

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

Moosavi Kashani. S¹*Zargar Balaye Jame. S²

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

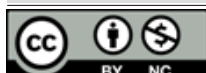
Moosavi Kashani. S , Zargar Balaye Jame. S. Predicting the Length of Stay of Patients in the Emergency Department of the Hospital using Machine Learning Models. Military Caring Sciences. 2024; 11 (2). 114-123.

Received: 31/07/2023

Received in revised form: 11/09/2023

Accepted: 12/09/2023

Published: 12/09/2024



©The Author(s)

Publisher: Aja University of Medical Sciences

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

سینا موسوی کاشانی^۱، ساناز زرگر بالای جمع^۲

چکیده

مقدمه: طول مدت اقامت بیماران در اورژانس بیمارستان، ابزار مهمی در جهت تخصیص بهینه 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 گذشته نگر است. 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 سازی 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 نشان دهنده‌ی آن است که علائم حیاتی، سن و شاخص شدت اضطراری سطح ۱ از جمله عواملی هستند که بیشترین تأثیر را در طول مدت اقامت داشتند، 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 نشان داد که مدل‌های مجموعه‌ای عملکرد خوبی در پیش‌بینی طول مدت اقامت 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، پیش‌بینی، طول مدت اقامت، یادگیری ماشینی

مجله علوم مراقبتی نظامی سال یازدهم ■ شماره ۲ ■ 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

رایج در سراسر جهان است که عواقب ناخواسته‌ای مانند از دست دادن منابع، استفاده بی‌رویه از زمان و نارضایتی پرسنل اورژانس و متقاضیان را به همراه دارد (۳). ارتباط قابل قبولی بین طول مدت اقامت بیماران در بخش اورژانس و شلوغی در این بخش وجود دارد (۴)، همچنین کاهش مدت اقامت بیماران پتانسیل قابل توجهی برای کاهش هزینه‌ها در بخش اورژانس دارد (۵). طول مدت اقامت (Length of Stay) بیماران در بخش اورژانس به عنوان فاصله‌ی زمانی بین رسیدن بیمار به بخش اورژانس تا زمانی که بیمار این بخش را ترک می‌کند گفته می‌شود که بازه‌ی آن برای هر بیمار تقریباً از ۴ تا ۴۸ ساعت متغیر است

خدمات درمانی به یکی از بزرگترین صنایع در سطح جهان تبدیل شده است و بخش اورژانس یکی از بخش‌های اصلی این خدمات را تشکیل می‌دهد و از ضرورت‌های بالایی برخوردار است (۱). بخش اورژانس برای ارائه مراقبت‌های 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 است و دارای چندین تعامل و تراکم تصمیم‌گیری بالا که باعث وقفه در این بخش است می‌باشد (۲). ازدحام بیش از حد در بخش اورژانس مشکلی

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

مطالعات انجام شده در این حوزه حاکی از آن است که برای پیش بینی طول مدت اقامت بیماران در بخش اورژانس از مدل های مجموعه ای (Ensemble) که دقت نسبتاً خوبی در پیش بینی دارند کمتر استفاده شده است و عملکرد الگوریتم های مدل های مجموعه ای 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 مدل های مجموعه ای با یکدیگر مقایسه گردید.

مواد و روش ها

این مطالعه از نوع کوهورت گذشته نگر است. جامعه آماری پژوهش حاضر در برگیرنده کلیه ی بیماران مراجعه کننده به اورژانس یکی از بیمارستان های 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 پرونده ی بیماران استخراج گردید. برای طبقه بندی هر تشخیص بیماری متخصصان مراقبت های بهداشتی از طبقه بندی بین المللی بیماری ها و ویرایش دهم (International Classification of Diseases)، استفاده می کنند که یک کد منحصر به فرد به هر تشخیص اختصاص می دهد (۱۴). برای مثال به بیماری های تنفسی کد J00-J99 تخصیص داده می شود. سایر متغیرهایی که در این پژوهش در نظر گرفته شده است می توان به نحوه ی ورود بیماران به اورژانس و شاخص شدت اورژانس آنان اشاره کرد. نحوه ی 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 بخش های اورژانس برای اولویت بندی بیماران بر اساس شدت وضعیت و سطح مراقبت لازم استفاده می شود؛ که بر اساس این شاخص بیماران در بدو ورود به اورژانس به ۵ سطح طبقه بندی می شوند. 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

(۶). این شاخص یک نشانگر مؤثر برای عملکرد بخش اورژانس و کیفیت تریاژ (Triage) است (۷). مطالعه ی 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 کوتاه تری دارند، بیشتر در معرض بیماری قرار می گیرند (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 می شود (۱۰)، بنابراین با استفاده از پیش بینی طول مدت اقامت بیماران در بخش اورژانس به عنوان ابزار پشتیبانی تصمیم، می توان به بهبود برنامه ریزی منابع اورژانسی و بیمارستانی و اطلاع رسانی به بیماران کمک کرد (۱۱). مطالعات گوناگونی در خصوص پیش بینی طول مدت اقامت بیماران در اورژانس انجام شده است، برای مثال مطالعه ی هیجری و اولوین (Hijry and Olawoyin) با استفاده از الگوریتم های یادگیری ماشین، مانند شبکه های عصبی مصنوعی (Artificial Neural Networks)، رگرسیون های 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 می دهد تا برای سطوح بالای ازدحام آماده شوند (۱۲). نتایج مطالعه ی کادری (Kadri) و همکاران حاکی از آن است که مدل های یادگیری عمیق نسبت به سایر مدل ها برای پیش بینی طول مدت اقامت بیماران در بخش اورژانس برتری دارد (۱۳). همچنین مطالعه ی رحمان (Rahman) و همکاران با هدف توسعه و اعتبار سنجی یک مدل پیش بینی 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 پیش بینی 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 بررسی های انجام شده نتایج حاکی از آن است که پیش بینی طول مدت اقامت بیماران در اورژانس بیمارستان، ابزار مهمی در جهت تخصیص بهینه ی منابع به بیماران بخش اورژانس، بهینه سازی عملیات و کاهش هزینه ها محسوب می شود. همچنین

یک پروژه داده کاوی ارائه می‌دهد. این فرایند کاربردی‌ترین روش مرجع داده کاوی است که دارای ۶ مرحله درک فضای کسب و کار، درک داده، آماده سازی داده‌ها، مدل سازی، ارزیابی و استقرار می‌باشد. (۱۶).

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

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

برای انجام پژوهش پیش رو از فرآیند استاندارد صنعت برای داده کاوی (Cross Industry Standard Process for Data Mining) استفاده شد. این فرآیند روشی مستقل از صنعت و یک رویکرد ساختاریافته و سیستماتیک برای برنامه ریزی و اجرای



شکل ۱- مراحل انجام پژوهش

به ۲ کلاس مورد متعادل قرار گرفت. در ادامه ۸۰ درصد از داده‌ها برای آموزش مدل و ۲۰ درصد هم برای ارزیابی مدل در نظر گرفته شد. با توجه به آنکه ماهیت مدل‌های استفاده شده از جنس درخت تصمیم بود، احتیاجی به نرمال‌سازی داده‌ها وجود نداشت.

در مرحله‌ی مدل‌سازی، با توجه به متغیر خروجی، نوع مدل‌های پژوهش حاضر از نوع یادگیری با ناظر و به صورت طبقه‌بندی (Classification) است. مدل‌های مجموعه‌ای شامل ترکیب چندین مدل برای بهبود دقت و استحکام پیش‌بینی‌ها می‌باشد. روش‌های مجموعه‌ای را می‌توان به دو دسته اصلی تقسیم کرد: (۱) Bagging (۲) Boosting (۱۸). روش Bagging یک روش مجموعه موازی است که شامل آموزش چندین نمونه از یک مدل در زیر مجموعه‌های مختلف داده‌های آموزشی است. سپس پیش‌بینی‌های این مدل‌ها با استفاده از میانگین‌گیری یا رأی‌گیری برای پیش‌بینی نهایی ترکیب می‌شوند. این روش با

در مرحله‌ی آماده سازی داده‌ها، متغیر هدف که طول مدت اقامت بیماران در بخش اورژانس بود به دو کلاس صفر و یک تبدیل شد. کلاس صفر معرف طول مدت اقامت بیمارانی بود که ۲۴ ساعت و یا کمتر از ۲۴ ساعت در بخش اورژانس اقامت داشتند، همچنین کلاس یک معرف طول مدت اقامت بیش از ۲۴ ساعت بود. همچنین لازم به ذکر است، تصدیق و اعتبار سنجی کلاس‌بندی طول مدت اقامت بیماران توسط متخصص اورژانس و معاونت درمان بیمارستان مورد بررسی 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 شدند. عدم تعادل در کلاس‌های صفر و یک مربوط به متغیر هدف یک مشکل جدی برای مسائل یادگیری ماشین از نوع طبقه‌بندی است، بنابراین با استفاده از الگوریتم Synthetic Minority Oversampling Technique (SMOTE) که می‌تواند با بیش نمونه‌گیری، نرخ عدم تعادل را بهبود ببخشد (۱۷)، فراوانی مربوط

از منحنی مشخصه عملکرد گیرنده برای مقایسه عملکرد مدل‌ها استفاده گردید. منحنی مشخصه عملکرد گیرنده (Receiver Operating Characteristic)، یک نمایش گرافیکی از عملکرد یک سیستم طبقه‌بندی کننده باینری است زیرا آستانه تمایز آن متفاوت است. 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).

انتخاب ویژگی یکی از مراحل مهم در تجزیه و تحلیل داده‌ها برای اعمال انتخاب مجموعه‌ای از ویژگی‌ها است که بیشترین تأثیر گذاری را بر روی هدف مسئله دارند (۲۲)؛ بنابراین برای ارزیابی میزان تأثیر گذاری متغیرها (انتخاب ویژگی) از الگوریتم Random Forest استفاده شد.

در پژوهش حاضر با استفاده از زبان برنامه نویسی پایتون و کتابخانه‌های مربوط به یادگیری ماشین از جمله، Numpy، Pandas و Scikit-learn نسبت به آماده سازی داده‌ها و ساخت مدل‌ها اقدامات لازم انجام گردید. این پژوهش با شناسه اخلاق IR.AJAUMS.REC.۱۴۰۱،۰۷۶ در دانشگاه علوم پزشکی آجا مصوب گردید. لازم به ذکر است، پژوهش حاضر صدمه جانی و روانی بر مشارکت کنندگان نداشت، همچنین سایر مفاد بیانی‌یهی هلسینکی و اصول اخلاق نشر (COPE) از ملاحظات اخلاقی مورد توجه در مطالعه حاضر بودند.

یافته‌ها

در مطالعه حاضر ۳۹۸ نفر شرکت داشتند. یافته‌های پژوهش نشان داد که ۷۴/۳۷ درصد از بیماران دارای طول مدت اقامت کمتر از ۲۴ ساعت و ۲۵/۶۳ درصد از بیماران دارای طول مدت اقامت بیش از ۲۴ ساعت در بخش اورژانس بودند. در این پژوهش، ۵۶/۰۳ درصد از بیماران را جنسیت مرد و ۴۳/۹۷ درصد از بیماران را جنسیت زن تشکیل داده است. همچنین نتایج نشان داد که میانگین سنی جنسیت مرد در کلاس صفر کمتر از میانگین سنی جنسیت زن است (جدول ۱).

استفاده از مدل‌های متعدد با واریانس بالا به کاهش بیش از حد برازش کمک می‌کند که در صورت میانگین، یکدیگر را خنثی می‌کنند، مانند جنگل تصادفی (Random Forest (RF))، روش Boosting یک روش مجموعه متوالی است که شامل آموزش مدل‌های ضعیف (مدل‌هایی با دقت پایین) به ترتیب است که در آن هر مدل بعدی از اشتباهات مدل قبلی یاد می‌گیرد. الگوریتم‌های تقویت‌کننده وزن‌های بالاتری را به نمونه‌های طبقه‌بندی اشتباه در داده‌های آموزشی اختصاص می‌دهند، به طوری که مدل‌های بعدی بیشتر بر اصلاح این 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 Light GBM، (Ada) Ada Boost، (CB) Cat Boost (۱۹) (LGBM).

در مرحله‌ی ارزیابی، معیارهای مهم در ارزیابی مسائل طبقه‌بندی عبارت‌اند از دقت (Accuracy)، صحت (Precision)، یادآوری (Recall) و در نهایت امتیاز $F1$ (score-F1) (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 است 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 است که کلاس واقعی از نقطه داده صفر و کلاس پیش بینی ۱ است 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

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)
	زن	۵۵/۳۵ (۱۹/۲)
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
	آمبولانس	۱۴۶
	سطح ۱	۳
شاخص شدت اورژانس	سطح ۲	۶۲
	سطح ۳	۲۲۷
	سطح ۴	۴

در طبقه بندی بین المللی داده‌های این پژوهش، بررسی‌ها نشان داد که بیشترین نوع ثبتي 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

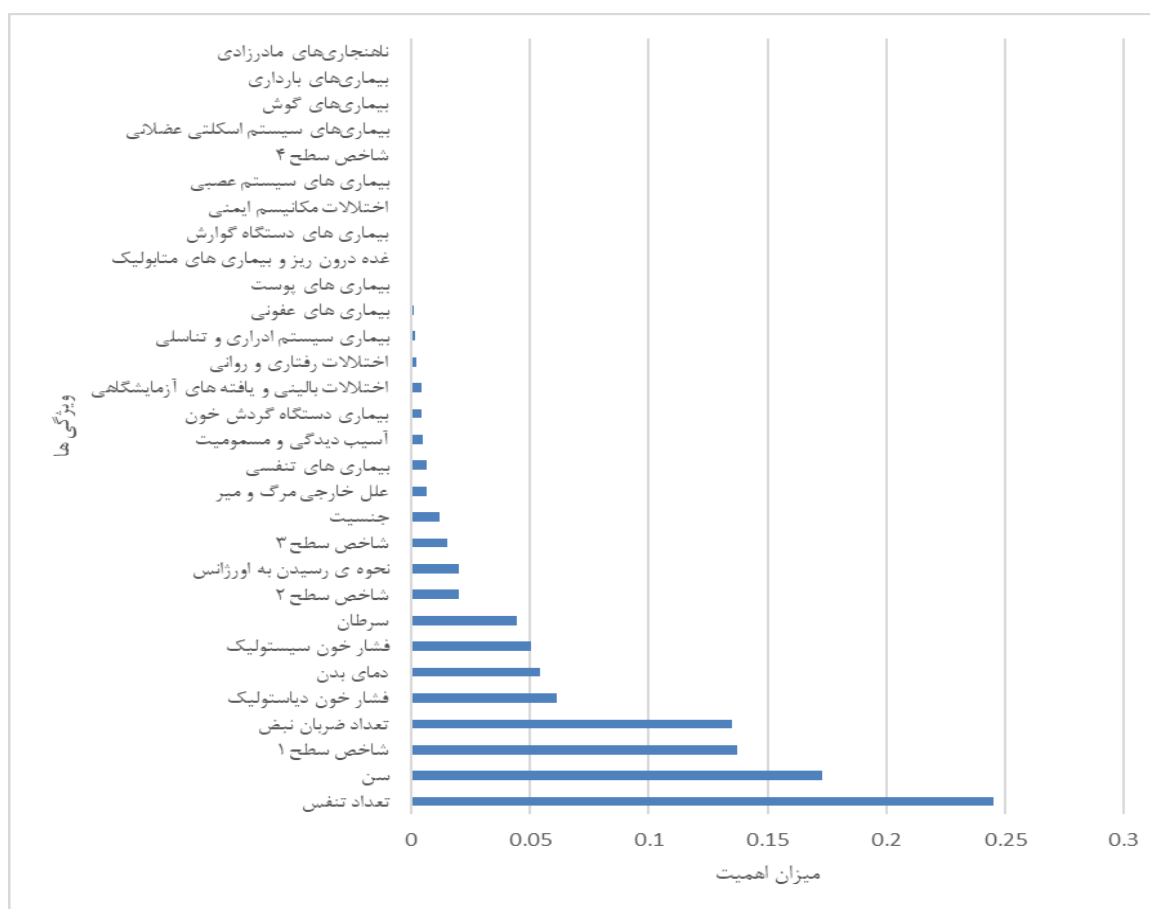
طبقه‌بندی بیماری‌ها	تعریف	کمتر از ۲۴ ساعت (کلاس صفر)	بیشتر از ۲۴ ساعت (کلاس یک)
A00-B99	بیماری‌های عفونی	۱۴	۳
C00-D48	سرطان	۳	۵
D50-D89	اختلالات مکانیسم ایمنی	۲	۱
E00-E89	غده درون‌ریز و بیماری‌های متابولیک	۵	۱
F01-F99	اختلالات رفتاری و روانی	۴	۲
G00-G99	بیماری‌های سیستم عصبی	۵	۱
H00-H59	بیماری‌های چشم	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 بین‌المللی 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 اضطرابی

سطح ۱ و سن از جمله عواملی هستند که بیشترین تأثیرگذاری را در متغیر هدف دارند (نمودار ۱).



نمودار ۱- میزان اهمیت متغیرها

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

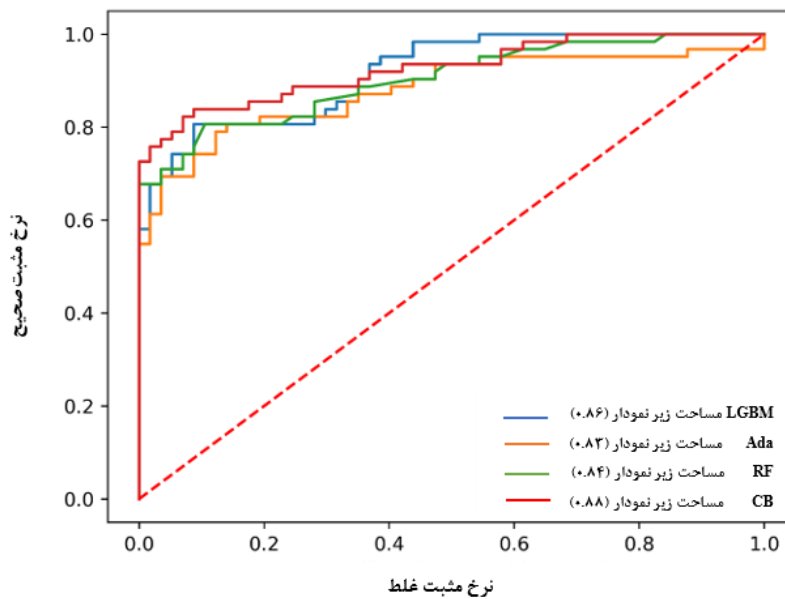
به ۲۹۶ بیمار افزایش یافت. در ادامه پس از بررسی اثر متغیرها و آماده‌سازی داده‌ها عملکرد مدل‌های Random Forest- Light GBM- Cat Boost-Ada Boost مورد ارزیابی قرار گرفت (جدول ۳).

جدول ۳- ارزیابی مدل‌ها

مدل	Accuracy	Precision	Recall	score-F ₁
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).



نمودار ۲- منحنی مشخصه عملکرد گیرنده مدل‌ها

بحث و نتیجه‌گیری

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 اضطرابی 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 اولوین (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 جمع‌آوری شده از

در رابطه با محدودیت‌های پژوهش حاضر لازم به ذکر است که نتایج آزمایش بیماران بخش اورژانس به 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 اورژانس مراجعه کرده بودند فاقد بیماری‌های پوستی، گوش، ناهنجاری‌های مادرزادی و همچنین بارداری بودند. به دلیل آنکه بیمارستان مورد مطالعه در پژوهش حاضر به صورت آموزشی و درمانی بود، نتایج حاصل از این پژوهش قابل تعمیم به سایر بیمارستان‌ها نیست و صرفاً برای همان بیمارستان حائز اهمیت است. در چهارچوب یافته‌های پژوهش حاضر یکی از پیشنهادها تبدیل مسئله از طبقه‌بندی به مسأله‌ی رگرسیون است تا مقدار مدت اقامت بیماران را به صورت یک عدد پیش‌بینی کند تا بتوان بر اساس مقدار پیش‌بینی شده تخصیص بهینه‌ی منابع در بخش اورژانس را مدیریت و بررسی 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 ۱۳/۰۴/۱۴۰۱ 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 رسیدن این پژوهش ما را یاری نموده‌اند کمال تشکر و قدردانی را داریم.

تضاد منافع

بدین وسیله کلیه نویسندگان تصریح می‌نمایند که هیچ گونه تضاد منافی در خصوص مطالعه حاضر وجود ندارد.

بیماران اشاره داشت. در مطالعه‌ی سارییر (Sariyer) و همکاران طول مدت اقامت بیماران بخش اورژانس توسط مدل‌های RF، شبکه‌های عصبی و لجستیک رگرسیون در حالی مورد پیش‌بینی قرار گرفت که علائم حیاتی بیماران در بدو ورود آنان به بخش اورژانس در نظر گرفته نشد (۱)، در حالی که در پژوهش حاضر علاوه بر طبقه‌بندی بین‌المللی بیماری‌ها و سایر اطلاعات، علائم حیاتی بیماران در نظر گرفته شد.

شناسایی بیمارانی که در مراکز درمانی در معرض خطر طول مدت اقامت بالا هستند، یک تکنیک ارزشمند به حساب می‌آید که نه تنها برای بخش اورژانس، بلکه برای سایر بخش‌ها نیز می‌تواند مفید باشد چرا که استفاده کارآمد از منابع بخش اورژانس و تخصیص مناسب آن به درک طول کلی اقامت از زمان پذیرش بیمار در اورژانس بستگی دارد (۲۳)؛ بنابراین به نظر می‌رسد که پیش‌بینی 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 جریان بیمار در محیط‌های چالش برانگیز ایفا می‌کند که با زمان انتظار، اقامت طولانی در بخش اورژانس و ازدحام بیش از حد مشخص می‌شود. این نقش‌ها پتانسیل کاهش اقامت در بیمارستان، مدت تریاژ و ازدحام بخش اورژانس 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 می‌کند و این موضوع هم بر کاهش ازدحام بیش از حد بخش اورژانس تأثیرگذار است. با توجه به مطالعه‌ی موری (Mowery) و همکاران (۸)، به نظر می‌رسد که طول 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.

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/emmermed-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. Vujović 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. Kursu 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