Military Caring Sciences
Volume 11, Issue 2, Summer 2024 (Serial Number: 40)
Original Article

Predicting the Length of Stay of Patients in the Emergency Department of the Hospital using Machine Learning Models

Moosavi Kashani. S[1]
*Zargar Balaye Jame. S[2]

1- MSc in Health Systems Engineering, School of Industrial and Systems, Tarbiat Modares University, Tehran, Iran

2- (*Corresponding Author) Ph.D. in Health Care Management, Associate Professor, Department of Management Sciences and Health Economics, School of Medicine, Aja University of Medical Sciences, Tehran, Iran , Email: sanazzargar@gmail.com

Abstract

Introduction: The length of patient stay in the emergency department is crucial for optimizing resource allocation, reducing costs, and enhancing operations.

Objective: This study aimed to predict patients' length of stay in the emergency department using machine learning models.

Material and Methods: This retrospective cohort study collected data from patients referred to the emergency room of a selected hospital in Tehran, including vital signs, diagnoses, and demographic information such as age and gender, during December 2022. After data preparation, ensemble models—Random Forest, Light GBM, Cat Boost, and Ada Boost—were employed to predict patients' length of stay.

Results: The study found that vital signs, age, and Emergency Severity Index level 1 significantly influence patient length of stay. The Cat Boost model, with an accuracy of 0.87, precision of 0.91, recall of 0.83, and F1-score of 0.87, outperformed other models in predictive performance.

Conclusion: This study demonstrated that ensemble models effectively predict emergency room patient length of stay, with Boosting methods outperforming Bagging methods.

Keywords: Emergency Service, Machine Learning, Length of Stay, Prediction

Moosavi Kashani. S , Zargar Balaye Jame. S. Predicting the Length of Stay of Patients in the Emergency Department of the Hospital using Machine Learning Models. Military Caring Sciences. 2024; 11 (2). 114-123.
Received: 31/07/2023 Received in revised form: 11/09/2023 Accepted: 12/09/2023 Published: 12/09/2024

پیش‌بینی طول مدت اقامت بیماران در اورژانس بیمارستان با استفاده از مدل‌های یادگیری ماشین

سینا موسوی کاشانی[1]، *ساناز زرگر بالای جمع[2]

چکیده

مقدمه: طول مدت اقامت بیماران در اورژانس بیمارســـتان، ابزار مهمی در جهت تخصیص بهینه‌ی منابع به بیماران بخش اورژانس، بهینه سازی عملیات و کاهش هزینه‌ها محسوب می‌شود.

هدف: مطالعه‌ی حاضر با هدف پیش‌بینی طول مدت اقامت بیماران در بخش اورژانس با استفاده از مدل‌های یادگیری ماشین انجام شد.

مواد و روش‌ها: این مطالعه از نوع کوهورت گذشـــته نگر اســـت. در این مطالعه (کوهورت گذشته نگر)، اطلاعات بیماران مراجعه کننده به اورژانس یکی از بیمارستان‌های منتخب شهر تهران که متشکل از علائم حیاتی، تشخیص و اطلاعات دموگرافیک مانند سن و جنس بود، در بازه‌ی زمانی یک ماهه‌ی آذر ۱۴۰۱ جمع آوری گردید. در ادامه پس از آماده ســـازی داده‌ها از مدل‌های Random Forest- Light GBM- Cat Boost- Ada Boost که از نوع مجموعه‌ای (Ensemble) هستند برای پیش‌بینی طول مدت اقامت بیماران مورد استفاده قرار گرفتند.

یافته‌ها: نتایج نشـــان دهنده‌ی آن اســـت که علائم حیاتی، سن و شاخص شـــدت اضطراری سطح ۱ از جمله عواملی هستند که بیشترین تأثیر را در طول مدت اقامت داشـــتند، همچنین مدل Cat Boost با معیارهای (Accuracy=۰/۸۷، Precision=۰/۹۱، Recall=۰/۸۳، F1-score= ۰/۸۷) بهترین عملکرد در میان دیگر مدل‌ها را داشت.

نتیجه‌گیری: این مطالعه نشـــان داد که مدل‌هـــای مجموعه‌ای عملکرد خوبی در پیش‌بینی طول مدت اقامـــت بیماران در اورژانس دارد همچنین در این پژوهش مدل‌هایی که از روش Boosting پیروی می‌کنند در مجموع نسبت به Bagging عملکرد بهتری دارند.

کلمات کلیدی: بخش اورژانس، پیش بینی، طول مدت اقامت، یادگیری ماشینی

مجله علوم مراقبتی نظامی سال یازدهم ■ شماره۲ ■ تابستان ۱۴۰۳ ■ شماره مسلسل ۴۰ ■ صفحات ۱۲۳ -۱۱۴

تاریخ دریافت: ۱۴۰۲/۰۵/۰۹
تاریخ ویرایش: ۱۴۰۲/۰۶/۲۰
تاریخ پذیرش: ۱۴۰۲/۰۶/۲۱
تاریخ انتشار: ۱۴۰۳/۰۴/۱۰

مقدمه

خدمـــات درمانی به یکی از بزرگترین صنایع در ســـطح جهان تبدیل شده اســـت و بخش اورژانس یکی از بخش‌های اصلی این خدمات را تشکیل می‌دهد و از ضرورت‌های بالایی برخوردار است (۱). بخـــش اورژانس برای ارائه مراقبت‌هـــای اضطراری جامع به جامعه در شرایط اضطراری و غیر اضطراری آماده و مجهز است و درب‌های آن به صورت ۲۴ ساعته و ۳۶۵ روز در سال باز می‌باشد، این بخش دارای عملیاتی منحصر به فرد اســـت و دارای چندین تعامـــل و تراکـــم تصمیم‌گیری بالا که باعـــث وقفه در این بخش است می‌باشد (۲). ازدحام بیش از حد در بخش اورژانس مشکلی رایج در سراسر جهان است که عواقب ناخواسته‌ای مانند از دست دادن منابع، استفاده بی‌رویه از زمان و نارضایتی پرسنل اورژانس و متقاضیـــان را به همـــراه دارد (۳). ارتباط قابل قبولی بین طول مـــدت اقامت بیماران در بخش اورژانس و شـــلوغی در این بخش وجود دارد (۴)، همچنین کاهش مدت اقامت بیماران پتانســـیل قابل توجهی برای کاهش هزینه‌هـــا در بخش اورژانس دارد (۵). طول مدت اقامت (Length of Stay) بیماران در بخش اورژانس به عنوان فاصله‌ی زمانی بین رســـیدن بیمـــار به بخش اورژانس تـــا زمانی که بیمار این بخش را ترک می‌کند گفته می‌شـــود که بازه‌ی آن برای هر بیمار تقریباً از ۴ تا ۴۸ ســـاعت متغیر اســـت

۱- دانشکده صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران
۲- گروه علوم مدیریت و اقتصاد سلامت دانشکده پزشکی، دانشگاه علوم پزشکی آجا، تهران، ایران (* نویسنده مسئول)،
آدرس الکترونیک: sanazzargar@gmail.com

(۶). این شاخص یک نشانگر مؤثر برای عملکرد بخش اورژانس و کیفیت تریاژ (Triage) است (۷). مطالعه‌ی موری (Mowery) و همکاران نشان داد که افزایش طول مدت اقامت بیماران در بخش اورژانس، افزایش مرگ و میر بیماران را به همراه دارد (۸) و در مقایسه با بیمارانی که مدت زمان کوتاه‌تری دارند، بیشتر در معرض بیماری قرار می‌گیرند (۹). امروزه از یادگیری ماشین (Machine Learning) به صورت گسترده در صنایع مختلف از جمله سلامت به منظور پیش بینی به عنوان ابزار تصمیم گیری و کشف الگوها استفاده می‌شود (۱۰)، بنابراین با استفاده از پیش بینی طول مدت اقامت بیماران در بخش اورژانس به عنوان ابزار پشتیبانی تصمیم، می‌توان به بهبود برنامه‌ریزی منابع اورژانسی و بیمارستانی و اطلاع‌رسانی به بیماران کمک کرد (۱۱). مطالعات گوناگونی در خصوص پیش بینی طول مدت اقامت بیماران در اورژانس انجام شده است، برای مثال مطالعه‌ی هیجری و اولاوین (Hijry and Olawoyin) با استفاده از الگوریتم‌های یادگیری ماشین، مانند شبکه‌های عصبی مصنوعی (Artificial Neural Networks)، رگرسیون‌های خطی (Linear Regression) و لجستیک (Logistic) به پیش بینی طول مدت اقامت بیماران پرداختند و اشاره داشتند که با پیش بینی این شاخص به پزشکان و پرستاران این امکان را می‌دهد تا برای سطوح بالای ازدحام آماده شوند (۱۲). نتایج مطالعه‌ی کادری (Kadri) و همکاران حاکی از آن است که مدل‌های یادگیری عمیق نسبت به سایر مدل‌ها برای پیش بینی طول مدت اقامت بیماران در بخش اورژانس برتری دارد (۱۳). همچنین مطالعه‌ی رحمان (Rahman) و همکاران با هدف توسعه و اعتبار سنجی یک مدل پیش‌بینی دقیق برای طول مدت اقامت بیماران بیش از ۴ ساعت در بخش اورژانس با استفاده از تکنیک داده کاوی (Data Mining) انجام شد. در این مطالعه یک الگوریتم درخت تصمیم (Decision Tree) برای پیش‌بینی بیماران با مدت زمان بیشتر از ۴ ساعت ساخته شد. در مجموع ۳۳ ویژگی مورد تجزیه و تحلیل قرار گرفت. نتایج نشان داد که دقت مدل درخت تصمیم ۸۵ درصد بود (۹). بنابراین با توجه به بررسی‌های انجام شده نتایج حاکی از آن است که پیش بینی طول مدت اقامت بیماران در اورژانس بیمارستان، ابزار مهمی در جهت تخصیص بهینه‌ی منابع به بیماران بخش اورژانس، بهینه سازی عملیات و کاهش هزینه‌ها محسوب می‌شود. همچنین مطالعات انجام شده در این حوزه حاکی از آن است که برای پیش بینی طول مدت اقامت بیماران در بخش اورژانس از مدل‌های مجموعه‌ای (Ensemble) که دقت نسبتاً خوبی در پیش بینی دارند کمتر استفاده شده است و عملکرد الگوریتم‌های مدل‌های مجموعه‌ای در این حوزه مقایسه نشده است. بنابراین این پژوهش با هدف پیش بینی طول مدت اقامت بیماران در بخش اورژانس انجام شد و در ادامه عملکرد مدل‌های مجموعه‌ای با یکدیگر مقایسه گردید.

مواد و روش ها

این مطالعه از نوع کوهورت گذشته نگر است. جامعه آماری پژوهش حاضر در برگیرنده کلیه‌ی بیماران مراجعه کننده به اورژانس یکی از بیمارستان‌های منتخب شهر تهران در آذرماه ۱۴۰۱ است. از میان ۱۰۰۰ بیمار مراجعه کننده به بخش اورژانس که بستری شده بودند، اطلاعات ۴۰۰ بیمار مورد بررسی قرار گرفت. اطلاعات جمع آوری شده متشکل از علائم حیاتی، تشخیص بیماری و اطلاعات دموگرافیک مانند سن و جنس بود که از پرونده‌ی بیماران استخراج گردید. برای طبقه‌بندی هر تشخیص بیماری، متخصصان مراقبت‌های بهداشتی از طبقه‌بندی بین المللی بیماری‌ها ویرایش دهم (International Classification of Diseases)، استفاده می‌کنند که یک کد منحصر به فرد به هر تشخیص اختصاص می‌دهد (۱۴). برای مثال به بیماری‌های تنفسی کد J00-J99 تخصیص داده می‌شود.

سایر متغیرهایی که در این پژوهش در نظر گرفته شده است می‌توان به نحوه‌ی ورود بیماران به اورژانس و شاخص شدت اورژانس آنان اشاره کرد. نحوه‌ی ورود بیمار به اورژانس دارای ۲ حالت، با استفاده از آمبولانس و یا با استفاده از ماشین شخصی بود.

شاخص شدت اورژانس (ESI: Emergency Severity Index) که معمولاً در بخش‌های اورژانس برای اولویت بندی بیماران بر اساس شدت وضعیت و سطح مراقبت لازم استفاده می‌شود؛ که بر اساس این شاخص بیماران در بدو ورود به اورژانس به ۵ سطح طبقه بندی می‌شوند. سطح ۱: نیاز به مداخله نجات دهنده فوری دارد، سطح ۲: خطر بالا و نیاز به ارزیابی پزشکی سریع دارد، سطح ۳: نیاز به ارزیابی فوری پزشکی دارد، سطح ۴: شرایط کمتر

فوری، می‌تواند مدت بیشـــتری برای ارزیابی و درمان صبر کند و سطح ۵: شرایط غیر اورژانسی که می‌تواند در یک محیط مراقبت اولیه مدیریت شود. (۱۵).

برای انجام پژوهـــش پیش رو از فرآیند اســـتاندارد صنعت برای داده کاوی (Cross Industry Standard Process for Data Mining) استفاده شد. این فرآیند روشی مستقل از صنعت و یک رویکرد ساختاریافته و سیســـتماتیک برای برنامه ریزی و اجرای یک پروژه داده کاوی ارائه می‌دهد. این فرایند کاربردی‌ترین روش مرجع داده کاوی اســـت که دارای ۶ مرحله درک فضای کسب و کار، درک داده، آماده سازی داده‌ها، مدل‌سازی، ارزیابی و استقرار می‌باشد. (۱۶).

مراحل انجام پژوهش حاضر در شکل یک به صورت کلی به تصویر کشیده شده است.

شکل۱- مراحل انجام پژوهش

در مرحلـــه‌ی آماده ســـازی داده‌ها، متغیر هـــدف که طول مدت اقامـــت بیماران در بخش اورژانس بود بـــه دو کلاس صفر و یک تبدیل شـــد. کلاس صفر معرف طول مدت اقامت بیمارانی بود که یا ۲۴ ســـاعت و یا کمتر از ۲۴ ســـاعت در بخش اورژانس اقامت داشتند، همچنین کلاس یک معرف طول مدت اقامت بیش از ۲۴ ســـاعت بود. همچنین لازم به ذکر است، تصدیق و اعتبار سنجی کلاس‌بندی طول مدت اقامت بیماران توســـط متخصص اورژانس و معاونت درمان بیمارســـتان مورد بررســـی قرار گرفت. داده‌های جمع آوری شده ارزش تکراری نداشتند اما ۲ نمونه از داده دارای ارزش مفقوده (Missing Value) بودند که از مجموعه‌ی داده‌ها حذف شـــدند. عدم تعادل در کلاس‌های صفـــر و یک مربوط به متغیر هدف یک مشـــکل جدی برای مسائل یادگیری ماشین از نوع طبقه‌بندی است، بنابراین با استفاده از الگوریتم (Synthetic Minority Oversampling Technique) که می‌تواند با بیش نمونه‌گیری، نرخ عدم تعادل را بهبود ببخشد (۱۷)، فراوانی مربوط به ۲ کلاس مورد متعادل قرار گرفت. در ادامه ۸۰ درصد از داده‌ها برای آمـــوزش مدل و ۲۰ درصد هم بـــرای ارزیابی مدل در نظر گرفته شـــد. با توجه به آنکه ماهیت مدل‌های اســـتفاده شـــده از جنس درخت تصمیم بود، احتیاجی به نرمال‌سازی داده‌ها وجود نداشت.

در مرحله‌ی مدل‌سازی، با توجه به متغیر خروجی، نوع مدل‌های پژوهـــش حاضر از نوع یادگیری با ناظـــر و به صورت طبقه‌بندی (Classification) اســـت. مدل‌های مجموعه‌ای شـــامل ترکیب چندین مدل برای بهبود دقت و اســـتحکام پیش‌بینی‌ها می‌باشد. روش‌هـــای مجموعه‌ای را می‌توان به دو دســـته اصلی تقســـیم کـــرد: ۱) Bagging ۲)Boosting (۱۸). روش Bagging یک روش مجموعه موازی اســـت که شـــامل آموزش چندین نمونه از یک مدل در زیر مجموعه‌های مختلف داده‌های آموزشـــی است. ســـپس پیش‌بینی‌های این مدل‌ها با استفاده از میانگین‌گیری یا رأی‌گیـــری برای پیش‌بینی نهایی ترکیب می‌شـــوند. این روش با

اســـتفاده از مدل‌های متعدد با واریانس بالا به کاهش بیش از حد بـــرازش کمک می‌کند که در صـــورت میانگین، یکدیگر را خنثی می‌کنند، مانند جنگل تصادفی (Random Forest (RF))، روش Boosting یک روش مجموعه متوالی اســـت که شامل آموزش مدل‌هـــای ضعیف (مدل‌هایی با دقت پایین) به ترتیب اســـت که در آن هـــر مـــدل بعدی از اشـــتباهات مدل قبلی یـــاد می‌گیرد. الگوریتم‌هـــای تقویت‌کننـــده وزن‌های بالاتری را بـــه نمونه‌های طبقه‌بندی اشـــتباه در داده‌های آموزشی اختصاص می‌دهند، به طوری که مدل‌های بعدی بیشـــتر بر اصلاح ایـــن خطاها تمرکز می‌کنند. سپس پیش‌بینی‌های این مدل‌ها با استفاده از میانگین وزنی برای پیش‌بینی نهایی ترکیب می‌شوند. این روش به کاهش سوگیری با تمرکز بر نمونه‌های دشواری که اغلب توسط مدل‌های قبلی به اشتباه طبقه‌بندی می‌شوند، کمک می‌کند مانند مدل‌های Cat Boost (CB)، Ada Boost (Ada)، Light GBM (LGBM) (۱۹).

در مرحله‌ی ارزیابی، معیارهای مهم در ارزیابی مسائل طبقه‌بندی عبارت‌انـــد از دقت (Accuracy)، صحت (Precision)، یادآوری (Recall) و در نهایت امتیاز F1 (F1-score) (۲۰).

$$Accuracy = \frac{T_p + T_N}{T_p + F_p + F_N + T_N}$$

$$Precision = \frac{T_p}{T_p + F_p}$$

$$Recall = \frac{T_p}{T_p + F_N}$$

$$F1 - score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

مثبت صحیح (TP) زمانی اســـت که هر دو کلاس واقعی و پیش بینی از نقاط داده ۱ است، منفی صحیح (TN) زمانی است که هر دو کلاس واقعی و پیش بینی از نقاط داده صفر است، مثبت غلط (FP) زمانی اســـت که کلاس واقعـــی از نقطه داده صفر و کلاس پیش بینی ۱ اســـت و در نهایت منفی غلط (FN) زمانی است که کلاس واقعی از نقطه داده ۱ و کلاس پیش بینی صفر است.

همچنین در پژوهش پیش رو علاوه بر معیارهای ارزیابی اشاره شده از منحنی مشخصه عملکرد گیرنده برای مقایسه عملکرد مدل‌ها اســـتفاده گردید. منحنی مشـــخصه عملکرد گیرنده (Receiver Operating Characteristic)، یک نمایش گرافیکی از عملکرد یک سیستم طبقه‌بندی کننده باینری است زیرا آستانه تمایز آن متفاوت اســـت. مساحت زیر این منحنی معیاری است که معمولاً برای ارزیابی عملکرد یک طبقه‌بندی کننده باینری در داده کاوی و یادگیری ماشین استفاده می‌شود و هر چه این معیار برای مدل بیشتر باشد، از عملکرد خوب مدل حکایت دارد (۲۱).

انتخـــاب ویژگی یکی از مراحل مهـــم در تجزیه و تحلیل داده‌ها برای اعمال انتخاب مجموعه‌ای از ویژگی‌ها اســـت که بیشـــترین تأثیـــر گذاری را بر روی هدف مســـئله دارند (۲۲)؛ بنابراین برای ارزیابـــی میزان تأثیر گذاری متغیرها (انتخاب ویژگی) از الگوریتم Random Forest استفاده شد.

در پژوهـــش حاضر با اســـتفاده از زبان برنامه نویســـی پایتون و کتابخانه‌هـــای مربوط بـــه یادگیری ماشـــین از جمله Numpy، Pandas و Scikit-learn نسبت به آماده سازی داده‌ها و ساخت مدل‌ها اقدامات لازم انجام گردید. این پژوهش با شناســـه اخلاق IR.AJAUMS.REC.۱۴۰۱.۰۷۶ در دانشگاه علوم پزشکی آجا مصوب گردید. لازم به ذکر اســـت، پژوهش حاضر صدمه جانی و روانی بر مشارکت کنندگان نداشت، همچنین سایر مفاد بیانیه‌ی هلســـینکی و اصول اخلاق نشـــر (COPE) از ملاحظات اخلاقی مورد توجه در مطالعه حاضر بودند.

یافته ها

در مطالعه حاضر ۳۹۸ نفر شـــرکت داشـــتند. یافته‌های پژوهش نشـــان داد که ۷۴/۳۷ درصد از بیماران دارای طول مدت اقامت کمتر از ۲۴ ســـاعت و ۲۵/۶۳ درصد از بیماران دارای طول مدت اقامـــت بیش از ۲۴ ســـاعت در بخـــش اورژانس بودنـــد. در این پژوهـــش، ۵۶/۰۳ درصـــد از بیماران را جنســـیت مرد و ۴۳/۹۷ درصد از بیماران را جنســـیت زن تشـــکیل داده است. همچنین نتایج نشـــان داد که میانگین ســـنی جنسیت مرد در کلاس صفر کمتر از میانگین سنی جنسیت زن است (جدول ۱).

جدول ۱- ویژگی‌های بیماران بخش اورژانس

متغیرها		کمتر از ۲۴ ساعت (کلاس صفر) میانگین (انحراف معیار)	بیشتر از ۲۴ ساعت (کلاس یک) میانگین (انحراف معیار)
سن (سال)		۴۸/۷۲ (۲۰/۹)	۵۷/۹۷ (۲۰/۵)
جنسیت	مرد	۴۳/۶۷ (۲۰/۸)	۵۶/۰۲ (۲۱/۸)
	زن	۵۵/۳۵ (۱۹/۲)	۶۰/۲۶ (۱۸/۹)
فشار خون سیستولیک		۱۲۷/۳ (۲۵/۴)	۱۲۸/۳ (۲۴/۲)
فشار خون دیاستولیک		۷۷/۴۴ (۱۶/۶)	۷۶/۵۳ (۱۳/۷)
تعداد تنفس		۱۸/۱۷ (۶/۱)	۱۸/۴۴ (۶/۸)
تعداد ضربان نبض		۸۱/۹۶ (۱۴/۰۴)	۸۶/۹۷ (۱۷/۸۳)
دمای بدن		۳۶/۹۵ (۰/۷)	۳۷/۰۹ (۰/۸)
نحوه‌ی ورود به اورژانس	شخصی	۱۵۰	۵۸
	آمبولانس	۱۴۶	۴۴
شاخص شدت اورژانس	سطح ۱	۳	۶
	سطح ۲	۶۲	۵۲
	سطح ۳	۲۲۷	۴۴
	سطح ۴	۴	۰

در طبقه بندی بین المللی داده‌های این پژوهش، بررسی‌ها نشان داد که بیشترین نوع ثبتی از بیماران مربوط به کد ۹۹ R00-R است که مربوط به بیماری اختلالات بالینی و یافته‌های آزمایشگاهی است. همچنین باید اشاره داشت که در پژوهش حاضر بیمارانی حضور داشتند که صرفاً در یک طبقه بندی از بیماری‌ها قرار نداشتند (جدول ۲).

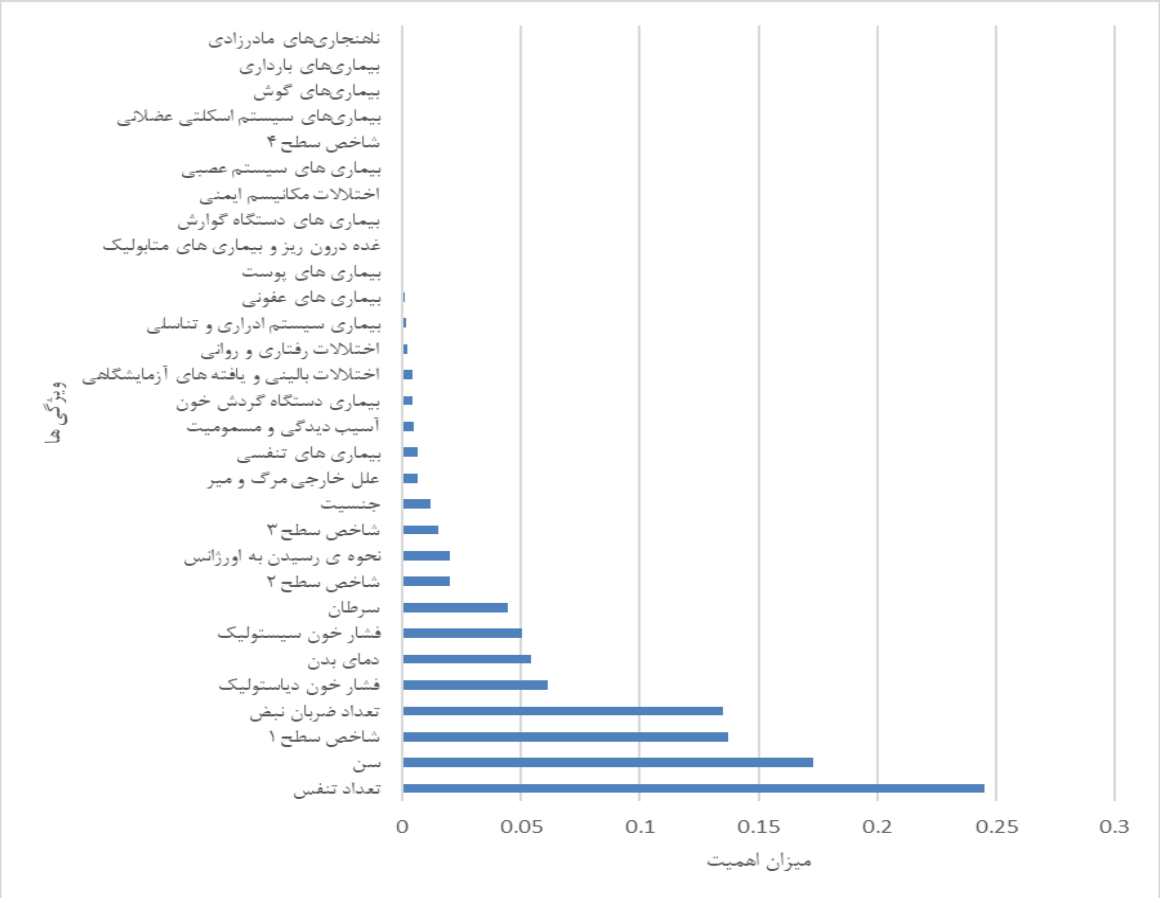
جدول ۲- طبقه‌بندی بین المللی بیماری‌ها برای بیماران بخش اورژانس

طبقه‌بندی بیماری‌ها	تعریف	کمتر از ۲۴ ساعت (کلاس صفر)	بیشتر از ۲۴ ساعت (کلاس یک)
A00-B99	بیماری‌های عفونی	۱۴	۳
C00-D48	سرطان	۳	۵
D50-D89	اختلالات مکانیسم ایمنی	۲	۱
E00-E89	غده درون‌ریز و بیماری‌های متابولیک	۵	۱
F01-F99	اختلالات رفتاری و روانی	۴	۲
G00-G99	بیماری‌های سیستم عصبی	۵	۱
H00-H59	بیماری‌های چشم	۱	۰
H60–H95	بیماری‌های گوش	۰	۰
I00-I99	بیماری‌های دستگاه گردش خون	۱۸	۱۳
J00-J99	بیماری‌های تنفسی	۱۷	۱۳

ادامه جدول ۲- طبقه‌بندی بین المللی بیماری‌ها برای بیماران بخش اورژانس

طبقه‌بندی بیماری‌ها	تعریف	کمتر از ۲۴ ساعت (کلاس صفر)	بیشتر از ۲۴ ساعت (کلاس یک)
K00-K95	بیماری‌های سیستم گوارش	۳	۲
L00-L99	بیماری‌های پوست	۰	۲
M00-M99	بیماری‌های سیستم اسکلتی عضلانی	۱	۰
N00-N99	بیماری‌های دستگاه تناسلی ادراری	۶	۵
O00-O99	بارداری، زایمان	۰	۰
P00-P96	پریناتال (Perinatal)	۰	۰
Q00-Q99	ناهنجاری‌های مادرزادی	۰	۰
R00-R99	اختلالات بالینی و یافته‌های آزمایشگاهی	۱۲۷	۳۹
S00-T88	آسیب‌دیدگی و مسمومیت	۵۵	۱۲
V00-Y99	علل خارجی مرگ و میر	۷۹	۱۷

پس از اعمال الگوریتم Random Forest برای بررسی میزان تأثیرگذاری متغیرها بر روی متغیر هدف، یافته‌ها نشان داد که علائم حیاتی مخصوصاً تعداد تنفس، شاخص شدت اضطراری سطح ۱ و سن از جمله عواملی هستند که بیشترین تأثیر گذاری را در متغیر هدف دارند (نمودار ۱).



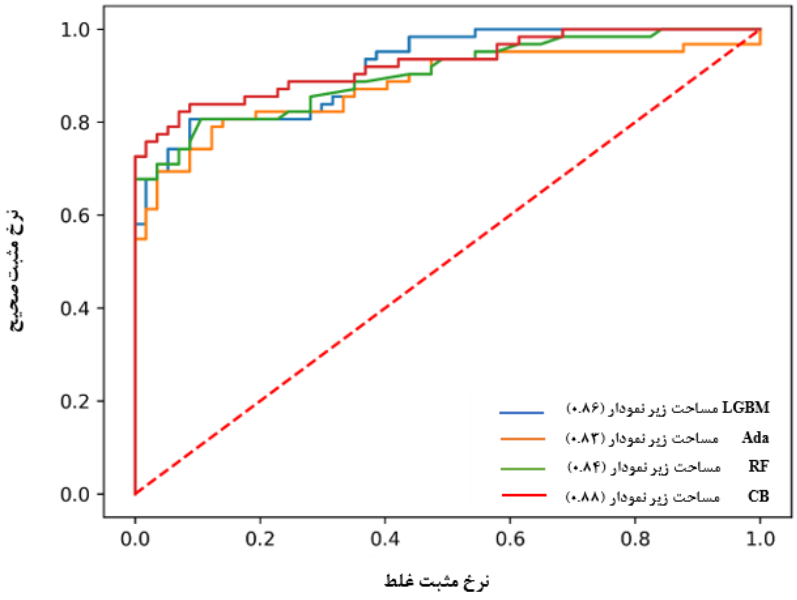
نمودار۱- میزان اهمیت متغیر ها

قبل از اعمال بیش نمونه‌گیری، از ۳۹۸ بیمار بررسی شده ۲۹۶ بیمار در کلاس صفر و ۱۰۲ بیمار در کلاس یک قرار داشتند که این مقادیر منجر به تمایل بیش از حد مدل‌ها به سمت کلاس صفر (Bias) می‌گردید؛ بنابراین پس از اعمال بیش نمونه‌گیری هر دو کلاس متعادل شدند و تعداد بیماران متعلق به کلاس یک به ۲۹۶ بیمار افزایش یافت. در ادامه پس از بررسی اثر متغیرها و آماده‌سازی داده‌ها عملکرد مدل‌های Random Forest- Light GBM- Cat Boost-Ada Boost مورد ارزیابی قرار گرفت (جدول ۳).

جدول ۳- ارزیابی مدل‌ها

مدل	Accuracy	Precision	Recall	score-F۱
Random Forest (RF)	۰/۸۴	۰/۹۲	۰/۷۷	۰/۸۴
Light GBM (LGBM)	۰/۸۵	۰/۹۱	۰/۸۱	۰/۸۵
Ada Boost (Ada)	۰/۸۳	۰/۸۵	۰/۸۰	۰/۸۳
Cat Boost (CB)	۰/۸۷	۰/۹۱	۰/۸۳	۰/۸۷

نتایج این پژوهش نشان داد که مدل Cat Boost عملکرد بالاتری نسبت به سایر مدل‌ها داشت. همچنین به منظور مقایسه‌ی بهتر از عملکرد مدل‌ها از منحنی مشخصه عملکرد گیرنده استفاده شد (نمودار ۲).



نمودار۲- منحنی مشخصه عملکرد گیرنده مدل ها

بحث و نتیجه گیری

این مطالعه با هدف پیش‌بینی طول مدت اقامت بیماران در بخش اورژانس و همچنین مقایسه‌ی عملکرد مدل‌های مجموعه‌ای انجام شد. نتایج این پژوهش نشان داد که علائم حیاتی، شاخص شدت اضطراری سطح ۱ و سن از جمله عواملی هستند که نقش مهمی در طول مدت اقامت بیماران در بخش اورژانس دارند، همچنین مدل‌هایی که از روش Boosting پیروی می‌کنند، نسبت به روش Bagging عملکرد بهتری دارند.

اگرچه مطالعه‌ای عیناً همسو با اهداف پژوهش در پیشینه‌ی پژوهش وجود نداشت، با این حال، یافته‌های پژوهش حاضر با نتایج مطالعه‌ی هیجری و اولاویین (Hijry and Olawoyin) (۱۲) و رحمان (Rahman) و همکاران (۹) همسو و همخوان است. از تفاوت پژوهش حاضر با سایر مطالعات این حوزه علاوه بر مدل‌های استفاده شده می‌توان به اطلاعات جمع آوری شده از

بیماران اشـــاره داشت. در مطالعه‌ی ساریر (Sariyer) و همکاران طول مدت اقامت بیماران بخش اورژانس توســـط مدل‌های RF، شبکه‌های عصبی و لجستیک رگرسیون در حالی مورد پیش بینی قرار گرفت که علائم حیاتی بیمـــاران در بدو ورود آنان به بخش اورژانس در نظر گرفته نشـــد (۱)، در حالی که در پژوهش حاضر علاوه بر طبقه‌بندی بین المللی بیماری‌ها و سایر اطلاعات، علائم حیاتی بیماران در نظر گرفته شد.

شناسایی بیمارانی که در مراکز درمانی در معرض خطر طول مدت اقامت بالا هســـتند، یک تکنیک ارزشمند به حساب می‌آید که نه تنها برای بخش اورژانس، بلکه برای ســـایر بخش‌ها نیز می‌تواند مفید باشـــد چرا که اســـتفاده کارآمد از منابـــع بخش اورژانس و تخصیص مناســـب آن به درک طول کلی اقامت از زمان پذیرش بیمار در اورژانس بستگی دارد (۲۳)؛ بنابراین به نظر می‌رسد که پیش‌بینـــی طول مدت اقامت بیمار در بخش اورژانس می‌تواند به مدیریت و تخصیص بهینه‌ی منابع در این بخش کمک کند.

نقش‌های گسترده و استراتژیک پرستاری نقشی حیاتی در بهبود جریـــان بیمار در محیط‌هـــای چالش برانگیز ایفـــا می‌کند که با زمان انتظـــار، اقامت طولانی در بخش اورژانـــس و ازدحام بیش از حد مشخص می‌شـــود. این نقش‌ها پتانسیل کاهش اقامت در بیمارســـتان، مدت تریاژ و ازدحام بخش اورژانـــس را دارند و در عین حال تجربه کلی کارکنان را بهبود می‌بخشند (۲۴)؛ بنابراین می‌توان نتیجه‌گیری کرد که پیش‌بینی طول مدت اقامت بیماران در بخش اورژانس به نقش استراتژیک و گسترده پرستاران کمک می‌کنـــد و این موضوع هم بر کاهـــش ازدحام بیش از حد بخش اورژانس تأثیرگذار است. با توجه به مطالعه‌ی موری (Mowery) و همکاران (۸)، به نظر می‌رســـد که طـــول مدت اقامت می‌تواند معیاری برای سایر پیامدها باشد، برای مثال از این معیار می‌توان به شدت بیماری و یا احتمال مرگ و میر پی برد.

در رابطه با محدودیت‌های پژوهش حاضر لازم به ذکر اســـت که نتایج آزمایـــش بیماران بخش اورژانس بـــه دلیل نواقصی که در پرونده‌ی آنان وجود داشت در نظر گرفته نشد. همچنین در بازه‌ی زمانی‌ای که اطلاعات بیماران جمع‌آوری شد، بیمارانی که به بخش اورژانـــس مراجعه کرده بودند فاقد بیماری‌های پوســـتی، گوش، ناهنجاری‌های مادرزادی و همچنین بارداری بودند. به دلیل آنکه بیمارســـتان مورد مطالعه در پژوهش حاضر به صورت آموزشـــی و درمانـــی بود، نتایج حاصل از این پژوهش قابل تعمیم به ســـایر بیمارستان‌ها نیســـت و صرفاً برای همان بیمارستان حائز اهمیت اســـت. در چهارچوب یافته‌های پژوهش حاضر یکی از پیشنهادها تبدیل مســـئله از طبقه‌بندی به مساله‌ی رگرسیون است تا مقدار مدت اقامت بیماران را به صورت یک عدد پیش بینی کند تا بتوان بر اساس مقدار پیش بینی شده تخصیص بهینه‌ی منابع در بخش اورژانس را مدیریت و بررســـی نمود. در مطالعات آتی می‌توان از شبکه‌های عصبی برای پیش‌بینی طول مدت اقامت بیماران با در نظر گرفتن جواب آزمایش آنان استفاده نمود.

تشکر و قدردانی

این مقاله برگرفته از طرح پژوهشی می‌باشد که در دانشگاه علوم پزشکی آجا در تاریخ ۱۴۰۱/۰۴/۱۳ با شماره ۲۴۷۳ مصوب شده است. بدین وسیله از کلیه شرکت‌کنندگان در پژوهش، مدیران و مسئولین محترم معاونت آموزشی دانشگاه علوم پزشکی ارتش که در به ثمر رســـیدن این پژوهش ما را یاری نموده‌اند کمال تشکر و قدردانی را داریم.

تضاد منافع

بدین وســـیله کلیه نویســـندگان تصریح می‌نمایند که هیچ گونه تضاد منافعی در خصوص مطالعه حاضر وجود ندارد.

References

1. Sariyer G, Taşar C, Cepe G. Use of data mining techniques to classify length of stay of emergency department patients. Bio-Algorithms and Med-Systems. 2019;15(1). 20180044. DOI: 10.1515/bams-2018-0044
2. Seow E. Leading and managing an emergency department-A personal view. J Acute Med. 2013; 3(3): 61-6. DOI: 10.1016/j.jacme.2013.06.001 PMID: 38620258 PMCID: PMC7147188.
3.Erenler AK, Akbulut S, Guzel M, Cetinkaya H, Karaca A, Turkoz B, Baydin A. Reasons for Overcrowding in the Emergency Department: Experiences and Suggestions of an Education and Research Hospital. Turk J Emerg Med. 2016; 14(2): 59-63. DOI: 10.5505/1304.7361.2014.48802 PMID: 27331171 PMCID: PMC4909875
4. Higginson I. Emergency department crowding. Emerg Med J. 2012; 29(6): 437-43. DOI: 10.1136/emermed-2011-200532 PMID: 22223713
5. Krochmal P, Riley TA. Increased health care costs associated with ED overcrowding. Am J Emerg Med. 1994; 12(3): 265-

6. DOI: 10.1016/0735-6757(94)90135-x. PMID: 8179727.
6. Cecchi E. Emergency department length of stay (ED-LOS) as synonymous with critical and clinical risk. Intern Emerg Med. 2022; 17(1): 191-2. DOI: 10.1007/s11739-021-02893-8. Epub 2021 Nov 25. PMID: 34822123 PMCID: PMC8613516
7. Bukhari H, Albazli K, Almaslmani S, Attiah A, Bukhary E, Najjar F, et al. Analysis of waiting time in emergency department of Al-Noor specialist Hospital, Makkah, Saudi Arabia. Open Journal of Emergency Medicine. 2014; 2(4): 67-73. DOI: 10.4236/ojem.2014.24012
8. Mowery NT, Dougherty SD, Hildreth AN, Holmes JHI, Chang MC, Martin RS, et al. Emergency department length of stay is an independent predictor of hospital mortality in trauma activation patients. J Trauma. 2011; 70(6): 1317-25. DOI: 10.1097/TA.0b013e3182175199 PMID: 21817968
9. Rahman MA, Honan B, Glanville T, Hough P, Walker K. Using data mining to predict emergency department length of stay greater than 4 hours: Derivation and single-site validation of a decision tree algorithm. Emerg Med Australas. 2020; 32(3): 416-21. DOI: 10.1111/1742-6723.13421 PMID: 31808312
10. Shailaja K, Seetharamulu B, Jabbar MA, editors. Machine learning in healthcare: A review. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA); 2018; 910-4. DOI: 10.1109/ICECA.2018.8474918
11. Etu EE, Monplaisir L, Arslanturk S, Masoud S, Aguwa C, Markevych I, et al. Prediction of length of stay in the emergency department for COVID-19 patients: A machine learning approach. IEEE access. 2022; 10: 42243-51. DOI: 10.1109/ACCESS.2022.3168045
12. Hijry H, Olawoyin R, editors. Application of Machine Learning Algorithms for Patient Length of Stay Prediction in Emergency Department During Hajj. 2020 IEEE International Conference on Prognostics and Health Management (ICPHM); 2020 8-10 June 2020. DOI: 10.1109/ICPHM49022.2020.9187055
13. Kadri F, Dairi A, Harrou F, Sun Y. Towards accurate prediction of patient length of stay at emergency department: A GAN-driven deep learning framework. J Ambient Intell Humaniz Comput. 2022: 1-15. DOI: 10.1007/s12652-022-03717-z PMID: 35132336 PMCID: PMC8810344
14. Steindel SJ. International classification of diseases, 10th edition, clinical modification and procedure coding system: Descriptive overview of the next generation HIPAA code sets. J Am Med Inform Assoc. 2010; 17(3): 274-82. DOI: 10.1136/jamia.2009.001230 PMID: 20442144 PMCID: PMC2995704
15. Platts-Mills TF, Travers D, Biese K, McCall B, Kizer S, LaMantia M, et al. Accuracy of the Emergency Severity Index Triage Instrument for Identifying Elder Emergency Department Patients Receiving an Immediate Life-saving Intervention. Acad Emerg Med. 2010; 17(3): 238-43. DOI: 10.1111/j.1553-2712.2010.00670.x PMID: 20370755
16. Wiemer H, Drowatzky L, Ihlenfeldt S. Data Mining Methodology for Engineering Applications (DMME)—A holistic extension to the CRISP-DM model. Applied Sciences. 2019; 9(12): 407. DOI: 10.3390/app9122407
17. Wang S, Dai Y, Shen J, Xuan J. Research on expansion and classification of imbalanced data based on SMOTE algorithm. Sci Rep. 2021; 11(1): 24039. DOI: 10.1038/s41598-021-03430-5 PMID: 34912009 PMCID: PMC8674253
18. Nasteski V. An overview of the supervised machine learning methods. Horizons B. 2017; 4: 51-62. DOI: 10.20544/HORIZONS.B.04.1.17.P05
19. Deshmukh J, Jangid M, Gupte S, Ghosh S, Ingle S, editors. Ensemble method combination: Bagging and boosting. Advanced computing technologies and applications: Proceedings of 2nd international conference on advanced computing technologies and applications—ICACTA 2020; 2020. Springer. DOI: 10.1007/978-981-15-3242-9_38
20. VujoviC Z. Classification model evaluation metrics. International Journal of Advanced Computer Science and Applications. 2021; 12(6): 599-606. DOI: 10.14569/IJACSA.2021.0120670
21. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006; 27(8): 861-74. DOI: 10.1016/j.patrec.2005.10.010
22. Kursa M, Rudnicki W. The All Relevant Feature Selection using Random Forest. 2011. DOI:10.48550/arXiv.1106.5112
23. Gul M, Guneri AF. Forecasting patient length of stay in an emergency department by artificial neural networks. Journal of Aeronautics and Space Technologies. 2015; 8(2): 43-8. DOI: 10.7603/s40690-015-0015-7
24. Sharma S, Rafferty AM, Boiko O. The role and contribution of nurses to patient flow management in acute hospitals: A systematic review of mixed methods studies. Int J Nurs Stud. 2020; 110: 103709. DOI: 10.1016/j.ijnurstu.2020.103709 PMID: 32745787