for each selected word were transformed into a two-dimensional array. Rows processed as missing values at this time because they did not contain words were replaced with 0.0. The variables generated through this process were assigned new names, as indicated in Table 4.

Table 4. Description of variables created by TF-IDF

	Type	Variable Name	Description
Isolation type	6 categories	i0, i1, i2 i3, i4, i5	Unknown, Air, Protect, Droplet, Contact, Blood
Bed Information	15 categories	ACU, ASU, CCU, EM, GW, NRICU, RS, POSSIBLE, SHARED ROOM, ASSIG, FIRST, ADD BEFORE, ADD WAIT, SEVERE, CON	Acute care unit, Acute stroke unit, Critical care unit, Emergency, General ward, National emergency department information system, Respiratory syncytial, Possible, Multiple rooms, Assignable, Preemptive, Admission postpone, Waiting for hospitalization, Severe, Confirm

3.2.3 Target variable

The "target value" of the waiting time prediction model is a numerical value denoting the waiting time. For the classification model regarding the likelihood of hospitalization in ED, the target value was standardized as 24 when it exceeded 24 hours for training. If the actual waiting time was < 24 hours, it was converted to the number "0". If it was ≥ 24 hours, it was converted to the number "1". The result value was set in binary format.

Finally, a total of 82 variables were used, including 61 structured data variables and 21 variables generated by NLP from unstructured text data.

Figure 4 depicts the overall flow of the prediction method and usage of XAI in this study.

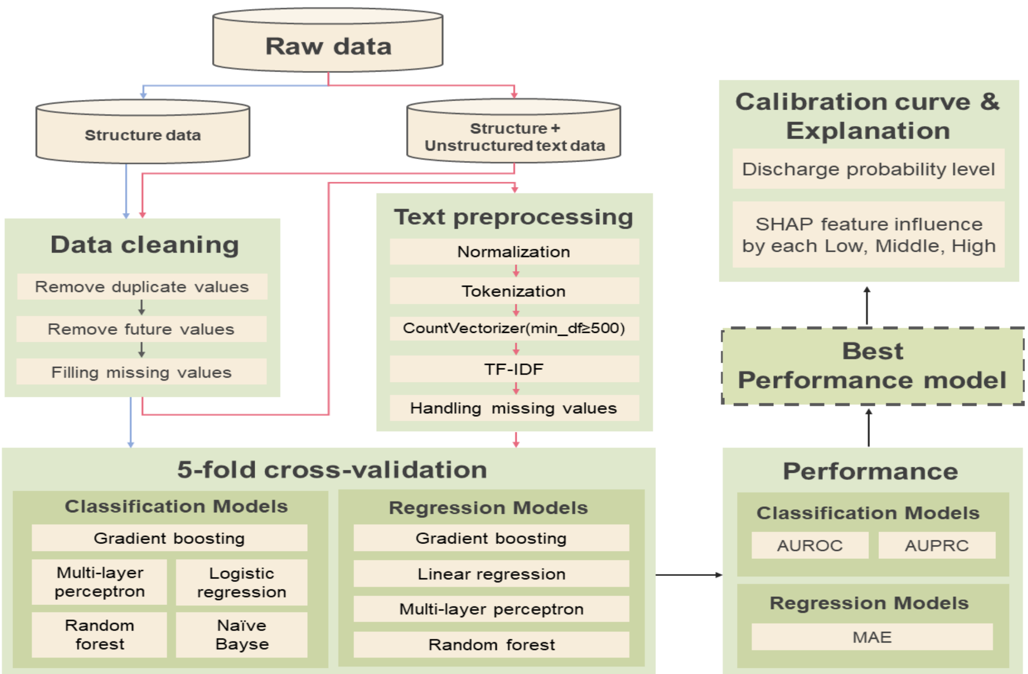

Figure 4. The overall flow of the prediction method and XAI usage in this study are depicted

3.3 Admission prediction ML-based models

In this study, two ML approaches were applied based on the research objectives. First, the ML classification model was used to predict the likelihood of hospitalization for patients waiting in the ED. Second, an ML regression model was used to predict the waiting time until hospitalization.

We evaluated five different ML models to predict the likelihood of hospitalization. These included the Gradient Boosting Machine (GBM) [43] and Extreme Gradient Boosting (XGB) [44] models. Typically, XGB aims to improve models by sequentially and iteratively training multiple decision trees on training data. It minimizes prediction errors and avoids overfitting through regularization techniques, ultimately improving generalization performance. Random Forest (RF) [45] is an ensemble learning algorithm based on decision trees, where multiple decision trees are combined to create a single model. During the construction of decision trees, features and data are randomly selected. Additionally, the prediction results are aggregated to prevent overfitting and achieve stable predictions. Multi-layer perceptron (MLP) [46] is a neural network architecture consisting of multiple hidden layers. MLP is capable of modeling non-linear correlations, making it suitable for addressing complex problems. The weight updates using backpropagation enable learning from diverse inputs and outputs. LR is a statistical technique [47] that models the correlation between inputs and outputs. It involves taking a linear combination of input features and weights, passing it through a logistic function to generate probability values between 0 and 1, and using a threshold to predict classes. LR is simple and interpretable; however, it is limited in modeling non-linear correlations among input features. Naïve Bayes (NB) is a supervised learning algorithm based on the principles of Bayes' theorem. This model assumes that input variables are independent and directly impact only one output variable, making it computationally efficient and quick to train [48]. Nonetheless, in many real-world scenarios, it may be unrealistic to assume such independence among input variables. This limitation can restrict the practical applicability of the model.

We used XGB, RF, MLP, and LR [49] among the ensemble techniques to determine the waiting time until hospitalization. Moreover, among the five classification models described earlier, XGB, RF, and MLP can be used for regression tasks. Thus, the same models were used to streamline the training process. Additionally, LR is a method used to model correlations between two or more variables, assuming the correlations between the variables are linear. Therefore, changes in one variable are proportionally related to changes in other variables. LR creates a prediction line based on these suppositions, making it simple and easy to interpret. However, with real-world data, complex non-linear correlations often exist, making it difficult to model.

The performance of these models was compared by training them on both structured and unstructured text data. Finally, we used XGB which resulted in it being the model with the best performance. For the hyperparameters, a grid search was used for fine-tuning.

3.4 Cross-validation

Between June 2018 and May 2022, a total of 49,266 patients who consulted at the hospital's ED in Seoul, South Korea, required hospitalization. Of these, 44,753 patients (90.8%) were admitted within 24 hours, while the remaining 4,513 patients (9.2%) were not admitted within 24 hours. Despite the imbalance in the data classes, our goal was to use 5-fold cross-validation without shuffling the data to assess the predictive performance of the admission prediction model on new patient data in the future. Data bias is conventionally reduced by 5-fold cross-validation in ML model evaluation, thus improving accuracy [50]. Five cross-validations were performed by dividing 44,753 patient data into five equal parts, with 80% of the data used for training and 20% used for validation.

3.4.1 Data transformation

We transformed the data at every fold validation for model training. The data were divided into categorical and numeric types based on variables. For the categorical type, the target encoder had to be applied. This encoding method is characterized by expressing the correlations between similar categories. However, this correlation is limited to only categories and targets. The advantage of this method is that it facilitates fast learning without compromising on the quantity of data; consequently, we implemented it in our model. To achieve balance, the mean was considered by setting the smoothing to 1.0 and minimum samples to 1. Furthermore, any columns with zero variance had to be dropped. For numeric types, a MinMaxScaler was used. This ensured that the value of each variable adhered to a specific range or rule; here, the data was converted to a range between 0 and 1.

3.4.2 Performance

The predictive performance of the ED hospitalization model was evaluated using the area under the receiver operating characteristic (AUROC) and the area under the precision-recall curve (AUPRC). The receiver operating characteristic (ROC) curve was used to indicate the performance of a classification model when determining a patient's hospitalization status that was represented in binary form as positive or negative. The X-axis represented the false positive rate (FPR), and the Y-axis represented the true positive rate (TPR) with a proportional correlation between the two. Measuring the AUROC was used to evaluate performance. The closer this value was to 1, the better the model and the higher the likelihood of hospitalization. The AUPRC was indicated by setting the X-axis to recall and the Y-axis to precision. These two values are inversely proportional to each other and form a downward curve towards the right. Both precision and recall are ideal models when close to 1; therefore, the closer the value of AUPRC is to 1, the better the model.

The performance of the predictive waiting time model for patients consulting at the ED was evaluated using the mean absolute error (MAE). MAE is the average of the absolute sum of the differences between the predicted waiting time by the regression model and the actual time. The error size was accordingly reflected. Because MAE is a highly sensitive indicator, the difference between the predicted ED waiting times by the model can be recognized immediately. In the given formula, e_i represented the difference between the actual value, y_i, and the predicted value, $ypred_i$. MAE was calculated as the average of the absolute differences divided by the total number of data points:

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |e_t| \quad (1)$$

$$* \ e_i = y_i - ypred_i$$

3.5 Calibration curve and SHapley Additive exPlanations

We utilized a calibration curve to identify the factors that impacted the determination of discharged patients in general ED scenarios. Additionally, we employed SHapley Additive exPlanations (SHAP) for XAI.

The calibration curve revealed the realistic predictive correlation between the predicted and actual probability of admission likelihood. Thus, if the patient's admission probability was 80%, it indicated that the actual admission probability was also 80%. We used the calibration curve of XGB with text data that had the highest performance among the models for predicting hospitalization. The X-axis of the calibration plot of this model was divided by 0.1 units. After calculating each section's event rate, the predicted and actual probabilities were displayed as a bar plot. Moreover, the proportion of non-hospitalized patients was similarly calculated and divided into three categories: Low (0–0.5), Medium (0.6–0.8), and High (0.9–1.0). The criteria for dividing the levels were based on the proportion of patients who were not admitted within 24 hours at each interval. The classes Low, Middle, and High comprised 1017, 173, and 245 patients, respectively.

SHAP is a technique that transparently reveals the internal functioning of a complex AI model. A Shapley value is calculated by measuring the average change in the presence or absence of a feature when combining multiple features, assisting in distinguishing the importance of each individual feature. Additionally, the importance of each feature can be calculated using the feature importance function in tree-based or boosting algorithms. However, this technique, used to identify variables, affects the prediction through permutation. Due to the estimation limit imposed by the degree of error, the importance of variables may vary each time the algorithm is executed, potentially leading to the oversight of dependencies between features. Therefore, models with correlations between features should avoid using the feature importance function. The Shapley value utilizes the concept of independence between variables as a key idea. It is used to calculate the impact of variables by comparing the results obtained when all combinations of variables related to a specific variable are input. The selection is made based on the effect on the target variable. Moreover, it can explain the negative and positive correlations between the value of the result and the variable. The technique should be selected based on the viewpoint. In our study, we chose to use SHAP because there was a high likelihood of a dependency between the features [51].

Each level was applied to the SHAP summary plot and plots bar. The SHAP summary plot can determine the magnitude of a feature based on its Shapley value. The plots bar calculates the global importance by averaging the absolute values of each Shapley value. Consequently, it is possible to grasp the detailed influence of the variable on the model. Among the Shapley values that were generated using the plots bar, the top eight values were depicted using radar plots.

4. Results

4.1 Performance of the ML-based predictive models

Table 5 and Table 6 show the AUROC and AUPRC metrics for each fold of a 5-fold cross-validation of ML models that have been trained to predict the likelihood of hospitalization within 24 hours after arrival at the ED. Comparing the two models, the models trained with unstructured text data outperformed those without unstructured text data by 6–10% and 20–28% on AUROC and AUPRC, respectively. The average values of AUROC and AUPRC for XGB, including text data, were the highest compared to other classification models. The AUROC and AUPRC were 0.922 (SD 0.030) and 0.687 (SD 0.085), respectively. By presenting the mean AUROC and AUPRC of each classification model using text data in Figure 5 and Figure 6, XGB evidently exhibited the best performance.

Table 5. Evaluation of 5-fold cross-validation for each ML model using AUROC and AUPRC

	RF		MLP		LR		NB		XGB	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
Fold 1	0.845	0.323	0.864	0.286	0.846	0.225	0.722	0.241	**0.228**	**0.412**
Fold 2	0.875	0.395	0.789	0.259	0.785	0.228	0.740	0.201	**0.227**	**0.437**
Fold 3	0.862	0.587	0.812	0.482	0.824	0.447	0.779	0.383	**0.875**	**0.604**
Fold 4	0.824	0.371	0.792	0.335	0.752	0.206	0.717	0.188	**0.837**	**0.400**
Fold 5	0.756	0.472	0.684	0.335	0.755	0.363	0.720	0.362	**0.822**	**0.542**
Mean (SD)	0.832 (0.042)	0.430 (0.092)	0.788 (0.058)	0.339 (0.077)	0.788 (0.035)	0.294 (0.095)	0.736 (0.023)	0.275 (0.082)	**0.800 (0.025)**	**0.480 (0.080)**

Table 6. Evaluation of 5-fold cross-validation for each ML-model with NLP using AUROC and AUPRC

	RF with NLP		MLP with NLP		LR with NLP		NB with NLP		XGB with NLP	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
Fold 1	0.891	0.525	0.896	0.474	0.890	0.436	0.759	0.360	**0.922**	**0.577**
Fold 2	0.948	0.703	0.936	0.701	0.920	0.609	0.892	0.630	**0.953**	**0.735**
Fold 3	0.946	0.780	0.930	0.726	0.932	0.723	0..913	0.759	**0.955**	**0.822**
Fold 4	0.897	0.601	0.893	0.575	0.856	0.512	0.799	0.411	**0.908**	**0.634**
Fold 5	0.833	0.612	0.835	0.585	0.826	0.558	0.783	0.511	**0.875**	**0.667**
Mean (SD)	0.902 (0.042)	0.644 (0.088)	0.898 (0.036)	0.612 (0.092)	0.885 (0.039)	0.568 (0.096)	0.829 (0.061)	0.534 (0.145)	**0.922 (0.030)**	**0.687 (0.085)**

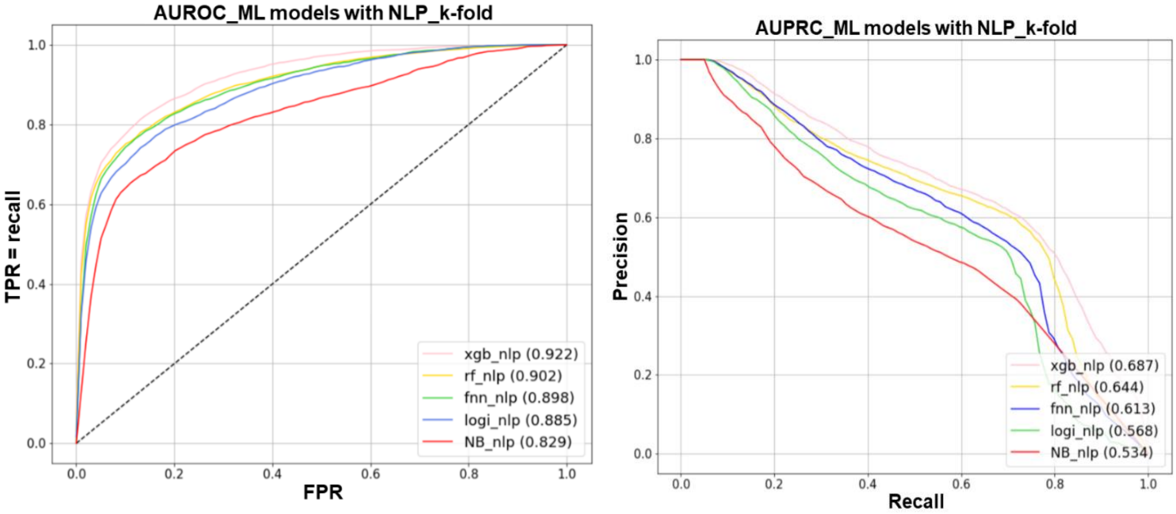

Figure 5, 6. AUROC and AUPRC curves of hospital admission prediction models with NLP

Table 7 and Table 8 show the MAE and average values for each fold of the regression ML model that predicts the waiting time from arrival at the ED to hospitalization. The MAE of the XGB model that included text data revealed the smallest difference when compared to that of the other models. The MAE of XGB with NLP revealed a difference of approximately 3.02 from the actual value and equated to 3 hours when converted into time. The difference was 30 minutes less than that of a normal XGB model, equating to 3 hours 30 minutes with a time conversion of 3.53.

Table 7. Evaluation by MAE of 5-fold cross-validation for each ML-model

	MAE			
	RF	MLP	LR	XGB
Fold 1	3.67	3.43	3.80	**3.53**
Fold 2	3.55	3.51	3.99	**3.55**
Fold 3	3.66	3.98	4.25	**3.52**
Fold 4	3.40	3.35	3.73	**3.38**
Fold 5	3.87	3.95	4.21	**3.66**
Mean (SD)	3.63 (0.154)	3.64 (0.267)	3.99 (0.209)	**3.53 (0.090)**

Table 8. Evaluation by MAE of 5-fold cross-validation for each ML-model with NLP

	MAE			
	RF with NLP	MLP with NLP	LR with NLP	XGB with NLP
Fold 1	3.35	3.00	3.55	**3.38**
Fold 2	2.66	2.90	3.48	**2.77**
Fold 3	2.69	2.71	3.45	**2.61**
Fold 4	3.03	3.00	3.51	**2.96**
Fold 5	3.54	3.67	4.02	**3.35**
Mean (SD)	3.05 (0.349)	3.05 (0.325)	3.60 (0.213)	**3.02 (0.310)**

We selected the XGB with text data as the final model. XGB has fast learning and classification speed due to parallel processing and strong durability with its overfitting regulation function. Furthermore, it demonstrates excellent predictive performance in both classification and regression tasks; therefore, it was hypothesized to predict the likelihood of hospitalization accurately and expected waiting time as a single model, making the process uncomplicated.

4.2 Variable importance through SHAP

The summary_plot of SHAP means that the performance contributing to the model's prediction increases as it moves up from the bottom of the y-axis. The x-axis represents the magnitude of each variable's impact on the outcome value. In our study, we found that the color red was associated with a delay in admission within 24 hours, while the color blue was found to have an impact on the admission process. For example, in the case of 'Bed information - Confirm', it most likely indicates a low likelihood of hospitalization. Conversely, in the case of 'ED admission day', where the blue color appears longer, it can be considered a variable that increases the likelihood of hospitalization within 24 hours. The SHAP results of XGB with the highest performing text data are shown in Figure 7. Table 2 and Table 4 were referred to for a description of the variables in Figure 7.

The SHAP summary plot demonstrates that the performance contributing to the model's prediction increases as it moves up from the bottom of the Y-axis. The X-axis represents the magnitude of each variable's impact on the outcome value.

In our study, we found that the color red was correlated with a delay in admission within 24 hours, while the color blue was found to have an impact on the admission process. For example, in the case of the variable "Bed information - Confirm," a low likelihood of hospitalization was indicated. Conversely, in the case of the variable "ED admission day," where the blue color appeared longer, the likelihood of hospitalization within 24 hours increased. The SHAP results of XGB with the highest-performing text data are shown in Figure 7. Table 2 and Table 4 are referred to for a description of the variables in Figure 7.

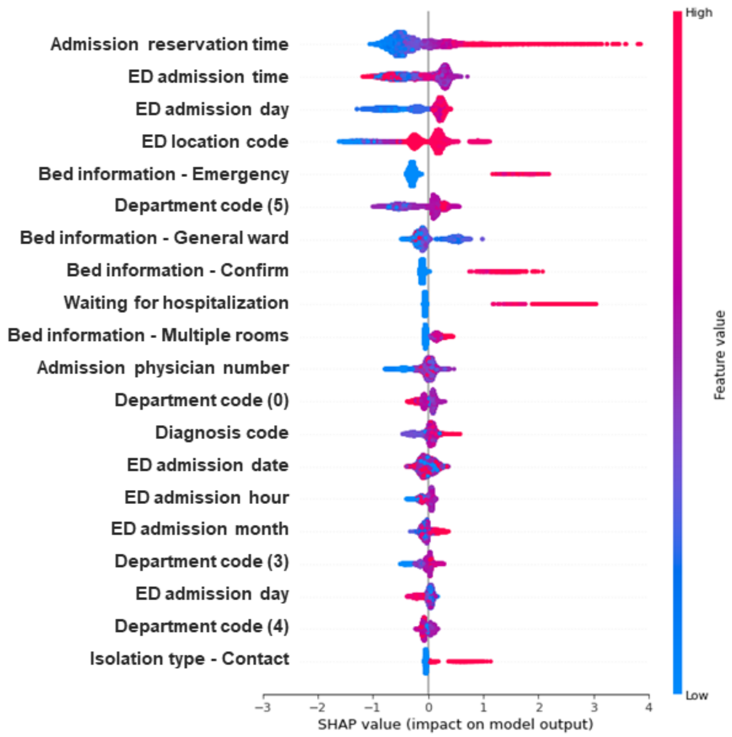

Figure 7. Shapley effect values for variables in the XGB model with text data

We analyzed the calibration curve to determine the statistical significance of the variables affecting patients who either had a low or high probability of being admitted within 24 hours of their arrival at the ED. In addition, the XGB model with the best performance was used to analyze the target value, set to “1,” of patients who were not hospitalized within 24 hours. The X-axis of the calibration curve was delineated as the average value, established as the predict_proba of the predicted value. The Y-axis delineated the average of the actual values (the fraction of positives correctly classified). By creating a histogram of the X- and Y-axes of the calibration curve, the ratio between the predicted value and the actual value was plotted, as presented in Figure 8.

As revealed in the histogram, the likelihood of events for observed discharged patients was higher than that of predicted discharged patients.

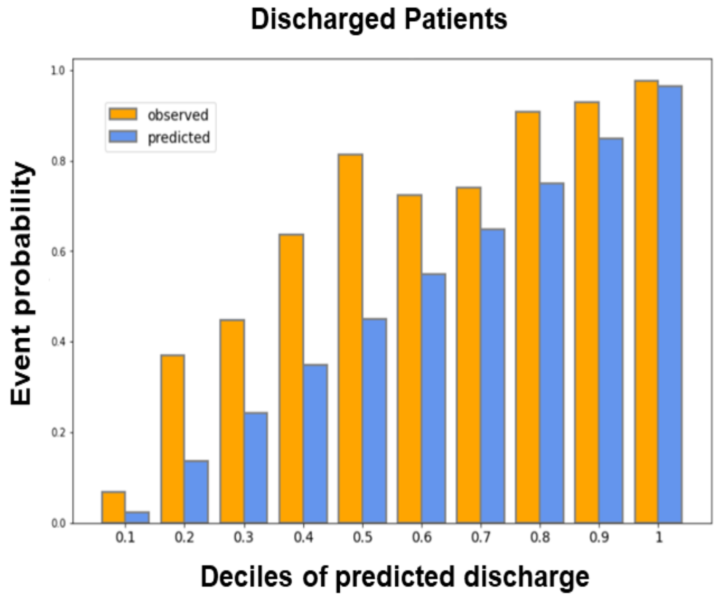

Figure 8 The X-axis is divided into deciles based on increments of 0.1 for the predicted discharge probability of the patients, while the Y-axis represents the actual discharge rate. The orange bar represents the rate of observed discharged patients, while the blue bar represents the rate of predicted discharged patients

The number of patients corresponding to each level was 10 186, 221, and 255 for the Low, Middle, and High level classes, respectively. Each level was confirmed by SHAP. We replaced the top eight quantifiable variables in the bar plots with radar plots, as depicted in Figure 9.

For the Low level class, “Admission reservation time” was the highest with a value of 0.47, followed by values for the “ED admission time” (0.35), “ED location code” (0.33), “ED admission day” (0.32), “EM” (0.3), “dept5” (0.28), “GW” (0.19), and “Date of admission” (0.15).

For the Middle level class, “ADD_WAIT” was the highest with a value of 1.18, followed by “EM” (1.11), “Admission reservation time” (0.94), “CON” (0.42), “ED admission time” (0.34), “GW” (0.31), “ED location code” (0.26), and “ED admission day” (0.23).

For the High level class, “Admission reservation time” was the highest with a value of 3.95, followed by values for the “EM” (1.06), “CON” (0.4), “ED location code” (0.38), “ED admission time” (0.3), “ADD_WAIT”

(0.28), “GW” (0.24), and “ED admission day” (0.22). Descriptions of the variables are mentioned in Table 2 and Table S4.

With these results, it is possible to demonstrate to the patient the likelihood of being admitted and inform them about their expected waiting time. If admission is not possible, the patient can be informed of the reason. As a result, EDs and hospitals will exhibit high operational efficiencies by implementing these methods.

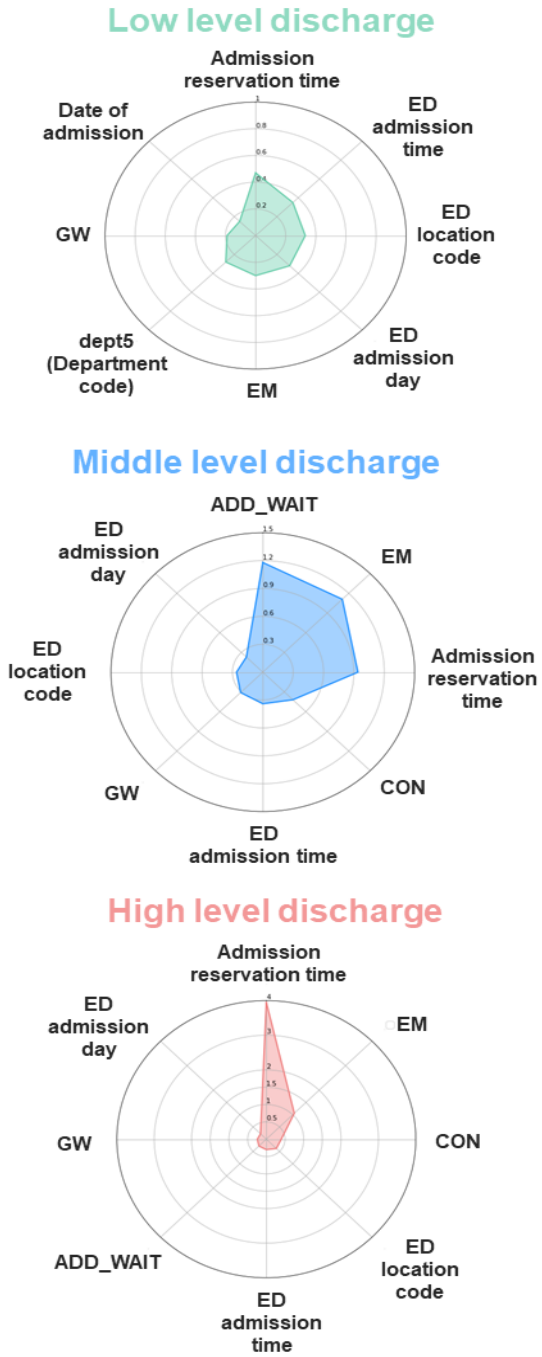

Figure 9. Radar plots of the influence of Shapley values of the top eight variables by Low, Middle, and High are shown. Green is Low, Blue is Middle, and Red is High

5. Discussions

The full dataset for our study utilized patient information obtained between June 2018 and May 2022 from a hospital in Seoul, South Korea, where patients were admitted through the ED. Moreover, this dataset included a specific subset of patients who were hospitalized in a designated COVID-19 environment. In addition to the economy and society, the recent global pandemic, COVID-19, has impacted the ED. After the outbreak of COVID-19, the waiting time in the ED has become irregular, as patients are only allowed to visit the hospital once their polymerase chain reaction (PCR) test results are confirmed in order to prevent the spread of infectious diseases. Furthermore, hospitals have experienced shortages of wards, as they have had to allocate infected and non-infected patients to separate wards [52]. Nationally, South Korea has established a centralized funding system [53], and this system offers patients considerable autonomy. Patients can consult at primary and secondary hospitals, as well as the ED, without requiring a referral. Additionally, if patients receive an initial diagnosis from a physician, they can consult at tertiary hospitals with a referral form issued by the diagnosing physician. Moreover, the government supports cost-sharing programs that assist in reducing the financial burden on patients during hospital consultations [53-54]. In contrast, the healthcare workforce regarding medical personnel remains below the average, as advised by the Organisation for Economic Co-operation and Development (OECD) [55]. This shortage of medical staff leads to inadequate patient care that is further exacerbated by the challenge of managing a large number of patients.

We compared ML-based models trained on data during this period. XGB revealed a high performance with an AUROC of 0.860 and an MAE of 3 hours 30 minutes. Additionally, we trained unstructured text data on XGB to predict the likelihood of hospitalization. Our model achieved an AUROC score of 0.922 and an MAE of 3 hours, demonstrating strong performance. Our model had an AUROC that was 6% higher and a MAE that was 30 minutes less than the existing XGB. Regarding text manually entered by doctors and nurses in the ED, such as ED sickbed information notes, this contains information that can be immediately updated based on new environmental changes, including COVID-19. This differed from simple computer-entered demographic, administrative, and clinical variables.

The text data contained important information regarding the patient's condition or hospitalization [33]. We applied TF-IDF to this data and further trained it on the existing XGB to exhibit improved performance in predicting hospitalization and estimated waiting time. By applying SHAP to the XGB model with added text data, we revealed the contribution of features based on the probability level of discharged patients. Informing the patient about the reason for not being hospitalized involves clearly counseling the patient about the condition. This practice enables the efficiency and accuracy of emergency physicians, ultimately reducing overcrowding in the ED.

The common variable that showed the greatest effect across the Low-, Middle-, and High-level classes was the "Admission reservation time". In the High level class, with the least likelihood of hospitalization, the effects of the parameters gradually decreased in the order of "EM", "CON", and "ED location code". By utilizing this information to enhance the process in the healthcare system, hospitals can appropriately elevate patient admission rates and decrease waiting times.

The study's findings revealed that the variables had a statistically significant influence on the results, as demonstrated by SHAP. These findings can improve the efficiency of the operation of the ED by enhancing the processes related to variables in unpredictable situations, such as COVID-19. High-performance models and the identification of adjustable variables can facilitate the maintenance of ML models and augment existing models. Our model is a useful and inexpensive tool that provides decisive clinical support to ED physicians constructed on readily available patient medical data.

Despite this considerable performance, the estimated waiting time provided by the ED to the patient and the time perceived by the patient may be different. These differences should be resolved through effective communication with patients in the hospital. Additionally, the ED should create a comfortable waiting room environment and provide feedback on estimated waiting times, notifying the patient of any updates in progress.

Furthermore, it is important to regularly check the waiting room area and apply a personalized approach to patient care. If these series of improvements are integrated, the overcrowding of the ED will be reduced [56-58].

Several limitations should be considered in this study. First is the limitation regarding the objective evaluation of the clinical usability and acceptability. It may be challenging to empirically appraise the actual value of the application of results of this independent single-center study for other hospitals and clinicians. However, this study was initiated in response to a request from physicians working in the ED to address overcrowding issues. To make the algorithms more user-friendly for practical use in the ED, we developed ML models that provided predictive results and worked on software development. Furthermore, a systematic evaluation of the effectiveness of the EMR system is planned.

The second limitation is the lack of validation for external generalizability. Our models can be applied to predict the likelihood of hospitalization and waiting time in the ED of tertiary hospitals in Korea. However, since each hospital has different systems, users of models may experience the inconvenience of having to learn new functionalities. Additionally, the EMR systems utilized by each hospital have varying interfaces and terminologies for input variables. This discrepancy may lead to reduced accuracy when comparing the models of our study to other models if they are implemented without modifications. Although validation is possible for commonly used variables, differences in variable definitions could impact accuracy. Nevertheless, this research was initiated in response to clinical requests from hospitals, indicating a high probability of its relevance in other hospitals.

The third limitation of developing prediction models for hospital admissions in the ED of Korean hospitals is the data imbalance caused by including mixed data from before and after the outbreak of COVID-19. This limitation may restrict the performance of the models of this study. However, according to the policies of the Korean government and hospitals, the ED needs to be sufficiently flexible to accommodate various diseases that are unrelated to COVID-19. Thus, a system has been implemented that requires COVID-19-suspected patients to be admitted after PCR testing. Moreover, numerous changes have been made to the ED system in response to the COVID-19 outbreak. In this context, we believe it is appropriate to use mixed data, including data related to COVID-19, for predicting sudden changes in the system that we aim to address. Therefore, we value the high performance of our study results. This study proposed prediction models for situational changes that may occur in the healthcare system due to various disease occurrences, including COVID-19.

The fourth limitation is regarding the sole use of XAI which does not suffice when counseling patients on non-admission. While XAI offers interpretable features, it only contributes individual variables to specific predictions, making it challenging to fully comprehend the overall behavior of the model. Consequently, particularly if a specific variable has a substantial impact on non-admission, the explanation given to the patient would still be inadequate. Moreover, XAI relies on the trained model and specific data distributions, thus making it arduous when generalizing SHAP values once new data is introduced. Exclusively relying on SHAP values to drive changes in specific areas of the ED is limiting. However, ED physicians can still provide patients with appropriate explanations regarding non-admission using this functionality, supporting the counseling of patients.

Despite these limitations, we have developed a high-performance predictive model. This is the first time that we identified the contribution of variables by dividing patients who were not hospitalized by levels.

6. Conclusion

We created models that included unstructured text data that can be easily overlooked to predict the likelihood of hospitalization within 24 hours and estimated waiting times for admission for patients using the ED. The data consisted of demographic, administrative, and clinical information that was easily obtained from registration at the ED until admission to the hospital. The XGB with text data model predicts the likelihood of admission and the expected waiting time within 24 hours of an ED length of stay. Moreover, via SHAP it includes a function that identifies the variables that have the greatest impact on difficult admissions. It has been confirmed that applying NLP to unstructured text data has a considerable effect on the target.

This model will contribute to improving the accuracy and speed of decision-making for ED physicians. The influence of variables on the number of patients who are not admitted within 24 hours, as identified through XAI, provides crucial information for ED physicians to explain to patients who cannot be admitted under the "24-hour Emergency Department Restriction Act".

By utilizing this information, the estimated average length of stay for all patients can be determined by considering the estimated waiting times. This lets doctors promptly notify patients about non-admission and transfer options, thus optimizing hospital operational costs and reducing waiting times. Furthermore, hospital managers can obtain information to optimize waiting times, improve hospital processes, and minimize operational costs. Such information greatly improves hospital operations and economic efficiency and can be particularly useful in mitigating ED overcrowding.

After implementing this model in a real ED, it is crucial to collect feedback from healthcare professionals and consistently update and optimize the model's performance by incorporating new data and receiving continuous feedback. The model should incorporate real-time patient monitoring to instantaneously integrate the patient's status into the model. Furthermore, utilizing a customized model that considers individual circumstances and medical conditions can improve the accuracy of admission predictions. These suggestions should be followed to further advance the future scope of research. In the future, if this model is applied and utilized within a monitoring system, such as a command center [34], it is postulated that hospital processes will be greatly improved.

Chapter 2.

Forecasting hospital room and ward occupancy using static and dynamic information concurrently

1. Introduction

The global healthcare market continues to grow, but the burden of healthcare costs on governments and individuals is reaching its limits. Consequently, there is increasing interest in the efficient utilization of limited resources in healthcare systems, and hospitals must develop approaches to maximize medical effectiveness within budgetary constraints [59-60]. One approach to this is optimizing the utilization of medical resources. Medical resources can be broadly categorized into three categories: human resources, physical capital, and consumables. The appropriate and optimized utilization of these resources is critical for improving healthcare quality and providing care to a larger number of patients [61-62].

Among the three medical resources, hospital beds are considered one of the physical capitals provided by hospitals to patients. These beds are allocated for various purposes such as rest, hospitalization, postsurgical recovery, and more. They constitute one of the factors that can directly influence the patient's internal satisfaction within the hospital. However, due to limited space, hospitals often have a restricted number of beds. Moreover, the number and functionality of beds are often fixed due to budgetary or environmental constraints, making it difficult to make changes. Nonetheless, if hospital administrators can evaluate BORs according to different time periods, they can predict the need for healthcare professionals and resources. On the basis of this information, hospitals can plan resources efficiently, reduce operational costs, and achieve economic objectives [63]. In addition, excessive BORs can exert a negative effect on the health of staff members and increase the possibility of exposure to infection risks. Hence, emphasizing only maintaining a high BORs may not necessarily lead to favorable outcomes for the hospital [24-25]. Considering these reasons, BOR prediction plays a vital role in hospitals and is recognized as a broadly understood necessity for resource optimization in the competitive medical field.

1.1 Prior Works

Hospital bed occupancy prediction has been investigated using various approaches recently. From studies predicting bed demand using mathematical statistics or regression equation models based on given data [11-14], the focus has shifted towards modeling approaches using time-series analysis. This approach observes recorded data over time to predict future values.

There is research that took an innovative approach using time-series analysis alongside the commonly used regression analysis for bed demand prediction. In this study, it demonstrates that using time-series prediction for bed occupancy yielded higher performance results than using a simple trend fitting approach [15]. In previous research, they used the ARIMA model for univariate data and a time-series model for multivariate data to predict BOR [16]. With the advancement of deep learning (DL) models that possess strong long-term memory capabilities, such as the recurrent neural network (RNN) and long short-term memory (LSTM), there has been an increase in studies applying these models to time-series data for prediction purposes. For instance, in a study by Kutafina [17], hospital BORs were predicted based on dates and public holiday data from government agencies and schools, without involving patients' personal information. That study used a nonlinear autoregressive exogenous (NARX) model to predict a short-term period of 60 days, with an aim to contribute to the planning of hospital staff. The model demonstrated good performance with an average mean absolute percentage error (MAPE) of 6.24%. In emergency situations such as the recent global COVID-19 pandemic, the sudden influx of infected patients can disrupt the hospitalization plans for patients with preexisting conditions [64]. Studies have been conducted using DL architectures to design models for predicting the BOR of patients with COVID-19 on a country-by-country basis. Some studies incorporated additional inputs such as vaccination rates and median ages to train the models [18]. Studies have also been conducted to focus on the short-term prediction of BOR during the COVID-19 period [19-20].

Table 1. Summary table of prior works

Citation Number	Year	Data Set	Method	Prediction Target
Mackay M, and Lee MD [11]	2007	De-identified data the date and time of patient admission and discharge between 1998 to 2000	Comparison of two compartment models through cross-validation	Entire hospital bed occupancy (annual average)
Littig, SJ, and Isken MW [12]	2007	Historical and real-time data warehouse. hospital information systems (the emergency department, financial, surgical scheduling, and inpatient tracking systems)	Computerized model of MLR[a] and LR[b]	Entire hospital short-term occupancy (24h or 72h) based on LOS[c]
Kumar A, and Mo J [13]	2010	The Bed Management between June 1, 2006 to June 1, 2007 1) In each class based on length of stay and admissions data 2) Historical previous year's same week admissions data 3) Relationship between identified variables to aid bed managers	The three methods are: 1) Poisson bed occupancy model 2) Simulation model 3) Regression model	The three prediction targets are: 1) Estimation of bed occupancy and optimal bed requirements in each class 2) Bed occupancy levels for every class for the following week 3) Weekly average number of occupied beds
Seematter-Bagnoud, L, et al [14]	2015	Inpatient stays data in 2010 (acute somatic care inpatients & outpatients)	Three models of hypothesis-based statistical forecasting of future trends	1) The number of hospital stays 2) Hospital inpatient days 3) Beds for medical stays
Farmer RD, and Emami J [15]	1990	Inpatient stays data for general surgery in the age group 15-44 years between 1969 to 1982	The two methods are: 1) Forecasting from a structural model 2) The time series or Box-Jenkins' method.	Entire hospital short-term daily bed requirements
Kim K, et al [16]	2014	Data warehouse between January, 2009 to June, 2012	The two methods are: 1) The ARIMA model for univariate data 2) The time-series model for multivariate data	Entire hospital bed occupancy (1 day and 1 week)
Kutafina E, et al [17]	2019	Inpatient stays data between October 14. 2002 to December 31. 2015 (patient identifier, time of admission, discharge, the name of the clinic the patient was admitted to. No personal information on the patients or staff was provided)	NARX model, a type of RNN	Entire hospital mid- term bed occupancy (60 days, bed pool in units of 30 beds)
Bouhamed H, Hamdi M, and Gargouri R [18]	2022	COVID-19 hospital occupancy data in 15 countries between December. 2021 to early January. 2022	The three models are: LSTM, GRU, and SRNN. Incorporate vaccination percentages and median ages of populations data to improve performance	Entire hospital bed occupancy
Bekker R, uit het Broek M, and Koole G	2021	Historical data publicly available until mid-October	The two methods are: 1) Using linear programming to predict admissions	1)Patients admission

[19]		2020	2) Fit remaining LOS and use results from queuing theory to predict occupancy	2) Entire hospital short-term bed occupancy from new and current
Farcomeni A, et al [20]	2021	Patients admitted to ICU between January to June 2020	The two methods are: 1) Generalized linear mixed regression model 2) Area-specific nonstationary integer autoregressive methodology	Entire hospital short-term intensive care bed occupancy

[a]MLR: multinomial logistic regression

[b]LR: linear regression

[c]LOS: length of stay

Although previous research has contributed to BOR prediction and operational planning at the hospital level, more detailed and systematic predictions are necessary for practical application in real-world operations. To address this issue, there have been studies that develop their own computer simulation hospital systems to not only predict bed occupancy but also execute scheduling for admissions and surgeries to enhance resource utilization [65-67]. Nevertheless, the existing studies have limitations of focusing solely on the overall BOR of the hospital. As an advancement to those studies, we aim to propose a strategy for predicting BOR at the level of each ward and room using various variables in a time-series manner. Interestingly, to our knowledge, this is the first study to apply DL to predict ward- and room-specific occupancy rates using time-series analysis.

1.2 Goal of this study

The aim of this study was to predict the BOR of hospital wards and rooms using time-series data from individual beds. Although overall bed occupancy prediction is useful for macro-level resource management in hospitals, resource allocation based on the prediction of occupancy rates for each ward and room is required for specific hospital scheduling and practicality. Through this approach, we aim to contribute to the efficient operational cost optimization of the hospital and ensure the availability of resources required for patient care.

We have developed DNN based time-series prediction models, among which one model combines data representing room-specific features (static data) with dynamic data to enhance prediction performance for room bed occupancy rates (RBORs). This model outperforms other RBOR prediction models, demonstrating a lower MAE of 0.003, a mean squared error (MSE) of 0.002, a root mean squared error (RMSE) of 0.005, and a higher R2 score of 0.074, indicating the highest performance among all RBOR models.

We developed six time-series prediction models, of which two were used for predicting ward bed occupancy rates (WBORs), and the other four focused on predicting RBORs. These models utilize LSTM architectures with strong long-term memory capabilities as their base structures. The WBOR models were used for predicting weekly and monthly occupancy rates, serving long-term hospital administrative planning purposes. Conversely, the RBOR models were designed for immediate and rapid occupancy planning and were trained with 3- and 7-day intervals. Each RBOR model was enhanced by combining static data, which represents room-specific features, to generate more sophisticated prediction models.

Figure 1 shows the potential application of our model as a form of web software in a hospital setting. Through an online dashboard, it can provide timely information regarding bed availability, enabling intelligent management of patient movements due to admissions and discharges. It facilitates shared responsibilities within the hospital and simplifies future resource planning [68].

In the Introduction section, we explored the importance of this research and investigated relevant previous studies, providing a general overview of the direction of our research. In the Methods section, we will provide descriptions of the dataset used and the structure of the DNN algorithm used and explain the model architecture and performance. In the Results section, we will present the performance and outcomes of this study, and finally, we will discuss the contribution, limitations, and potential avenues for improvement of the research.

Figure 1. A virtual dashboard of the status and forecast of WBOR and RBOR. (1) On the first screen, the overall BOR of the hospital is shown, along with the number of beds in use and available. Moreover, a predictive graph displays the anticipated WBOR for selected dates. (2) The second screen presents the WBOR for individual beds, indicating their statuses such as "in use," "reserved," "empty," and "cleaning." Detailed information about each room is also displayed

2. Methods

We intended to predict the occupancy rates of individual hospital wards and rooms based on the information accumulated in individual bed-level data on an hourly basis, aggregated on a daily basis. For this purpose, we developed six time-series models. As the base model, we applied LSTM, which is suitable for sequence data. This model addresses the limitation of long-term memory loss in traditional RNNs and was chosen because of its suitability for training bed data represented as sequence data.

There were two WBOR prediction models, which were trained at 7- and 30-day intervals to predict the occupancy rate for the next day. Moreover, there were four RBOR prediction models, similar to the ward models, using LSTM as the base model and were trained at 3- and 7-day intervals. Furthermore, as another approach, each RBOR prediction model was augmented with static data, and four DL algorithms were proposed for the final comparison of their performances in predicting RBOR.

2.1 Materials

2.1.1 Study setting

This was a retrospective, single-center cohort study. Data were collected from AMC, with information on the occupancy status of each bed recorded at hourly intervals between May 27, 2020, and November 21, 2022. The dataset comprised a total of 54,632,684 records. This study used ethically preapproved data and underwent AMC IRB review (IRB 2021-0321), which was conducted according to the 2008 Declaration of Helsinki. De-identified data used in the study were extracted from ABLE, the AMC clinical research data warehouse.

A total of 57 wards, encompassing specialized wards, 1411 rooms, including private and shared rooms, and 4990 beds, were included in this study. Wards and rooms with specific characteristics such as the intensive care unit (ICU), new-born, and nuclear medicine treatment room were excluded from the analysis as their occupancy prediction using simple and general variables did not align with the direction of this study.

2.1.2 Supporting data

Supporting data for public holidays were added in our dataset. We considered that holidays have both a recurring pattern with specific dates each year and a distinctive characteristic of being nonworking days, which could affect occupancy rates. Based on Korean public holidays, which include "Chuseok, Hangeul Proclamation Day, Children's Day, National Liberation Day, Memorial Day, Buddha's Birthday, Independence Movement Day, and Constitution Day," there were 27 days that corresponded to public holidays during the period covered by the dataset. We denoted these dates with a value of "1" if they were public holidays and "0" if they were not, based on the reference date.

2.1.3 Preprocessing and description of variables

Among the variables representing individual beds, the reference date, ward and room information, patient's occupancy status, bed cleanliness status, and detailed room information were available. Based on the recorded date of bed status, we derived additional variables such as the reference year, reference month, reference week (week of the year), reference day, and reference day of the week.

Room data were derived from the input information representing the cleanliness status of beds. This variable has two possible states, viz., "admittable" or "discharge." If neither of these states was indicated, it implied that a patient was currently hospitalized in the bed. As the status of hospitalized patients was indicated by missing values, we replaced them with the number “1” to indicate the presence of a patient in the bed, and “0” otherwise. The sum of all “1” values represents the current number of hospitalized patients. The count of beds in each room indicates the capacity of each room. The target variable BOR was calculated by dividing the number of patients in the room by the room capacity, resulting in a room-specific patient occupancy rate variable. The ward data were subjected to a similar process as that of the room data, with the difference being that we generated ward-specific variables such as ward capacity and WBOR using the same approach. The static room data consisted of 14 variables, including the title of the room and the detailed information specific to each room.

For the variables in the ward and room data, we disregarded the units of the features and converted them into numerical values for easy comparison, after which we performed normalization. Regarding the variables representing detailed room information, we converted them to numerical values where “Yes” was represented as “1,” and “No” was represented as “0.”

The final set of variables used in this study was categorized into date, ward, room, and detailed room information. Table 2 shows the detailed descriptions of the variables used in our training, including all the administrative data related to beds that are readily available in the hospital.

Table 2. Description of variables by category

	Variable	Type	Description
Date			
	year	3 categories	The reference year for bed status
	month	12 categories	The reference month for bed status
	week	53 categories	The reference week for bed status
	day	31 categories	The reference date for bed status
	weekday	7 categories	The reference day of the week for bed status
	holiday	2 categories	Holiday status
Ward			
	ward abbreviation	57 categories	Abbreviations for entire ward names
	ward capacity	numeric	Number of available ward beds
	ward bed capacity	numeric	Number of patients currently admitted to the ward
	ward occupancy rate	numeric	Ward bed capacity divided by ward capacity
Room			
	room abbreviation	1411 categories	Abbreviations for entire room names
	room capacity	numeric	Number of available room beds
	room bed capacity	numeric	Number of patients currently admitted to the room
	room occupancy rate	numeric	Room bed capacity divided by room capacity
Room static feature			
	room code	34 categories	Room grade code
	nuclear	2 categories (N/Y)	Nuclear medicine room availability
	sterile	2 categories (N/Y)	Sterile room availability
	isolation	2 categories (N/Y)	Isolation room availability
	EEG testing	2 categories (N/Y)	EEG testing room availability
	observation	2 categories (N/Y)	Observation room availability
	kidney	2 categories (N/Y)	Kidney transplant room availability
	liver	2 categories (N/Y)	Liver transplant room availability
	sub-ICU	2 categories (N/Y)	Sub-ICU room availability
	special	2 categories (N/Y)	Special room availability
	small single	2 categories (N/Y)	Small single room availability
	short-term	2 categories (N/Y)	Short-term room availability
	psy-double	2 categories (N/Y)	Psychiatry department double room availability
	psy-open	2 categories (N/Y)	Psychiatry department open room availability

The explanation of the classification for generating the datasets for training each model is provided in Table 3. The static features of the detailed room information were combined with the room dataset, which has sequence characteristics, to generate a separate dataset termed Room+Static.

Table 3. Dataset classification and included variables

Dataset Name	Variables
Ward dataset	ward abbreviation, year, month, week, day, weekday, holiday, ward capacity, ward bed capacity, ward occupancy rate
Room dataset	room abbreviation, year, month, week, day, weekday, holiday, room capacity, room bed capacity, room occupancy rate
Static dataset	14 static variables related to detailed room information
Room+Static dataset	room abbreviation, year, month, week, day, weekday, holiday, room capacity, room bed capacity, 14 static variables related to detailed room information, room occupancy rate

2.1.4 Separation

Each dataset was split into training, validation, and test sets for training and evaluation of the model. The training set consisted of 32,153 rows (67.8%) with data from May 27, 2020, to December 2021. The validation set, used for parameter tuning, included 7085 rows (15.0%) with data from January to June 2022. Finally, the test set comprised 8208 rows (17.2%) with data from July 2022 to November 21, 2022.

2.2 DL algorithms

We used various DL algorithms for in-depth learning. In the following subsections, we shall provide explanations for each model algorithm used in our research.

2.2.1 LSTM network

The RNN [69] is a simple algorithm that passes information from previous steps to the current step, allowing it to iterate and process sequential data. However, it encounters difficulties in handling long-term dependencies, such as those found in time-series data, due to the vanishing gradient problem. To address this issue, LSTM [70] was developed. LSTM excels in handling sequence data and is commonly used in natural language processing, machine translation, and time-series data analysis. LSTM consists of input gate, output gate, and forget gate. And the 'cell state,' is carefully controlled by each gate to determine whether the memory should be retained or forgotten for the next time step.

2.2.2 Bidirectional LSTM network

Although RNN and LSTM possess the ability to remember previous data, they have a limitation in that their results are primarily based on the immediate past patterns because the input is processed in a sequential order. This limitation can be overcome through a network architecture known as bidirectional long short-term memory (BiLSTM) [71]. BiLSTM allows end-to-end learning, minimizing the loss on the output and simultaneously training all parameters. It also has the advantage of performing well even with long data sequences. Because of its suitability for models that require knowledge of dependencies from both the past and future, such as LSTM-based time-series prediction, we selected BiLSTM as the base model.

2.2.3 Attention mechanism

The attention mechanism [72-73] refers to the process of incorporating the encoder's outputs into the decoder at each time step of predicting the output sequence. Rather than considering the entire input sequence, it focuses more on the relevant components that are related to the predicted output, allowing the model to focus on important areas. This mechanism helps minimize information loss in datasets with long sequences, enabling better learning and improving the model's performance. It has been widely utilized in areas such as text translation and speech recognition. Nevertheless, as it is still based on RNN models, it has the drawbacks of slower speed and not being completely free from information loss issues.

2.2.4 Combining static and dynamic features

Data can exhibit different characteristics even at the same time. For instance, in data collected at 1-h intervals for each hospital bed, we can distinguish between "dynamic data," which include features that change over time such as the bed condition, date, and patient occupancy, and "static data," which consists of information that remains constant, such as the ward and room number.

DL allows us to utilize all the available information for prediction. Therefore, for predicting RBOR, we investigated an approach that combines dynamic and static data using an LSTM-based method [74]. This approach demonstrated better performance than LSTM alone [75]. Our approach involves adding a layer that incorporates static data as an input to the existing room occupancy prediction model.

2.3 Model architecture

2.3.1 Base model

Our objective was to predict the intermediate-term occupancy rates of wards and rooms within the hospital to contribute to hospital operations planning. Among the six DL models we developed, four used a base model that incorporates BiLSTM as the fundamental structure. BiLSTM was chosen as the base model because it has demonstrated improved prediction performance compared with that of traditional LSTM models. BiLSTM models process input sequences bidirectionally, allowing for better information collection and representation.

Moreover, we have enhanced the performance of our models by adding an attention layer to BiLSTM. The attention layer assigns higher weights to features that exert a significant impact on the prediction, allowing the model to focus on relevant information and gather necessary input features. This helps improve the accuracy of the prediction. Furthermore, the attention layer reduces the amount of information processed, resulting in improved computational efficiency. Ultimately, this contributes toward enhancing the overall performance of the model.

The window length of the input sequence was divided into three different intervals, viz., 3, 7, and 30 days. The WBOR model was trained on sequences with a window length of 7 and 30 days, whereas the RBOR model was trained on sequences with a window length of 3 and 7 days. The first layer of our model consisted of BiLSTM, which was followed by the LeakyReLU activation function. LeakyReLU is a linear function that has a small gradient for negative input values, similar to ReLU. It helps the model converge faster. After applying this process once again, the AttentionWithContext layer was applied, which focuses on important components of input sequence data and transforms outputs obtained from the previous layer. After applying the activation function again, a dense layer with one neuron was added for generating the final output. The sigmoid function was used to limit the output values between 0 and 1. Finally, our model was compiled using the MSE loss function, Adam optimizer, and MAE metric. The parameters for each layer were selected based on accumulated experience through research. Figure 2 visually represents the above-described structure.

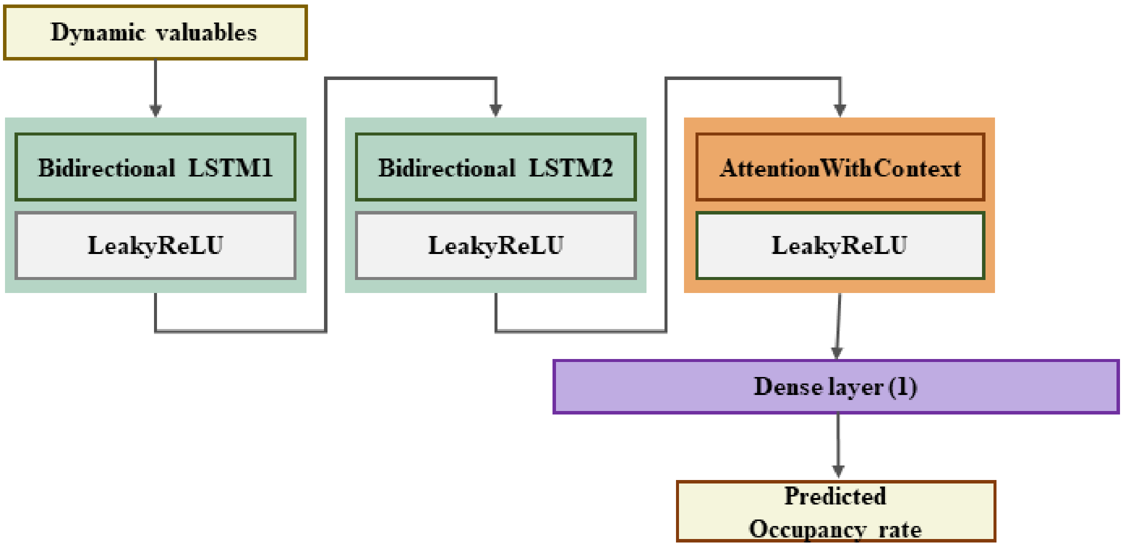

Figure 2. Base model architecture

2.3.2 Combining dynamic and static data using DL model

The accumulated bed data, which were collected on a time basis, were divided into dynamic and static data of the rooms, which were then inputted separately. To improve the performance of the BOR prediction model, we designed different DL architectures for the characteristics of these two types of data.

We first used a base model based on BiLSTM to learn the time-series data and then focused the model's attention using the dense layer to process fixed-size inputs. To prevent overfitting, we applied the dropout function to randomly deactivate neurons in two dense layers. The hidden states of the two networks were combined, and the resulting output was passed to a single layer, combining the time dynamic and static data.

Finally, the hidden states of the two networks were combined, and the combined result was passed to a single layer to effectively integrate the dynamic and static data. This allowed us to utilize the information from both the dynamic and static data for BOR prediction. This architecture is illustrated in Figure 3.

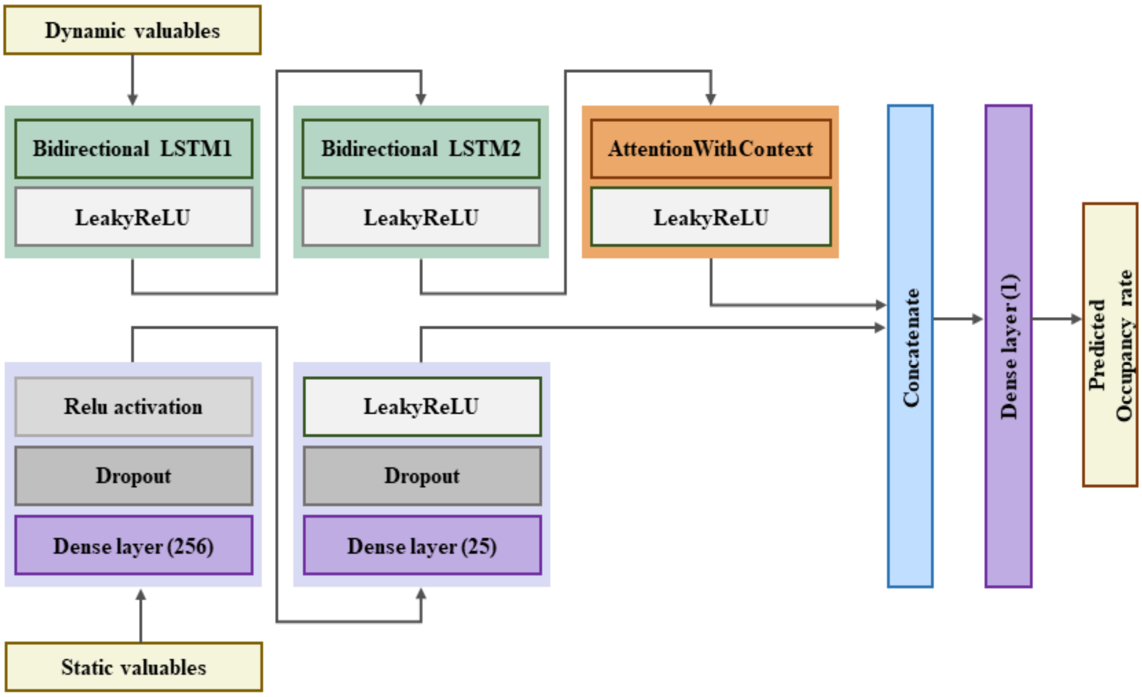

Figure 3. Model architecture combining static and dynamic variables

2.4 Evaluation

We selected various metrics to evaluate the performance of time-series data predictions. Among them, MAE represents the absolute difference between the model's predicted values and the actual BOR. We also considered MSE, which is sensitive to outliers. Moreover, to address the limitations of MSE and provide a penalty for large errors, we opted for RMSE. We also used the R2 score to measure the correlation between the predicted and actual values.

MAE is a commonly used metric to evaluate the performance of time-series prediction models. MAE is intuitive and easy to calculate, making it widely used in practice. Because MAE uses absolute values, it is less sensitive to outliers in the occupancy rate values for specific dates. MAE is calculated using the following formula:

$$e_i = y_i - ypred_i$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i|$$

MSE is a metric that evaluates the magnitude of errors by squaring the differences between the predicted and actual values and then taking the average. It is calculated using the following formula:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (|e_i|)^2$$

RMSE is used to address the limitations of MSE where the error scales as a square, providing a more intuitive understanding of the error magnitude between the predicted and actual values. It penalizes large errors, making it less sensitive to outliers. RMSE is calculated using the following formula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (|e_i|)^2}$$

The R2 score is used to measure the explanatory potential of the prediction model, and it is calculated using the following formula:

$$R^2 score = 1 - \frac{SSR}{SST}$$

$$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$SSR = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Here, SSR represents the sum of squared differences between the predicted and actual values, and SST represents the sum of squared differences between the actual values and the mean value of actual values.

3. Results

We compared the performance of six different DL time-series prediction models, which are termed Ward 7 Days (W7D), Ward 30 Days (W30D), Room 3 Days (R3D), Room 7 Days (R7D), Room Static 3 Days (RS3D), and Room Static 7 Days (RS7D). They were evaluated using three time periods: train, validation, and test.

The prediction performance of WBOR and RBOR was compared, which showed that they were more accurate at predicting WBOR, with MAE values of 0.06–0.07. The W7D model, which used 7 days of ward data to predict the next day's ward occupancy, had an MAE of 0.057, an MSE of 0.007, and an RMSE of 0.082, showing high accuracy. The R2 score was also 0.582, which was approximately 0.124 higher than that (0.458) of the W30D model, indicating that the variables in that model explained occupancy reasonably well.

We next compared the performance of the four models for RBOR prediction, among them, the RS7D model, which was trained on a 7-day time step by integrating static and dynamic data, showed the best performance. It achieved an MAE of 0.123, an MSE of 0.052, an RMSE of 0.226, and an R2 score of 0.320. In particular, the R2 score outperformed the R3D model by 0.026. These data are summarized in Table 4. Regarding the WBOR prediction model, the model with a shorter training unit, W7D, demonstrated better performance. However, regarding the RBOR prediction model, the model with a longer training unit of 7 days, which incorporated detailed room-specific information, exhibited slightly higher performance than the model with a shorter 3-day training unit. The model with the added room-specific information still demonstrated superior performance overall.

Table 4. MAE, MSE, RMSE, and R2 score of occupancy prediction models

	Model	MAE	MSE	RMSE	R2 Score
Ward					
	Ward 7 Days	**0.057**	**0.007**	**0.082**	**0.582**
	Ward 30 Days	0.062	0.009	0.093	0.458
Room					
	Room 3 Days	0.126	0.053	0.231	0.294
	Room 7 Days	0.123	0.052	0.227	0.317
	Room Static 3 Days	0.124	0.057	0.239	0.246
	Room Static 7 Days	**0.123**	**0.051**	**0.226**	**0.320**

We visualized the predicted and actual occupancy for each model and investigated the occupancy trends since July 2022 on our test dataset. First, we selected a specific ward in W7D to demonstrate the change in WBOR over 2 months. The right panel of Figure 4 shows the WBOR change over 5 months from July 2022 in W30D. The blue line represents the actual occupancy value, and the red line represents the predicted occupancy value by the model. This provides an at-a-glance view of the overall predicted occupancy level for each month and allows hospital staff to observe trends to obtain a rough understanding of WBOR.

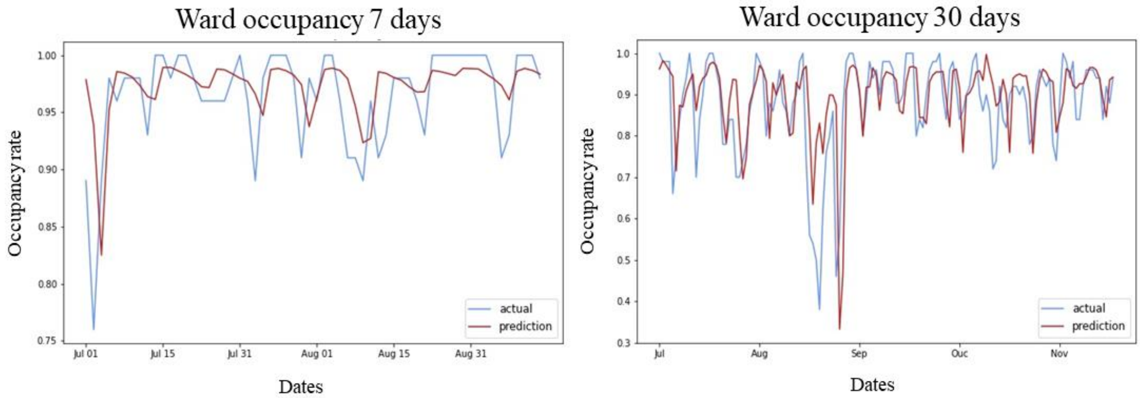

Figure 4. Examples of predicted and actual BOR for the 2 months of July through August 2022 for W7D and the 5 months of July through November 2022 for W30D

Figure 5 shows the graph of occupancy rate values for a randomized specific room, displaying the predicted and actual values for the four RBOR prediction models, with two graphs for each model. The left graph shows the occupancy rate changes over 5 months from July to November 2022, and the right graph shows the occupancy rate for the months of July and August, providing a detailed view of RBOR. By examining the trends of the predicted and actual values for the four models in this period for a specific room, we can observe that the models maintain a similar trend to the actual occupancy rate.

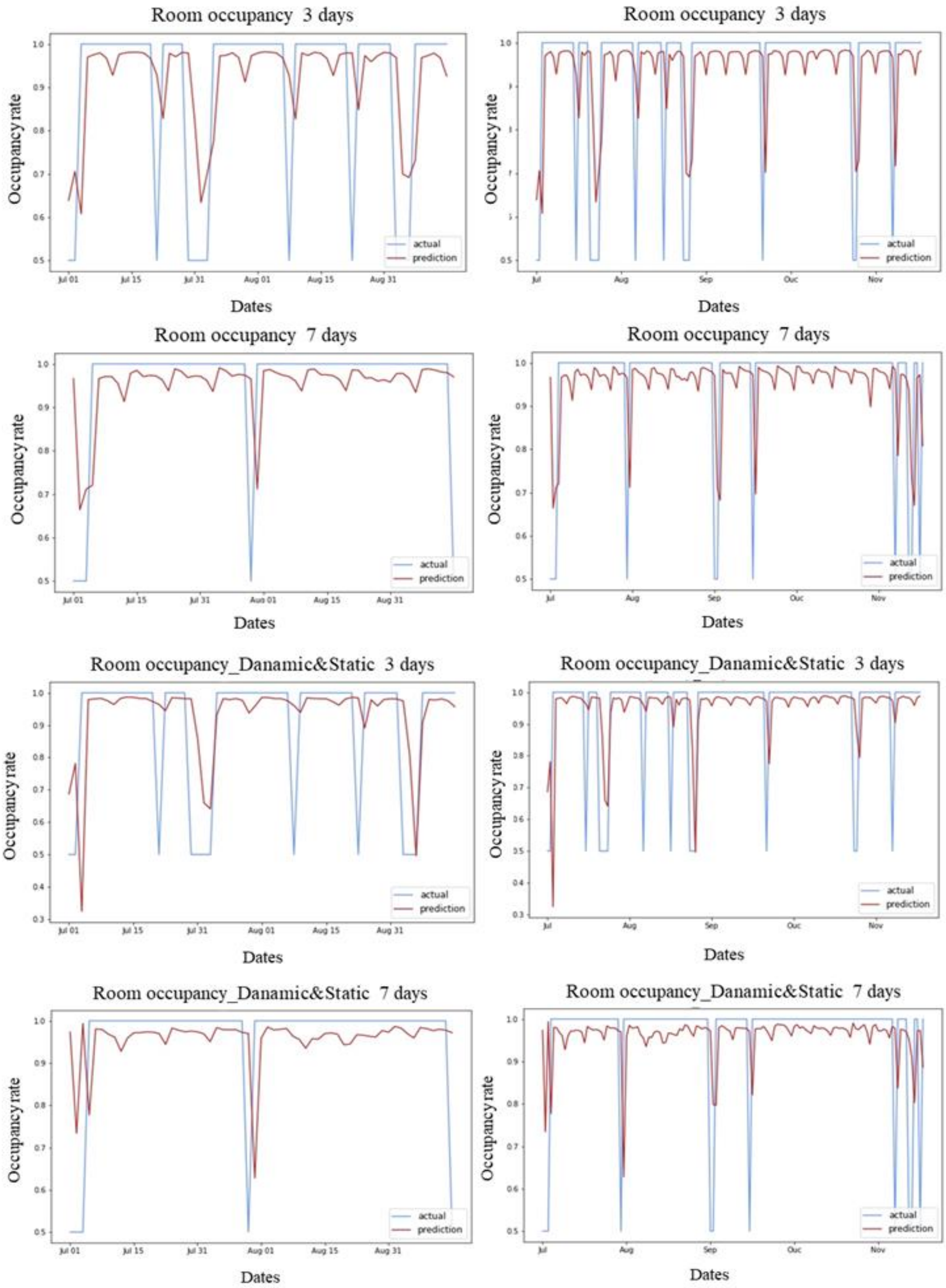

Figure 5. Example of predicted and actual BOR for the 2-month period between July and August 2022 and the 5-month period from July to November.

4. Discussions

The entire dataset of this study consisted of administrative data collected at AMC at an hourly interval for each ward from May 27, 2020, to November 21, 2022. To improve the hospital's challenges, we developed a model to predict the occupancy rate of wards and rooms. Our aim was to contribute toward administrative and financial planning for bed management within the hospital.

During the specified period, we compared the results of using DL models to predict the overall BOR for each ward and individual rooms. In the case of WBOR prediction, the MAE of the 7-day window model was approximately 0.057, demonstrating a remarkably close prediction to the occupancy compared to that of the 30-day window model, with a difference of approximately 0.005. Furthermore, the MSE and RMSE were 0.007 and 0.082, respectively, indicating high accuracy in the predictions. Moreover, the R2 score of 0.582 indicated that the model had better explanatory potential than the average. For the individual RBOR prediction, among the four models, the RS7D model performed the best, exhibiting an MAE of approximately 0.03, which was remarkably lower than that (0.123) of the other models. Moreover, the MSE and RMSE were significantly lower than those of the RBOR model, with a difference of 0.002 and 0.013, respectively. The R2 score of 0.320 indicated that it had higher explanatory potential than the RS3D model, being higher by 0.07.

Finally, we visualized the predicted and actual values on a graph for a specific period and observed that each model captured the trend of the actual BOR quite well. Although the models were less accurate in predicting low occupancy periods, they followed the general trend closely. Overall, these findings demonstrate that our DL models effectively predicted BOR for both wards and individual rooms, with certain models demonstrating superior performance in different scenarios.

4.1 Limitations

Although the models in this study demonstrated good performance in following the trends of BOR and achieved good results, there are several limitations to this research.

First, there are limitations in the data. Although we utilized administrative data and detailed room information available from the hospital to enable the models to capture occupancy trends, the relationship between the variables and the model's explanatory potential showed room for improvement, as indicated by the R2 score. To achieve higher prediction accuracy, it would be beneficial to incorporate diverse data sources and real-time updated information.

Second, there is variability in external factors. Hospital BORs are heavily influenced by external environmental factors. Sudden events such as environmental factors and outbreaks of infectious diseases such as COVID-19 can render accurate prediction of bed occupancy challenging [64, 76]. Furthermore, seasonal effects and accidents can increase the number of patients. Sufficient collection of long-term data on these external factors would be necessary, but such uncertainties can reduce the accuracy of predictions.

Despite these limitations, our study has demonstrated a significant level of adherence to trends in predictions of individual ward and room occupancy. More detailed variables and a longer period of data accumulation would be required to predict the specific number of beds.

5. Conclusion

We have presented models that can predict the occupancy rates of wards and individual hospital rooms using artificial neural networks based on time-series data. The predicted results of these models demonstrated a high level of accuracy in capturing the future trends of BOR. In particular, we presented four RBOR models with structure and window changes to compare their performance and found that the RS7D model showed the best performance. Our results can be implemented as web application on hospital online dashboards, as depicted in

Figure 1 [68]. In fact, Johns Hopkins University has been applying these methods in their command center to monitor hospital capacity and achieve effectiveness in patient management planning [77].

Furthermore, predicting BORs supports patient admission and discharge planning, helping to alleviate overcrowding in emergency departments and reduce patient waiting times. Staff members can effectively schedule patient admissions and discharges and minimize waiting times by understanding the BOR, providing urgent treatment to emergency patients. Moreover, providing appropriate information to patients waiting in the emergency department can increase patient satisfaction and facilitate efficient transitions to hospital admission [78-79]. By applying artificial intelligence models that combine BOR prediction, which contributes toward reducing emergency department waiting times, with individual patient admission and discharge prediction, hospitals can achieve resource optimization and cost savings, resulting in improved patient satisfaction.

Conclusion

We have developed AI models to improve hospital processes using EMR. We applied this research to predict and mitigate ED overcrowding and hospital bed occupancy, aiming to reduce hospital operating costs. Our models have demonstrated excellent performance and can contribute to optimizing hospital operations. This paper is divided into two parts, with each study conducted in the following order: 1) Predicting the likelihood of hospitalization within 24 hours for ED patients and estimating waiting times, 2) Evaluating the impact of unstructured text data in ED bed notes, and 3) Predicting individual ward and room BORs.

In chapter 1, we developed models to predict the likelihood of hospitalization within 24 hours for ED patients and estimate waiting times, while also examining the impact of unstructured text data. These models were created to assist ED physicians in making rapid decisions and have demonstrated outstanding performance in predicting the likelihood of hospitalization within 24 hours and waiting times for patients in the ED. Furthermore, the utilization of unstructured text data has enhanced the model's performance, and through XAI, we confirmed the significance of unstructured text in ED bed notes.

In chapter 2, we involved a study to predict the BORs of individual hospital wards and rooms. Various datasets were created by combining time-series data related to bed occupancy recorded at hourly intervals with static data for the rooms. Using these datasets, we developed two models for predicting WBORs and four models for predicting RBORs. These models exhibited high performance, with the model that combined both dynamic and static data and predicted on a weekly basis showing the best performance. This further emphasized the significance of static data.

Research aimed at improving hospital processes should produce practical outcomes that can be utilized within the healthcare facility. Therefore, we have not only proposed models that significantly enhance the economic efficiency of hospital and ED operations but have also introduced a practical virtual web application for use. By applying AI models to hospital processes, healthcare professionals and administrators can streamline operations and efficiently allocate limited medical resources, leading to improvements in healthcare services and the potential for cost savings.

References

[1] Reuben, D. B., & Cassel, C. K. Physician stewardship of health care in an era of finite resources. Jama, 2011;306(4): 430-431

[2] Wennberg, J. E., Fisher, E. S., & Skinner, J. S. Geography And The Debate Over Medicare Reform: A reform proposal that addresses some underlying causes of Medicare funding woes: geographic variation and lack of incentive for efficient medical practices. Health affairs, 2002;21(Suppl1): W96-W112

[3] Swensen, S. J., Dilling, J. A., Mc Carty, P. M., Bolton, J. W., & Harper Jr, C. M. The business case for health-care quality improvement. Journal of patient safety, 2013;9(1): 44-52

[4] Arha, G.: Reducing wait time prediction in hospital emergency room: lean analysis using a random forest model, 2017

[5] Parker, C.A., Liu, N., Wu, S.X., Shen, Y., Lam, S.S.W., Ong, M.E.H.: Predicting hospital admission at the emergency department triage: A novel prediction model. The American journal of emergency medicine, 2019;37(8): 1498–1504

[6] Shirakawa, T., Sonoo, T., Ogura, K., Fujimori, R., Hara, K., Goto, T., Hashimoto, H., Takahashi, Y., Naraba, H., Nakamura, K., et al.: Institution-specific machine learning models for prehospital assessment to predict hospital admission: prediction model development study. JMIR medical informatics, 2020;8(10), 20324

[7] Hong, W.S., Haimovich, A.D., Taylor, R.A.: Predicting hospital admission at emergency department triage using machine learning. PloS one, 2018;13(7), 0201016

[8] Sun, Y., Teow, K.L., Heng, B.H., Ooi, C.K., Tay, S.Y.: Real-time prediction of waiting time in the emergency department, using quantile regression. Annals of emergency medicine, 2012;60(3):299–308

[9] Roquette, B.P., Nagano, H., Marujo, E.C., Maiorano, A.C.: Prediction of admission in pediatric emergency department with deep neural networks and triage textual data. Neural Networks, 2020;126:170–177

[10] Zhang, X., Kim, J., Patzer, R.E., Pitts, S.R., Patzer, A., Schrager, J.D.: Prediction of emergency department hospital admission based on natural language processing and neural networks. Methods of information in medicine, 2017;56(05):377–389

[11] Mackay M, and Lee MD. Using Compartmental Models to Predict Hospital Bed Occupancy. Department of Psychology. University of Adelaide 2007;1-18.

[12] Littig, SJ, and Isken MW. Short term hospital occupancy prediction. Health care management science 2007;10:47-66

[13] Kumar A, and Mo J. Models for bed occupancy management of a hospital in Singapore. Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management. IIEOM & cosponsored by INFORMS, 2010;1-6

[14] Seematter-Bagnoud, L, et al. Comparison of different methods to forecast hospital bed needs. European Geriatric Medicine 2015;6.3:262-266

[15] Farmer RD, and Emami J. Models for forecasting hospital bed requirements in the acute sector. Journal of Epidemiology & Community Health 1990;44.4: 307-312

[16] Kim K, et al. Predicting patient volumes in hospital medicine: A comparative study of different time series forecasting methods. Northwestern University, Illinois, USA, Scientific Report 2014;6(1):1-20

[17] Kutafina E, et al. Recursive neural networks in hospital bed occupancy forecasting. BMC medical informatics and decision making 2019;19.1:1-10

[18] Bouhamed H, Hamdi M, and Gargouri R. Covid-19 Patients' Hospital Occupancy Prediction During the Recent Omicron Wave via some Recurrent Deep Learning Architectures. International Journal of Computers Communications & Control 2022;17.3

[19] Bekker R, uit het Broek M, and Koole G. Modeling COVID-19 hospital admissions and occupancy in the Netherlands. European journal of operational research 2023;304.1: 207-218

[20] Farcomeni A, et al. An ensemble approach to short-term forecast of COVID-19 intensive care occupancy in Italian regions. Biometrical Journal 2021;63.3:503-51

[21] Sun, B.C., Hsia, R.Y., Weiss, R.E., Zingmond, D., Liang, L.-J., Han, W., McCreath, H., Asch, S.M.: Effect of emergency department crowding on outcomes of admitted patients. Annals of emergency medicine, 2013;61(6):605–611

[22] Ramirez-Nafarrate, A., Hafizoglu, A.B., Gel, E.S., Fowler, J.W.: Optimal control policies for ambulance diversion. European Journal of Operational Research, 2014;236(1):298–312

[23] Pines, J.M., Iyer, S., Disbot, M., Hollander, J.E., Shofer, F.S., Datner, E.M.: The effect of emergency department crowding on patient satisfaction for admitted patients. Academic Emergency Medicine, 2008;15(9):825–831

[24] Keegan AD. Hospital bed occupancy: more than queuing for a bed. Medical Journal of Australia 2010;193.5: 291-293

[25] Kaier K, Mutters NT, & Frank U. Bed occupancy rates and hospital-acquired infections—should beds be kept empty?. Clinical Microbiology and Infection 2012;18.10: 941-945

[26] Morganti, K.G., Bauhoff, S., Blanchard, J.C., Abir, M., Iyer, N., Smith, A., Vesely, J.V., Okeke, E.N., Kellermann, A.L.: The evolving role of emergency departments in the united states. Rand health quarterly, 2013;3(2)

[27] Asplin, B.R., Magid, D.J., Rhodes, K.V., Solberg, .I., Lurie, N., Camargo Jr, C.A.: A conceptual model of emergency department rowding. Annals of emergency medicine, 2013;42(2): 173–180

[28] Kim, K.W., Kim, S.H., Park, K.N., Kim, H.J., Oh, S.H., Lee, J.Y., Lee, J.M., Im, J.Y.: Mid-term effects of tertiary hospital beds expansion on emergency department overcrowding. Journal of the Korean Society of Emergency Medicine, 2014;25(6):722–729

[29] Fatovich, D.M., Nagree, Y., Sprivulis, P.: Access block causes emergency department overcrowding and ambulance diversion in perth, western australia. Emergency medicine journal, 2005;22(5):351–354

[30] Sei-Chang, O.: Examination of the emergency medical response system in korea and suggestions for improvements relating to transport. IATSS Research, 2004;28(2):32–40

[31] Kim, S.J., Kim, H., Park, Y.H., Kang, C.Y., Ro, Y.S., Kim, O.H.: Analysis of the impact of the coronavirus disease epidemic on the emergency medical system in south korea using the korean triage and acuity scale. Yonsei Medical Journal, 2021;62(7):631

[32] Lateef, F.: Patient expectations and the paradigm shift of care in emergency medicine. Journal of emergencies, trauma and shock, 2011;4(2):163

[33] Meystre, S., Haug, P.J.: Automation of a problem list using natural language processing. BMC medical informatics and decision making, 2005;5(1):1–14

[34] Martinez, D.A., Kane, E.M., Jalalpour, M., Scheulen, J., Rupani, H., Toteja, R., Barbara, C., Bush, B., Levin, S.R.: An electronic dashboard to monitor patient flow at the johns hopkins hospital: communication of key performance indicators using the Donabedian model. Journal of medical systems, 2018;42(8):1–8

[35] Hsia, R.Y., Asch, S.M., Weiss, R.E., Zingmond, D., Liang, L.-J., Han, W., McCreath, H., Sun, B.C.: Hospital determinants of emergency department left without being seen rates. Annals of emergency medicine, 2011;58(1):24–32

[36] Kane, E.M., Scheulen, J.J., P¨uttgen, A., Martinez, D., Levin, S., Bush, B.A., Huffman, L., Jacobs, M.M., Rupani, H., Efron, D.T.: Use of systems engineering to design a hospital command center. The Joint Commission Journal on Quality and Patient Safety, 2019;45(5):370–379

[37] Boudreaux, E.D., O'Hea, E.L.: Patient satisfaction in the emergency department: a review of the literature and implications for practice. The Journal of emergency medicine,2004;26(1):13–26

[38] Welch, S.J.: Twenty years of patient satisfaction research applied to the emergency department: a qualitative review. American Journal of Medical Quality, 2010;25(1):64–72

[39] Gartner, D., Padman, R.: Machine learning for healthcare behavioural or: Addressing waiting time perceptions in emergency care. Journal of the Operational Research Society,2020;71(7):1087 1101

[40] Janke, A.T., Overbeek, D.L., Kocher, K.E., Levy, P.D.: Exploring the potential of predictive analytics and big data in emergency care. Annals of emergency medicine, 2016;67(2):227–236

[41] Newman-Toker, D.E., McDonald, K.M., Meltzer, D.O.: How much diagnostic safety can we afford, and how should we decide? a health economics perspective. BMJ quality & safety 22(Suppl 2), 2013;11–20

[42] Feldman, R., Sanger, J., et al.: The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge university press, New York, USA, 2007

[43] Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of statistics, 2001;1189–1232

[44] CHEN, T., GUESTRIN, C.: Xgboost: A scalable tree boosting system; proceedings of the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. F, 2016

[45] Breiman, L.: Random forests. Machine learning, 2001;45(1):5–32

[46] Battiti, R.: First-and second-order methods for learning: between steepest descent and newton's method. Neural computation, 1992;4(2):141–166

[47] Dreiseitl, S., Ohno-Machado, L.: Logistic regression and artificial neural network classification

models: a methodology review. Journal of biomedical informatics, 2002;35(5-6):352–359

[48] Saritas, M.M., Yasar, A.: Performance analysis of ann and naive bayes classification algorithm

for data classification. International journal of intelligent systems and applications in engineering, 2019;7(2):88–91

[49] Montgomery, D.C., Peck, E.A., Vining, G.G.: Introduction to Linear Regression Analysis. John Wiley & Sons, Hoboken, USA, 2021

[50] Kohavi, R., et al.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai, 1995;vol. 14:pp. 1137–1145. Montreal, Canada

[51] Adadi, A., Berrada, M.: Peeking inside the black-box: a survey on explainable artificial intelligence (xai). IEEE access 6, 2018;52138–52160

[52] Vincent, J.-L., Creteur, J.: Ethical aspects of the covid-19 crisis: How to deal with an overwhelming shortage of acute beds. European Heart Journal: Acute Cardiovascular Care, 2020; 9(3):248–252

[53] Song, Y.J.: The south korean health care system. Jmaj, 2009;52(3):206–209

[54] Park, K., Park, J., Kwon, Y.D., Kang, Y., Noh, J.-W.: Public satisfaction with the healthcare system performance in south korea: Universal healthcare system. Health Policy, 2016;120(6):621–629

[55] Organization, W.H., et al.: World health statistics, 2020

[56] Katz, K., Larson, B., Larson, R.: Prescription for the waiting-in-line blues entertain, enlighten, and engage. Oper Manag Crit Perspect Bus Manag 2, 2003;160

[57] Hui, M.K., Zhou, L.: How does waiting duration information influence customers' reactions to waiting for services? 1. Journal of Applied Social Psychology, 1996;26(19):1702–1717

[58] Dub´e, L., Schmitt, B.H.: The temporal dimension of social episodes: Position effect in time judgments of unfilled intervals. Journal of Applied Social Psychology. 1996;26(20):1816–1826

[59] Reuben DB, Cassel CK. Physician stewardship of health care in an era of finite resources. JAMA 2011;306:430–431

[60] CMS, D. National Health Expenditure Projections 2011-2021. *Centers for Medicare and Medicaid Services* 2012;1-23.

[61] World Health Organization. The world health report 2000: health systems: improving performance. World Health Organization; 2000

[62] Kabene SM, et al. The importance of human resources management in health care: a global context. Human resources for health 2006;4.1:1-17

[63] Page K, Barnett AG, & Graves N. What is a hospital bed day worth? A contingent valuation study of hospital Chief Executive Officers. BMC health services research 2017;17: 1-8

[64] Rosenbaum L. Facing Covid-19 in Italy—ethics, logistics, and therapeutics on the epidemic's front line. New England Journal of Medicine 2020;382.20:1873-1875

[65] Caro JJ, et al. Predicting hospital resource use during COVID-19 surges: a simple but flexible discretely integrated condition event simulation of individual patient-hospital trajectories. Value in Health 2021;24.11:1570-1577

[66] Schmidt R, Geisler S, and Spreckelsen C. Decision support for hospital bed management using adaptable individual length of stay estimations and shared resources. BMC medical informatics and decision making 2013;13.1: 1-19

[67] Hancock WM, and Walter PF. The use of computer simulation to develop hospital systems. ACM SIGSIM Simulation Digest 1979;10.4:28-32

[68] Shahpori R, et al. An on-line dashboard to facilitate monitoring of provincial ICU bed occupancy in Alberta, Canada. Journal of Hospital Administration 2014;3.1: 47-54

[69] Rumelhart DE, Hinton GE, and Williams RJ. Learning representations by back-propagating

errors. nature 1986;323.6088:533-536

[70] Hochreiter S, and Schmidhuber J. Long short-term memory. Neural computation 1997;9.8:1735-1780

[71] Schuster M, and Paliwal KK. Bidirectional recurrent neural networks. IEEE transactions on Signal Processing 1997;45.11:2673-2681

[72] Bahdanau D, Cho K, and Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv 2014;1409.0473

[73] Luong, MT, Pham H, and Manning CD. Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025 2015;

[74] Leontjeva A, and Kuzovkin I. Combining static and dynamic features for multivariate sequence classification. 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE 2016;

[75] Esteban C, et al. Predicting clinical events by combining static and dynamic information using recurrent neural networks. 2016 IEEE International Conference on Healthcare Informatics (ICHI). Ieee 2016;93-101

[76] Vincent JL, and Creteur J. Ethical aspects of the COVID-19 crisis: How to deal with an overwhelming shortage of acute beds. European Heart Journal: Acute Cardiovascular Care 2020;9.3:248-252

[77] Martinez DA, et al. An electronic dashboard to monitor patient flow at the Johns Hopkins Hospital: communication of key performance indicators using the Donabedian model. *Journal of medical systems 2018;*42:1-8

[78] Gartner D, and Padman R. Machine learning for healthcare behavioural OR: Addressing waiting time perceptions in emergency care. Journal of the Operational Research Society 2020;71.7:1087-1101

[79] Welch SJ. Twenty years of patient satisfaction research applied to the emergency department: a qualitative review. American Journal of Medical Quality 2010;25.1:64-72

국문 요약

전 세계적으로 헬스케어 시장 규모가 확대되고 의료 수준이 향상됨에 따라 의료 비용이 증가되고 있습니다. 동시에 제한된 예산을 가진 국가와 개인의 의료 비용 부담은 증가하고 있습니다. 이 문제를 해결하기 위해 의료비 낭비를 줄이고 제어하기 위한 다양한 접근 방법이 제안되고 있습니다. 그 중에서도 병원 프로세스 개선은 환자에게 커다란 혜택을 주는 서비스와 치료를 제공하는 중요한 방법 중 하나입니다. 특히 프로세스의 최적화는 병원과 환자 모두에게 비효율성의 낭비를 줄이는 효과적인 방법 중 하나로 여겨집니다.

최근 인공지능을 의료 분야에 적용하는 연구가 활발히 진행되고 있으며, 이를 통해 의료 자원의 한정성을 극복하려는 노력이 늘어가고 있습니다. 다양한 방법 가운데 응급실 과밀화 완화 및 병원의 병상 점유율을 예측하는 방법이 특히 주목받고 있습니다. 응급실의 과밀화는 사망률 상승, 대기 시간 증가, 치료 실수, 진단 및 절차 지연과 같은 문제를 야기할 수 있습니다. 또한, 높은 병상 점유율은 의료 직원의 건강에 부정적인 영향을 미치고 감염 위험도를 증가시킬 수 있습니다.

본 연구에서는 전자의무기록(EMR)을 활용하여 병원 프로세스를 개선하는 인공지능 모델을 개발했습니다. 우리 모델은 뛰어난 성능을 보이며 병원 운영 최적화에 기여할 수 있다는 것을 입증했습니다. 본 논문은 두 부분으로 구성되며, 각 연구는 다음과 같습니다. 1) 응급실 경유 환자의 24시간 이내 입원 가능성 및 대기 시간 예측, 2) 응급실 병상 노트의 비구조적 텍스트 데이터 영향 평가, 3) 개별 병동 및 병실 이용률 예측의 순서대로 진행되었습니다.

1장에서는 응급실 경유 환자의 24시간 이내 입원 가능성과 예상 대기 시간을 예측하는 모델을 만들고, 비정형 텍스트 데이터의 영향을 확인했습니다. 이 모델은 응급실 의사가 신속한 의사결정을 지원하기 위해 개발되었으며, 뛰어난 성능으로 응급실 경유 환자의 24시간 이내 입원 가능성과 대기 시간을 예측합니다. 더불어, 비정형 텍스트 데이터의 활용을 통해 모델의 성능을 향상시키고, XAI를 통해 변수 영향력을 확인함으로써 응급실 병상 노트의 비정형 텍스트가 중요한 역할을 한다는 사실을 입증했습니다.

2장에서는 개별 병동 및 병실의 병상 점유율을 예측하는 연구가 진행되었습니다. 1시간 간격으로 기록된 병상과 관련된 시계열 데이터와 병실의 정적 데이터를 결합하여 다양한 데이터셋을 구성하였으며, 이를 활용하여 2개의 병동 병상 이용률 예측 모델과 4개의 병실 병상 이용률 예측 모델을 개발하였습니다. 이러한 모델은 높은 정확성을 보이며, 특히 병실의 동적 및 정적 데이터를 결합하여 일주일 간격으로 예측한 모델이 가장 우수한 결과를 보여주었습니다. 이를 통해 정적 데이터의 중요성도 보여주었습니다.

병원 프로세스 개선을 위한 연구는 실질적인 결과물을 병원에서 활용 가능하도록 생산해야 합니다. 따라서 우리는 병원 및 응급실 운영의 경제적 효율성을 크게 향상시키는데 도움이 되는 모델을 제안하기 위해 실질적으로 활용 가능한 가상 웹 애플리케이션도 제안했습니다. 인공지능 모델을 병원 프로세스에 적용하면 의료진과 병원 관리자가 프로세스를 간소화하고 제한된 의료 자원을 효율적으로 활용할 수 있으므로 의료 서비스의 향상과 함께 비용 절감 효과를 기대할 수 있습니다.

중심 단어: EMR, Emergency department, Artificial intelligence, Natural language processing, Time series forecasting, Combining static and dynamic variables