پیش‌بینی مرگ‌ومیر بیماران بستری در بخش‌های مراقبت ویژه: مقایسه سه تکنیک داده‌کاوی: *یک گزارش کوتاه*

زهرا محمدی تقی‌آباد[1]، مریم احمدی[1*]، علیرضا آتشی[2]

۱- گروه مدیریت اطلاعات سلامت، دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران.
۲- گروه سلامت الکترونیک، دانشکده مجازی، دانشگاه علوم پزشکی تهران، تهران، ایران.

[*] نویسنده مسئول: تهران، دانشگاه علوم پزشکی ایران، دانشکده مدیریت و اطلاع‌رسانی، گروه مدیریت اطلاعات سلامت.
تلفن: ۸۸۷۹۳۸۰۵-۰۲۱
E-mail: Ahmadi.m@iums.ac.ir

چکیده

دریافت: ۱۴۰۰/۰۱/۲۸ ویرایش: ۱۴۰۰/۰۲/۰۴ پذیرش: ۱۴۰۰/۰۶/۲۵ آنلاین: ۱۴۰۰/۰۷/۰۱

زمینه و هدف: تحلیل پیامدها در بخش‌های مراقبت ویژه از چالش‌های مهم حوزه سلامت می‌باشد. از این‌رو هدف مطالعه حاضر پیش‌بینی مرگ‌ومیر بیماران بخش‌های مراقبت ویژه (ICUs) با استفاده از چند تکنیک داده‌کاوی است.

روش بررسی: در این مطالعه نرم‌افزار (Weka software, version 3.9.2, University of Waikato, New Zealand) و داده‌های ۸۷۴ بیمار ICUs از دی ۱۳۹۶ تا پایان اسفند ۱۳۹۷ برای توسعه مدل‌ها بکار گرفته شد. بر مبنای چند الگوریتم منتخب داده‌کاوی، مدل‌های پیش‌بینی ایجاد و از شاخصه AUC برای مقایسه عملکرد استفاده شد.

یافته‌ها: بر مبنای ۱۹ متغیر شناخته شده با اهمیت بیشتر از جمله نمره کمای گلاسکو، مدل‌های پیش‌بینی توسعه یافت. اگرچه AUC مدل‌های KNN، شبکه عصبی مصنوعی و درخت تصمیم (به‌ترتیب ۸۱/۵٪، ۷۷/۵٪ و ۷۴/۳٪) قابل قبول بود اما صحت الگوریتم درخت تصمیم J۴۸ (۷۴/۲٪) بالاتر گزارش شد.

نتیجه‌گیری: نتایج این مطالعه بیانگر برتری مدل KNN، نسبت به سایر مدل‌های پیش‌بینی مرگ‌ومیر بیماران ICUs بود.

کلمات کلیدی: داده‌کاوی، بخش‌های مراقبت ویژه، مدل پیش‌بینی.

مقدمه

بیماران در شرایط بحرانی نیازمند مراقبت‌های ویژه برای بقا هستند.[1] اما به‌علت کمبود و یا عدم دسترسی به متخصصان امکان نادیده گرفته شدن برخی جزئیات مهم وجود دارد. آمارها نشان می‌دهد با ارایه خدمات صحیح احتمالا می‌توان مرگ‌ومیر بیماران بخش‌های مراقبت ویژه ((Intensive care units (ICUs)) را تا بیش از ۳۰٪ کاهش داد.[2] بنابراین، احتیاج به تمایز بیماران نیازمند به مراقبت خاص‌تر در ICUs، با کمک شیوه‌های دقیق افزایش یافته است.[3]

در دهه‌های گذشته ابزارهای کشف خودکار، با تحلیل داده‌های خام انبوه به اطلاعات سطح بالا در کمک به تصمیم‌گیری، موثر شناخته شده‌اند.[2و4] به‌طور ویژه ICUs برای پیاده‌سازی یک سیستم مبتنی بر داده جهت شناسایی روابط حجم بالای داده‌ها برای اهداف تشخیصی و پیش‌بینی با استفاده از مدل‌های داده‌کاوی مناسب به‌نظر می‌رسد.[5]

برخی محدودیت‌ها در جمع‌آوری داده‌های پیچیده، چالش‌های مرتبط با روش‌های مختلف داده‌کاوی و معتبر بودن نتایج تحلیل از نظر علمی، تاکیدی بر لزوم ارزیابی دقیق مدل‌های پیش‌بینی است.

امروزه روش پیش‌بینی مبتنی بر تکنیک‌های جدید محاسباتی، مانند شبکه‌های عصبی مصنوعی (Arterial neural networks ((ANNs) و درخت تصمیم ((Decision tree (DT)، منجر به توسعه تعدادی از مدل‌های پیش‌بینی در ICUs شده‌اند، اما عملکرد

[Downloaded from journals.tums.ac.ir on 2024-07-04]

این مدل‌ها همچنان مورد بحث است.[۲و۶و۹]

با توجه به اهمیت پیش‌بینی وضعیت نهایی بیمار در اولویت‌بندی، اخذ تصمیمات درمانی و تخصیص منابع، هدف این مطالعه ایجاد مدل پیش‌بینی مرگ‌ومیر ICUs بر مبنای الگوریتم‌های یادگیری ماشین و نهایتا ارزیابی مدل‌های پیش‌بینی است.

روش بررسی

نمونه پژوهش این مطالعه گذشته‌نگر از رکوردهای ۲۴ ساعت اولیه ۸۷۴ بیمار بستری در ICUs در بیمارستان‌های رسول اکرم و فیروزگر از دی ماه ۱۳۹۶ تا پایان اسفند ۱۳۹۷، جمع‌آوری گردید.

ابتدا به‌منظور پاک‌سازی داده‌ها تعداد رکوردهای خالی برای هر عامل محاسبه و با مراجعه به سیستم اطلاعات بیمارستانی (Hospital information system (HIS)) رکوردهای خالی برای هر بیمار براساس اطلاعات موجود تکمیل شد. سپس، تعداد متغیرهایی با بیش از ۸۵٪ مفقودی از مطالعه خارج شدند. مابقی داده‌ها به‌ازای هر عامل موثر مفقود با مشورت پزشکان به متوسط میزان نرمال تغییر یافت تا مجموعه داده نهایی برای پردازش آماده شود. قابل ذکر است، اطلاعات شناسایی بیمار برای حفظ محرمانگی حذف شدند.

بیماران به دو دسته بیماران ترخیصی و متوفی تقسیم‌بندی شدند. برای ساخت مدل پیش‌بینی مرگ‌ومیر بیماران ICUs، از الگوریتم J۴۸، K نزدیک‌ترین همسایه (K Nearest Neighbor (KNN)) با تعداد هفت همسایه و شبکه عصبی (پرسپترون چند لایه) تحت نرم‌افزار Weka software, version 3.9.2, University of Waikato, New Zealand استفاده شد. به‌منظور ارزیابی عملکرد مدل‌های پیش‌بینی، مساحت تحت منحنی مشخصه عملکرد Area under the curve (AUC) گزارش شد. AUC بیان‌گر آن است که مدل به چه میزان می‌تواند بیماران زنده و متوفی را براساس مقادیر ویژگی و حساسیت افتراق دهد.[۱]

یافته‌ها

در این مطالعه، ۲۶۲ نفر (۲۹٪/۹۸) از بیماران متوفی و ۶۱۲ نفر (۷۰٪/۰۲) از بیماران زنده مانده‌اند. بیماران دیابتیک ۱۵۳ نفر (۱۷٪/۵۱)، بیماران ونتیله ۴۰۳ نفر (۴۶٪/۱۱) و بیماران پس از جراحی ۵۰۱ نفر (۵۷٪/۳۲) بوده‌اند. همچنین، نمونه پژوهش شامل ۵۰۹ نفر مرد (۵۸٪/۲۴) و ۳۶۵ نفر زن (۴۱٪/۷۶) و بیشترین رنج سنی افراد نمونه متعلق به گروه ۷۵-۵۱ سال (۹۹٪/۴۱) بود.

مجموعه داده‌ها متشکل از ۳۱ متغیر و یک متغیر وضعیت نهایی فرد بستری بود. از WEKA نسخه ۱۹ برای شناسایی عوامل موثر استفاده شد که بر این اساس نمره کمای گلاسکو (Glasgow Coma Scale (GCS)) بیشترین اهمیت را داشت.

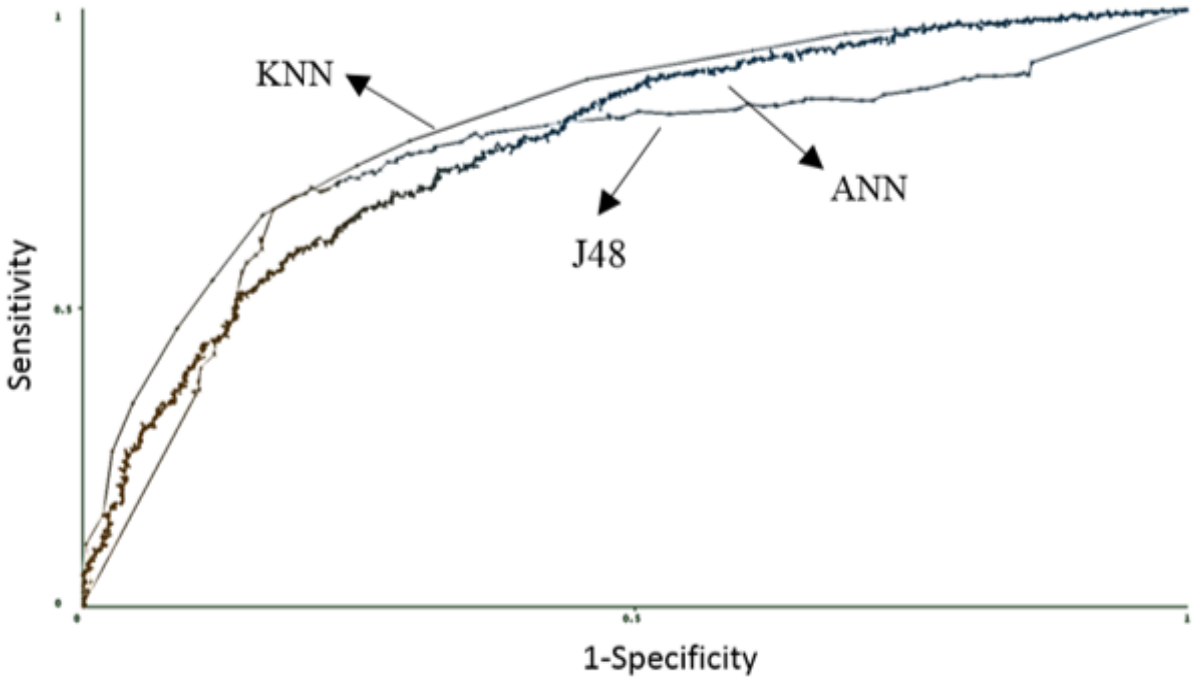
وضعیت ونتیلاسیون، بیمار پس از جراحی، جنسیت، درجه حرارت، کراتینین سرم، دیابت، Blood urea nitrogen (BUN)، سن، اعتیاد، شاخصه International Normalized Ratio (INR) ، Partial Thromboplastin Time (PTT)،PH، آلبومین، هموگلوبین، گلوکز، تعداد نبض، هماتوکریت و فشار اکسیژن (PO2) به‌ترتیب در رتبه‌های بعدی از نظر اهمیت قرار گرفتند و مدل‌ها بر مبنای این عوامل توسعه یافت. سایر عوامل مانند شمار گلوبول سفید، فشارخون، تعداد تنفس، بیلی‌روبین و سدیم سرمی حایز اهمیت شناخته نشدند.

در جدول ۱ حساسیت، ویژگی، صحت، دقت و AUC تمام مدل‌های پیش‌بینی نتایج ICUs ارایه شده است. بر این اساس الگوریتم KNN با ۷۷/۳٪ بیشترین دقت را داشت.

جدول ۱: حساسیت، ویژگی، صحت، دقت، نرخ خطا و AUC الگوریتم‌ها

الگوریتم‌ها	حساسیت	ویژگی	صحت (درصد)	دقت (درصد)	نرخ خطا (درصد)	AUC (درصد)
J48	۰/۷۰۰۳	۰/۷۷۹۴	۷۴/۲	۷۳/۱	۲۵/۸	۷۴/۳
KNN	۰/۶۵۶۴	۰/۸۳۴۹	۷۲/۲	۷۷/۳	۲۴/۸	۸۱/۵
ANN	۰/۷۰۷۱	۰/۷۲۵۴	۷۱/۵	۶۸/۷	۲۸/۵	۷۷/۵

[Downloaded from journals.tums.ac.ir on 2024-07-04]



شکل ۱: منحنی ROC مدل‌های پیش‌بینی مرگ‌ومیر بیماران ICUs

براساس AUC بهترین مدل ایجاد شده به‌ترتیب، KNN، ANN و در نهایت J۴۸ بود (شکل ۱). به‌طور معمول، AUC قابل قبول از نظر توسعه‌دهندگان مدل ۷۰٪ است.

بحث

به‌علت نیاز به کالیبراسیون و عدم دقت کافی مدل‌های رایج پیش‌بینی مرگ‌ومیر بیماران ICUs، تکنیک‌های داده‌کاوی می‌توانند علاوه بر تعیین متغیرهای موثر، برای ایجاد مدل پیش‌بینی نیز مورد استفاده قرار گیرند.[۲] اما باید توجه داشت پیش‌بینی وضعیت نهایی بیماران ICUs و شناسایی فاکتورهای تاثیرگذار بر آن، براساس داده‌ها و تکنیک داده‌کاوی به‌کارگرفته، ممکن است متفاوت باشد.

جدول ۲ مقایسه‌ای اجمالی بین مطالعات مشابه با این پژوهش را نمایش می‌دهد. در مطالعات جمعیت مورد مطالعه دارای بیماری زمینه‌ای خاص هستند.[۶و۸و۱۰] مدل‌های پیش‌بینی مرگ‌ومیر در این مطالعه علاوه بر حجم نمونه کمتر، براساس اطلاعات موجود در پرونده الکترونیک اکثر بیماران ICUs و بدون محدودیت از لحاظ ابتلا به بیماری زمینه‌ای توسعه یافت.

نتایج مطالعه نشان داد تکنیک‌های یادگیری ماشین موجب بهبود تصمیمات در سیستم‌های مراقبتی و درمانی می‌شوند اما به‌علت عدم بررسی برخی از این الگوریتم‌ها، مانند KNN در بسیاری از این سیستم‌ها، پیشنهاد می‌شود با استفاده از داده‌های مراقبت‌های ویژه در دسترس، نتایج این مطالعه با نتایج حاصل از ارزیابی وضعیت نهایی بیماران با سایر ابزارهای تعیین وخامت بیماری‌ها مقایسه شده تا از اعتبار و تعمیم‌پذیری نتایج اطمینان حاصل گردد. بنابراین با توجه به پر هزینه بودن خدمات ICUs و اهمیت پیش‌بینی دقیق وضعیت نهایی بیماران، توسعه مدل‌های بومی پیش‌بینی مرگ‌ومیر بیماران بستری در ICUs ضروری به‌نظر می‌رسد.

[Downloaded from journals.tums.ac.ir on 2024-07-04]

جدول ۲: مطالعات مشابه با مطالعه حاضر در پیش‌بینی نتایج ICUs

نویسندگان	سال انتشار	تعداد بیماران	الگوریتم دسته‌بندی	AUC (درصد)	شایع‌ترین تشخیص	تعداد عوامل شناخته شده موثر
Kim *et al.*[2]	۲۰۱۱	۲۳٤٤٦	DT ANN SVM	۸۹/۲ ۸۷/٤ ۸۷/٦		۱۵
Lin *et al.*[6]	۲۰۱۹	۱۹۰٤٤	RF SVM ANN	۸٦/٦ ۸٦/۱ ۸۳/۳	نارسایی حاد کلیوی	۱۷
Anand *et al.*[8]	۲۰۱۸	٤۱۱۱	RF	۷۸/۷	دیابت	۵
Hsieh *et al.*[9]	۲۰۱۸	۳٤۱	ANN	۸٤/٦		۱۶
Jaimes *et al.*[10]	۲۰۰۵	۵۳۳	ANN	۸۷/۸۲	عفونت	
Awad *et al.*[11]	۲۰۱۷	۱۱۷۲۲	RF	۹۰		۱۵
Our study	۲۰۱۹	۸۷٤	ANN KNN DT	۷۷/۵ ۸۱/۵ ۷٤/۳		۱۹

به‌منظور تایید صحت و دقت مدل‌های پیش‌بینی در حوزه سلامت، بایستی این‌گونه مطالعات، با تعداد بیمار بیشتر و بیمارستان‌هایی با افرادی مبتلا به بیماری‌های مختلف و بستری در ICUs انجام شود.

سپاسگزاری: این مطالعه برگرفته از پایان‌نامه مقطع کارشناسی ارشد تحت عنوان "مقایسه‌ی مدل‌های پیش‌بینی مرگ‌ومیر بیماران بستری در بخش‌های مراقبت ویژه براساس تکنیک‌های داده‌کاوی و مدل آپاچی" با کد اخلاق IUMS/SHMIS-98-275-14-987 می‌باشد که با حمایت دانشگاه علوم پزشکی و خدمات بهداشتی درمانی ایران اجرا شده است.

References

1. Ramon J, Fierens D, Güiza F, Meyfroidt G, Blockeel H, Bruynooghe M, et al. Mining data from intensive care patients. *Adv Eng Inform* 2007;21(3):243-56.
2. Kim S, Kim W, Park RW. A comparison of intensive care unit mortality prediction models through the use of data mining techniques. *Healthc Inform Res* 2011;17(4):232-43.
3. Pantet O, Faouzi M, Brusselaers N, Vernay A, Berger MM. Comparison of mortality prediction models and validation of saps ii in critically ill burns patients. *Ann Burns Fire Disasters* 2016;29(2):123-9.
4. Lu H-Y, Li T-C, Tu Y-K, Tsai J-C, Lai H-S, Kuo L-T. Predicting long-term outcome after traumatic brain injury using repeated measurements of glasgow coma scale and data mining methods. *J Med Syst* 2015;39(2):14-24.
5. Silva Á, Cortez P, Santos MF, Gomes L, Neves J. Mortality assessment in intensive care units via adverse events using artificial neural networks. *Artif Intell Med* 2006;36(3):223-34.
6. Lin K, Hu Y, Kong G. Predicting in-hospital mortality of patients with acute kidney injury in the icu using random forest model. *Int J Med Inform* 2019;125:55-61.
7. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the super icu learner algorithm (sicula): A population-based study. *Lancet Respir Med* 2015;3(1):42-52.
8. Anand RS, Stey P, Jain S, Biron DR, Bhatt H, Monteiro K, et al. Predicting mortality in diabetic icu patients using machine learning and severity indices. *AMIA Jt Summits Transl Sci Proc* 2018;2018:310.
9. Hsieh Hsieh M-H, Hsieh M-J, Chen C-M, Hsieh C-C, Chao C-M, Lai C-C. An artificial neural network model for predicting successful extubation in intensive care units. *J Clin Med* 2018;7(9):240.
10. Jaimes F, Farbiarz J, Alvarez D, Martínez C. Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Crit Care* 2005;9(2):1-7.
11. Awad A, Bader-El-Den M, McNicholas J, Briggs J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Inform* 2017;108:185-95.

[Downloaded from journals.tums.ac.ir on 2024-07-04]

TUMJ Tehran University Medical Journal, October 2021; Vol. 79, No. 7: 568-572 Brief Report

Prediction of mortality in patients admitted to intensive care units, A comparison of three data mining techniques: *a brief report.*

Zahra Mohammadi Taghiabad M.Sc.[1]
Maryam Ahmadi Ph.D.[1*]
Alireza Atashi Ph.D.[2]

1- Department of Health Information Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran.
2- Department of Electronic Health, Virtual School, Tehran University of Medical Sciences, Tehran, Iran.

[*]Corresponding author: Department of Health Information Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran.
Tel: +98-21-88793805
E-mail: Ahmadi.m@iums.ac.ir

Abstract

Received: 17 Apr. 2021 Reised: 24 Apr. 2021 Accepted: 16 Sep. 2021 Available online: 23 Sep. 2021

Background: Early outcome prediction of hospitalized patients is critical because the intensivists are constantly striving to improve patients' survival by taking effective medical decisions about ill patients in Intensive Care Units (ICUs). Despite rapid progress in medical treatments and intensive care technology, the analysis of outcomes, including mortality prediction, has been a challenge in ICUs. Hence, this study aims to predict the mortality of patients admitted to ICUs using data mining techniques.

Methods: In this study, among the cases of patients who were admitted to ICUs of the Rasoul Akram and Firoozgar hospitals of Tehran City, Iran, from December 2017 to March 2018, the first 24 hours of the ICUs admission data of 874 cases were gathered. A new model based on the standard methodology CRISP was developed. In the modeling section, two well-known data mining techniques called artificial neural network (ANN), K nearest neighbor (KNN) and decision tree (DT) were used. WEKA 3.9.2 open-source software was implemented for data analysis. Finally, according to the accuracy, sensitivity, specificity criteria and AUC-ROC Curve, the superior model was introduced.

Results: Based on the WEKA results, 19 variables had the most impact on the mortality prediction of patients admitted to ICUs including Glasgow Coma Scale (GCS), mechanical ventilation, surgical service at ICUs admission, gender, temperature, serum creatinine, diabetes, Blood urea nitrogen (BUN), age, addiction, International Normalized Ratio (INR), PH, Partial Thromboplastin Time (PTT), albumin, hemoglobin, glucose, pulse rate, hematocrit (HCT), PO2. Based on the created models, some rules have been extracted which can be used as a pattern to predict the probability of mortality. Although the AUC of the three models was acceptable (KNN 81.5%, ANN 77.5% and DT 74.3%), but the accuracy of decision tree J48 (74.2%) was higher.

Conclusion: The study indicated that in the KNN model, the rules derived from it can be effective in mortality prediction in patients admitted to ICUs.

Keywords: data mining, intensive care units, prediction models.

Copyright © 2021 Tehran University of Medical Sciences. Published by Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Non-commercial uses of the work are permitted, provided the original work is properly cited.

[Downloaded from journals.tums.ac.ir on 2024-07-04]