



Review Article

The Role of Artificial Intelligence in Predicting Acute Kidney Injury: A Systematic Review of Machine Learning and Deep Learning Approaches

Masoumeh Kheirkhahzadeh¹, Amin Golabpour^{2*}

1. Assistant Professor of Artificial Intelligence, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Iran
2. Assistant Professor of Medical Informatics, School of Allied Medical Sciences, Shahrood University of Medical Sciences, Shahrood, Iran

***Corresponding author:** Amin Golabpour, School of Allied Medical Sciences, Shahrood University of Medical Sciences, Shahrood, Iran. Email: a.golapour@shmu.ac.ir

DOI: [10.66224/nkums.18.1.91](https://doi.org/10.66224/nkums.18.1.91)

How to Cite this Article:

Kheirkhahzadeh M, Golabpour A. The Role of Artificial Intelligence in Predicting Acute Kidney Injury: A Systematic Review of Machine Learning and Deep Learning Approaches. J North Khorasan Univ Med Sci. 2026;18(1): 91-102
DOI: [10.66224/nkums.18.1.91](https://doi.org/10.66224/nkums.18.1.91)

Received: 22 July 2025

Accepted: 06 December 2025

Keywords:

Acute Kidney Injury
Artificial Intelligence
Black-box Algorithms
Machine Learning
White-box Algorithms

Abstract

Introduction: Acute kidney injury (AKI) is a common and potentially reversible condition in hospital settings and is associated with adverse outcomes, such as increased mortality, prolonged hospitalization, and higher healthcare costs. In this context, artificial intelligence (AI) algorithms have emerged as promising approaches for the early identification of patients at risk of AKI. This study aimed to systematically review research that has applied AI models for AKI prediction, with a focus on their performance, interpretability, and clinical evaluation.

Methods: A systematic search was conducted in PubMed, Scopus, and Web of Science databases. Studies were included if they applied machine learning or deep learning algorithms to predict AKI. Extracted data included the type of algorithm, sample size, number of features, performance metrics, degree of interpretability, and presence of clinical evaluation. The models were categorized into white-box and black-box algorithms.

Results: Among the 208 included studies, most reported an AUROC greater than 0.90. AUROC values reached up to 0.94 for white-box models and up to 0.97 for black-box models. Regression analysis indicated that dataset size had a significant impact on the performance of white-box models, whereas no such relationship was observed in black-box algorithms. Many studies also reported accuracy values ranging from 75% to 95%.

Conclusions: Despite their higher interpretability, white-box models were used less frequently. Enhancing transparency, increasing clinical expert involvement, and conducting real-world clinical evaluations are essential for the broader adoption and practical implementation of AI-based models in AKI prediction.



مرور نظام‌مند نقش هوش مصنوعی در پیش‌بینی آسیب حاد کلیوی: مقایسه عملکرد و قابلیت تبیین‌پذیری

معصومه خیرخواه‌زاده^۱، امین گلاب‌پور^{۲*}

^۱ استادیار هوش مصنوعی، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز، اهواز، ایران
^۲ استادیار انفورماتیک پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شاهرود، شاهرود، ایران

*نویسنده مسئول: امین گلاب‌پور، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شاهرود، شاهرود، ایران. ایمیل: a.golapour@shmu.ac.ir

DOI: 10.66224/nkums.18.1.91

چکیده مقدمه: آسیب حاد کلیوی (AKI) یکی از مشکلات رایج و بالقوه برگشت‌پذیر در محیط بیمارستانی است و پیامدهایی مانند افزایش مرگ‌ومیر، بستری طولانی‌مدت و هزینه‌های درمانی را به همراه دارد. در این میان، الگوریتم‌های هوش مصنوعی به‌عنوان رویکردی نویدبخش برای شناسایی زودهنگام بیماران در معرض خطر مطرح شده‌اند. هدف این مطالعه، مرور نظام‌مند پژوهش‌هایی است که از مدل‌های هوش مصنوعی در پیش‌بینی AKI استفاده کرده‌اند، با تمرکز بر عملکرد، تبیین‌پذیری و ارزیابی بالینی. مواد و روش‌ها: جست‌وجوی نظام‌مند در پایگاه‌های PubMed، Scopus، و Web of Science انجام شد. مقالاتی انتخاب شدند که از الگوریتم‌های یادگیری ماشین یا یادگیری عمیق برای پیش‌بینی AKI استفاده کرده بودند. داده‌های استخراج‌شده شامل نوع الگوریتم، حجم نمونه، تعداد ویژگی‌ها، شاخص‌های عملکرد، میزان تبیین‌پذیری و وجود ارزیابی بالینی بود. همچنین، مدل‌ها در دو گروه جعبه‌سفید و جعبه‌سیاه طبقه‌بندی شدند. یافته‌ها: از میان ۲۰۸ مقاله واردشده، اکثر مطالعات AUROC بالاتر از ۰/۹۰ را گزارش کردند. مقدار AUROC در مدل‌های جعبه‌سفید تا ۰/۹۴ و در مدل‌های جعبه‌سیاه تا ۰/۹۷ افزایش یافته بود. تحلیل رگرسیون نشان داد در مدل‌های جعبه‌سفید، اندازه مجموعه داده اثر معناداری بر عملکرد داشت، درحالی‌که این ارتباط در مدل‌های جعبه‌سیاه دیده نشد. بسیاری از مطالعات دقتی بین ۷۵ تا ۹۵ درصد گزارش کرده بودند. نتیجه‌گیری: باوجود تبیین‌پذیری بالاتر، مدل‌های جعبه‌سفید کمتر استفاده شده‌اند. ارتقای شفافیت، مشارکت متخصصان و انجام ارزیابی بالینی برای پذیرش و کاربرد گسترده مدل‌های هوش مصنوعی ضروری است.	تاریخ دریافت: ۱۴۰۴/۰۴/۳۱ تاریخ پذیرش: ۱۴۰۴/۰۹/۱۵ واژگان کلیدی: آسیب حاد کلیوی الگوریتم‌های جعبه‌سفید الگوریتم‌های جعبه‌سیاه هوش مصنوعی یادگیری ماشین
--	---

مقدمه

پرخطر گزارش کرده‌اند. برای مثال، مطالعه‌ای روی ۵۱۲ بیمار بزرگسال تحت عمل جراحی قلب بین سال‌های ۲۰۱۵ تا ۲۰۱۶ نشان داد که افزایش سطح کراتینین طی ۴۸ ساعت پس از عمل به‌عنوان AKI شناخته شده و عوامل خطر مهمی مانند سطح کراتینین پیش از عمل، قند خون ناشتا و سطح هموگلوبین مرتبط با بروز AKI شناسایی شدند [۴]. بروز بالای AKI در میان بیماران تحت عمل جراحی قلب، بیماران ICU و کسانی که دچار حوادثی نظیر غرق‌شدگی یا شوک سپتیک شده‌اند، نشان‌دهنده اهمیت پیشگیری و تشخیص زودهنگام آن است. از آن‌جاکه تشخیص بالینی AKI غالباً با تأخیر صورت می‌گیرد و شاخص‌های آزمایشگاهی رایج نظیر کراتینین حساسیت کافی ندارند، کاربرد روش‌های هوش مصنوعی و یادگیری ماشین به‌عنوان ابزارهای پیش‌بینی‌کننده می‌تواند نقشی کلیدی در بهبود پیامدهای بالینی و کاهش هزینه‌های درمان ایفا کند.

کاربرد هوش مصنوعی و یادگیری ماشین در پزشکی به‌طور

آسیب حاد کلیوی (Acute Kidney Injury/ AKI) یک اختلال شایع و خطرناک است که به صورت کاهش ناگهانی عملکرد کلیه‌ها ظرف چند ساعت تا چند روز بروز می‌کند. این اختلال بالقوه برگشت‌پذیر است و می‌تواند پیامدهای جدی داشته باشد. تخمین زده شده است که سالانه حدود ۱۳/۳ میلیون نفر در سراسر جهان به AKI مبتلا می‌شوند که بیش از ۱/۷ میلیون مورد مرگ‌ومیر ناشی از آن گزارش شده است [۱]. این عارضه در حدود ۱۰ تا ۲۰ درصد از بیماران بستری و تا ۵۰ درصد از بیماران بستری در بخش مراقبت‌های ویژه مشاهده می‌شود [۲]. حتی شکل‌های خفیف AKI نیز می‌توانند با افزایش معنادار خطر مرگ، نارسایی مزمن کلیه (Chronic kidney disease/CKD)، بیماری‌های قلبی-عروقی و بستری مجدد همراه باشند [۳]. در صورت پیشرفت به CKD و نارسایی مزمن کلیه، نیاز به انجام دیالیز ضروری است که باعث کاهش کیفیت زندگی بیماران می‌شود.

در ایران نیز مطالعات متعددی شیوع بالای AKI را در گروه‌های

بر لزوم توسعه مدل‌های شفاف‌تر، انجام ارزیابی‌های بالینی واقعی و توجه بیشتر به اشتراک‌گذاری داده‌ها و کد تأکید می‌کنند تا قابلیت استفاده بالینی این فناوری‌ها ارتقا یابد [۸].

مطابق با مقررات اتحادیه اروپا، به‌ویژه قانون هوش مصنوعی (AI Act ۱۶۸۹/۲۰۲۴)، سیستم‌های هوش مصنوعی در حوزه پزشکی، از جمله الگوریتم‌های پیش‌بینی AKI، در دسته‌بندی «پرخطر» قرار دارند و ملزم به رعایت الزامات خاصی هستند [۱۶]. یکی از این الزامات، شفافیت و قابلیت تبیین (Explainability) مدل‌ها است تا تصمیمات اتخاذشده توسط سیستم برای کاربران بالینی به‌روشنی قابل فهم باشد. این هدف می‌تواند از طریق استفاده از مدل‌های جعبه‌سفید (White-Box) یا بهره‌گیری از روش‌های تبیینی مانند SHAP و LIME روی مدل‌های پیچیده‌تر محقق شود. علاوه‌براین، اعتبارسنجی و ارزیابی بالینی در محیط‌های واقعی، همراه با مشارکت فعال متخصصان بالینی در تمامی مراحل طراحی، پیاده‌سازی و تفسیر سیستم‌های هوش مصنوعی، ضروری است تا اطمینان حاصل شود که مدل‌ها با استانداردهای بالینی هم‌خوانی دارند و از بروز مخاطرات ناشی از تفسیر یا تصمیم‌گیری نادرست جلوگیری شود. همچنین، باید امکان نظارت انسانی بر فرایند تصمیم‌گیری فراهم شود، به‌گونه‌ای که تصمیم‌نهایی همواره با تأیید پزشک باشد و الگوریتم‌ها نتوانند به‌طور خودسرانه توصیه‌های درمانی ارائه دهند. چنین رویکردی نه‌تنها دقت فنی مدل‌ها را تضمین می‌کند، بلکه از نظر رعایت اصول اخلاقی، حقوقی و حفظ ایمنی بیماران نیز اعتبار و اعتماد لازم را به همراه خواهد داشت.

با گسترش کاربرد هوش مصنوعی در پزشکی، به‌ویژه در حوزه پیش‌بینی و پیش‌آگاهی AKI، توجه به جنبه‌هایی فراتر از دقت الگوریتمی اهمیت یافته است. در سال‌های اخیر، مطالعات متعددی از الگوریتم‌های یادگیری ماشین و یادگیری عمیق برای شناسایی بیماران در معرض خطر AKI استفاده کرده‌اند؛ باین‌حال، پرسش مهم این است که آیا این مدل‌ها علاوه‌بر عملکرد فنی، ویژگی‌هایی چون قابلیت تبیین، شفافیت، اعتبارسنجی بالینی، مشارکت متخصصان بالینی در طراحی مدل و امکان نظارت انسانی را نیز رعایت کرده‌اند یا خیر. این مطالعه با هدف مرور نظام‌مند تحقیقات انجام‌شده در این زمینه، به بررسی میزان تحقق این شاخص‌های کلیدی در مطالعات هوش مصنوعی مرتبط با AKI می‌پردازد.

روش کار

این مطالعه به‌صورت مرور نظام‌مند طراحی شد تا بررسی کند مقالات اصیل منتشرشده در زمینه پیش‌بینی AKI با استفاده از هوش مصنوعی (AI) و یادگیری ماشین (ML)، تا چه میزان از الگوریتم‌های تفسیرپذیر استفاده کرده‌اند و آیا ارزیابی‌های آماری و بالینی مناسب در آنها صورت گرفته است. این مرور براساس اصول PRISMA و بدون هیچ‌گونه محدودیت زمانی انجام شد.

در مرحله اول، مفاهیم کلیدی مورد بررسی در این مطالعه شامل آسیب حاد کلیه و تمامی واژه‌ها و اصطلاحات مترادف و مرتبط با آن بود تا پوشش جامعی از مطالعات مرتبط فراهم شود. همچنین، هدف اصلی مطالعه بر پیش‌بینی (Prediction)، پیش‌نگری (Forecasting)، و پیش‌آگاهی (Prognosis) وضعیت AKI متمرکز بود تا مدل‌ها و

چشمگیری افزایش یافته است [۵-۷]. از جمله این کاربردها، پیش‌بینی آسیب حاد کلیوی (AKI) است که در سال‌های اخیر مورد توجه قرار گرفته است [۸، ۹] و به‌عنوان ابزاری نویدبخش برای بهبود تشخیص زودهنگام و پیش‌آگاهی این عارضه شناخته می‌شود. با توجه به این که معیارهای رایج مانند سطح کراتینین سرم معمولاً با تأخیر افزایش می‌یابند، استفاده از الگوریتم‌های یادگیری ماشین که قادر به تحلیل هم‌زمان حجم زیادی از داده‌های بالینی، آزمایشگاهی و تصویربرداری هستند، می‌تواند در شناسایی زودهنگام بیماران در معرض خطر مؤثر باشد [۱۰]. مدل‌هایی مانند شبکه‌های عصبی عمیق (Deep Neural Networks)، جنگل تصادفی (Random Forest)، و الگوریتم XGBoost با دقت قابل‌قبولی برای پیش‌بینی وقوع AKI توسعه یافته‌اند و در مطالعات متعدد عملکرد بهتری نسبت به معیارهای سنتی نشان داده‌اند [۱۱].

الگوریتم DeepMind توسعه‌یافته توسط Google Health توانست وقوع AKI را تا ۴۸ ساعت قبل با دقتی بین ۵۵/۸ درصد تا ۹۰ درصد پیش‌بینی کند [۱۲]. همچنین، استفاده از مدل‌های ترکیبی که شامل متغیرهای حیاتی (Vital Signs)، داروها، سوابق بیماری و داده‌های آزمایشگاهی هستند، نشان داده که می‌توانند در سطوح مختلف بستری، از بخش اورژانس تا بخش مراقبت‌های ویژه به تصمیم‌گیری بالینی کمک کنند. این مدل‌ها علاوه‌بر کاهش تأخیر در تشخیص، با ارائه هشدارهای به‌موقع می‌توانند امکان انجام مداخلات درمانی مؤثرتر را فراهم کنند. باین‌حال، چالش‌هایی همچون تعمیم‌پذیری مدل‌ها در مراکز مختلف، تفسیرپذیری نتایج و نیاز به یکپارچگی با سامانه‌های اطلاعات سلامت بیمارستانی همچنان مطرح هستند.

در دو دهه گذشته مرورهای نظام‌مند متعددی به بررسی کاربرد الگوریتم‌های هوش مصنوعی، به‌ویژه الگوریتم‌های یادگیری ماشین، در پیش‌بینی AKI پرداخته‌اند. مطالعه‌ای در سال ۲۰۲۳ که به بررسی ده مطالعه با بیش از ۲۴۲٬۰۰۰ بیمار پرداخته بود نشان داد که مدل‌های یادگیری عمیق (در بخش مراقبت‌های ویژه AUC برابر ۰/۹۰۷ و در بخش بستری برابر ۰/۷۹۷) به‌طور مشابه با رگرسیون لجستیک (در بخش مراقبت‌های ویژه AUC برابر ۰/۸۷۷ و در بخش بستری برابر ۰/۷۸۹) عمل می‌کنند. باوجوداین، پژوهشگران نسبت به خطر بیش‌برازش (Overfitting) و معماری «جعبه سیاه» این مدل‌ها و همچنین محدودبودن مطالعات به جمعیت‌های متمرکز هشدار دادند [۱۳]. بررسی دیگر، ۴۶ مطالعه را مورد نقد قرار داد و آشکار ساخت که مدل‌هایی مانند جنگل تصادفی و XGBoost محبوب هستند، اما تنها پنج مطالعه دارای اعتبارسنجی خارجی بودند و اغلب فاقد شفافیت لازم در روش‌های تحلیل و اشتراک‌گذاری کد داده‌ها بودند [۱۴].

علاوه‌بر این، استفاده از روش‌های توضیح‌پذیری مدل (Explainable AI) مانند SHAP (SHapley Additive) و LIME (Local Interpretable Model-) (agnostic Explanations) هنوز در مطالعات AKI رایج نیست [۸]. شواهدی نیز وجود دارد که نشان می‌دهد مدل‌های یادگیری ماشین در پیش‌بینی مرگ‌ومیر و نیاز به دیالیز در بیماران مبتلا به AKI با دقت بالا عمل می‌کنند، اما مشکلاتی مانند قابلیت تعمیم‌پذیری و عدم حضور داده‌های چندمرکزی همچنان وجود دارد [۱۵]. به‌طور کلی، این مرورها

بررسی‌ها و تحلیل‌های مطالعات لحاظ شدند تا گستره وسیعی از رویکردهای محاسباتی پوشش داده شود.

با توجه به این مفاهیم، استراتژی جست‌وجو طراحی و در سه پایگاه داده معتبر بین‌المللی شامل PubMed، Scopus، و Web of Science (ISI) اجرا شد. در این جست‌وجو از ترکیب کلیدواژه‌های عنوان، چکیده و کلیدواژه‌ها استفاده شد. جزئیات جست‌وجو در جدول شماره ۱ قابل مشاهده است.

الگوریتم‌های هوش مصنوعی بتوانند در تشخیص زودهنگام و بهبود نتایج بالینی مؤثر باشند.

علاوه بر این، انواع مختلف الگوریتم‌های هوش مصنوعی و یادگیری ماشین که در این حوزه کاربرد دارند، شامل یادگیری ماشین (Machine Learning)، یادگیری عمیق (Deep Learning)، ماشین بردار پشتیبان (Support Vector Machine)، درخت تصمیم (Decision Tree)، شبکه‌های عصبی (Neural Networks)، الگوریتم‌های تقویتی مانند XGBoost و سایر روش‌های مرتبط در

جدول ۱. متن جست‌وجو در سه پایگاه داده

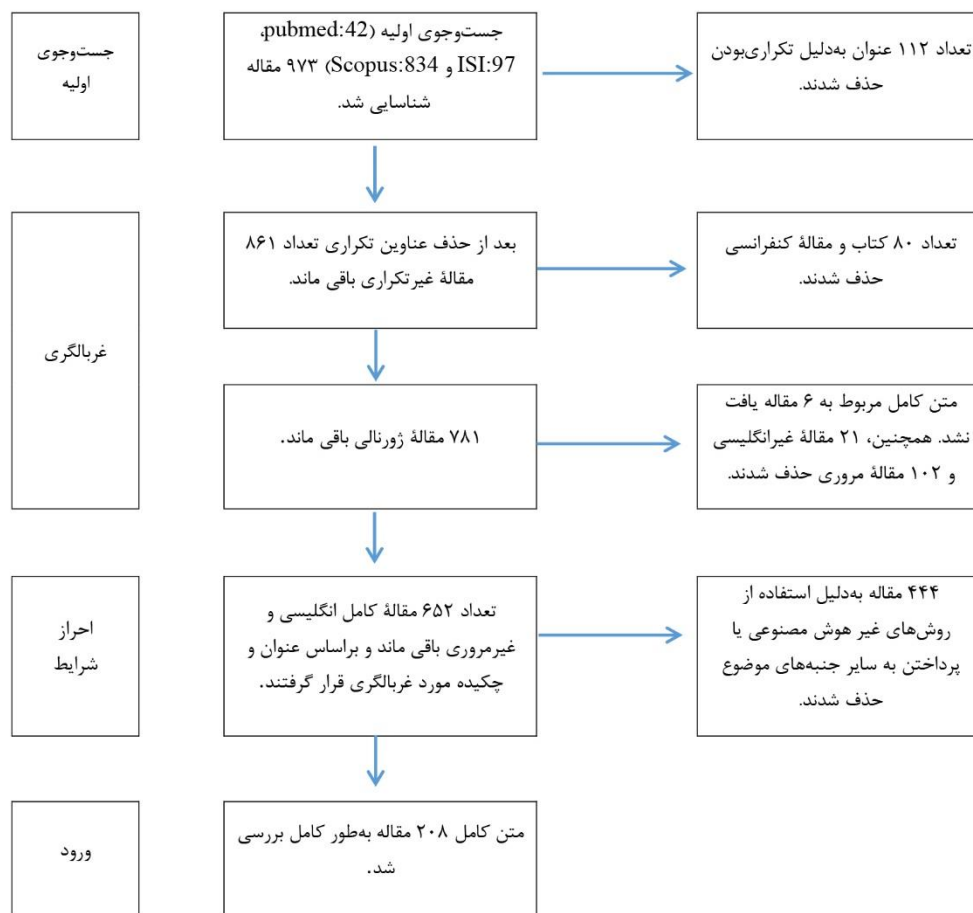
Database	Search Query
PubMed	("Acute Kidney Injuries"[Title/Abstract] OR "Acute Renal Injury"[Title/ABSTRACT] OR "Acute Kidney Failure"[Title/ABSTRACT] OR "Acute Ren. Fail"[Title/ABSTRACT] OR "Acute Kidney Insufficiency"[Title/ABSTRACT] OR "Acute Renal Insufficiency") AND ("forecasting"[Title/ABSTRACT] OR "Prediction"[Title/ABSTRACT] OR "Prognoses") AND ("Support vector machine"[Title/ABSTRACT] OR "neural network"[Title/ABSTRACT] OR "decision tree"[Title/ABSTRACT] OR "Artificial intelligence"[Title/ABSTRACT] OR "Fuzzy"[Title/ABSTRACT] OR "machine learning"[Title/ABSTRACT] OR "Data mining"[Title/ABSTRACT] OR "genetic algorithm"[Title/ABSTRACT] OR "metaheuristic"[Title/ABSTRACT] OR "Meta heuristic"[Title/ABSTRACT] OR "meta-heuristic"[Title/ABSTRACT] OR "evolutionary algorithm"[Title/ABSTRACT] OR "Deep Learning"[Title/ABSTRACT] OR "Hierarchical Learning"[Title/ABSTRACT] OR "Multi-layer Learning"[Title/ABSTRACT] OR "Representation Learning"[Title/ABSTRACT] OR "Reinforcement Learning"[Title/ABSTRACT] OR "Generative Adversarial Networks"[Title/ABSTRACT] OR "Variational Autoencoders"[Title/ABSTRACT] OR "Restricted Boltzmann Machines"[Title/ABSTRACT] OR "Deep Belief Networks"[Title/ABSTRACT] OR "Flow-Based Models"[Title/ABSTRACT] OR "Autoregressive Models"[Title/ABSTRACT] OR "Diffusion Models"[Title/ABSTRACT] OR "Random Forest"[Title/ABSTRACT] OR "Boosting"[Title/ABSTRACT] OR "AdaBoost"[Title/ABSTRACT] OR "Gradient Boosting Machine"[Title/ABSTRACT] OR "XGBoost"[Title/ABSTRACT] OR "LightGBM"[Title/ABSTRACT] OR "CatBoost"[Title/ABSTRACT] OR "Q-Learning"[Title/ABSTRACT] OR "Deep Q-Network"[Title/ABSTRACT] OR "Optimization algorithms"[Title/ABSTRACT] OR "Image Processing"[Title/ABSTRACT] OR "Machine Vision"[Title/ABSTRACT] OR "medical imaging")
Scopus	TITLE-ABS-KEY ("Acute Kidney Injuries" OR "Acute Renal Injury" OR "Acute Kidney Failure" OR "Acute Ren. Fail" OR "Acute Kidney Insufficiency" OR "Acute Renal Insufficiency") AND TITLE-ABS-KEY ("forecasting " OR "Prediction" OR "Prognoses") AND TITLE-ABS-KEY ("Support vector machine" OR "neural network" OR "decision tree" OR "Artificial intelligence" OR "Fuzzy" OR "machine learning" OR "Data mining" OR "genetic algorithm" OR "metaheuristic" OR "Meta heuristic" OR "meta-heuristic" OR "evolutionary algorithm" OR "Deep Learning" OR "Hierarchical Learning" OR "Multi-layer Learning" OR "Representation Learning" OR "Reinforcement Learning" OR "Generative Adversarial Networks" OR "Variational Autoencoders" OR "Restricted Boltzmann Machines" OR "Deep Belief Networks" OR "Flow-Based Models" OR "Autoregressive Models" OR "Diffusion Models" OR "Random Forest" OR "Boosting" OR "AdaBoost" OR "Gradient Boosting Machine" OR "XGBoost" OR "LightGBM" OR "CatBoost" OR "Q-Learning" OR "Deep Q-Network" OR "Optimization algorithms" OR "Image Processing" OR "Machine Vision" OR "medical imaging")
Web of Science	TS= ("Acute Kidney Injuries" OR "Acute Renal Injury" OR "Acute Kidney Failure" OR "Acute Ren. Fail" OR "Acute Kidney Insufficiency" OR "Acute Renal Insufficiency") AND TS= ("forecasting " OR "Prediction" OR "Prognoses") AND TS= ("Support vector machine" OR "neural network" OR "decision tree" OR "Artificial intelligence" OR "Fuzzy" OR "machine learning" OR "Data mining" OR "genetic algorithm" OR "metaheuristic" OR "Meta heuristic" OR "meta-heuristic" OR "evolutionary algorithm" OR "Deep Learning" OR "Hierarchical Learning" OR "Multi-layer Learning" OR "Representation Learning" OR "Reinforcement Learning" OR "Generative Adversarial Networks" OR "Variational Autoencoders" OR "Restricted Boltzmann Machines" OR "Deep Belief Networks" OR "Flow-Based Models" OR "Autoregressive Models" OR "Diffusion Models" OR "Random Forest" OR "Boosting" OR "AdaBoost" OR "Gradient Boosting Machine" OR "XGBoost" OR "LightGBM" OR "CatBoost" OR "Q-Learning" OR "Deep Q-Network" OR "Optimization algorithms" OR "Image Processing" OR "Machine Vision" OR "medical imaging")

مطالعه حذف شدند. این موارد شامل مرورها، نامه به سردبیر، گزارش‌های موردی، مقالات کنفرانسی، پروتکل‌های پژوهشی و پایان‌نامه‌ها بود. همچنین، مقالاتی که تمرکز اصلی آنها بر تشخیص بیماری به جای پیش‌بینی بوده یا صرفاً از روش‌های آماری سنتی بدون بهره‌گیری از الگوریتم‌های هوش مصنوعی استفاده کرده بودند، از تحلیل خارج شدند.

فرایند جست‌وجو و غربالگری مطالعات در قالب نمودار جریان PRISMA در شکل ۱ نمایش داده شده است. این نمودار مراحل شناسایی، حذف مقالات تکراری، غربالگری اولیه، بررسی متن کامل و انتخاب نهایی مطالعات وارد شده به مرور نظام‌مند را به صورت مرحله‌به‌مرحله نشان می‌دهد.

در این مطالعه مرور نظام‌مند، به منظور بررسی استفاده از الگوریتم‌های هوش مصنوعی در پیش‌بینی AKI، تنها مقالات اصیل (Original Research) وارد مرور شدند. مقالات واجد شرایط باید به طور مشخص به پیش‌بینی یا پیش‌آگهی AKI پرداخته و در آنها از حداقل یکی از روش‌های هوش مصنوعی یا یادگیری ماشین، از جمله الگوریتم‌هایی مانند جنگل تصادفی، شبکه‌های عصبی، ماشین بردار پشتیبان، XGBoost، AdaBoost، یا روش‌های یادگیری عمیق استفاده شده باشد. مقالات تنها در صورتی در نظر گرفته شدند که به زبان انگلیسی نگاشته شده باشند و در این مرور محدودیت زمانی برای انتخاب مقالات اعمال نشد.

از سوی دیگر، مقالاتی که شرایط لازم برای ورود را نداشتند، از



شکل ۱. نمودار جریان انتخاب مطالعات براساس چهارچوب PRISMA

علاوه بر این، بررسی شد که مدل طراحی شده از نوع «جعبه سیاه» بوده است یا «جعبه سفید». همچنین، مشخص شد که آیا در فرایند طراحی، تحلیل یا ارزیابی مدل، مشارکت متخصصان بالینی براساس معیارهای مشخص صورت گرفته است یا خیر. معیار مشارکت شامل حضور فعال متخصصان بالینی، مانند پزشکان داخلی، نفرولوژیست‌ها یا سایر گروه‌های مرتبط، در مراحل تأیید متغیرهای ورودی، انتخاب شاخص‌های بالینی مناسب، بازبینی و تأیید قوانین استخراج شده از مدل‌های جعبه سفید و نیز مشارکت در ارزیابی یا اعتبارسنجی بالینی مدل بود. نکته مهم دیگر در تحلیل، تمایز بین مطالعاتی بود که صرفاً براساس معیارهای آماری (مانند حساسیت یا ویژگی) مدل را ارزیابی کرده‌اند و آنهایی که ارزیابی بالینی عملی یا اعتبارسنجی در محیط واقعی نیز انجام داده‌اند. مجموعه این معیارها، امکان تحلیل چندلایه و واقع‌بینانه از کیفیت، قابلیت تفسیر و کاربردپذیری بالینی مدل‌های هوش مصنوعی در زمینه پیش‌بینی AKI را فراهم می‌سازد.

در گام بعدی، با استفاده از داده‌های استخراج شده، تحلیل آماری توصیفی به منظور شناسایی ویژگی‌های کلیدی مطالعات انجام شد. در این مرحله، پراستفاده‌ترین الگوریتم‌های هوش مصنوعی در حوزه پزشکی، به همراه میانگین حجم داده‌ها (تعداد بیماران در هر مطالعه) و میانگین تعداد ویژگی‌های مورد استفاده تعیین و گزارش شد. این تحلیل‌ها، علاوه بر سنجش برتری آماری و دقت پیش‌بینی مدل‌های هوشمند، باعث شناخت بهتر مزایا و محدودیت‌های هر دو

در جستجوی اولیه، تعداد ۹۷۳ عنوان مقاله شناسایی شد. پس از حذف ۱۱۲ مورد تکراری با استفاده از نرم‌افزار EndNote، ۸۶۱ مقاله برای غربالگری عنوان و چکیده وارد شدند. در این مرحله، مقالات نامرتبط، کنفرانسی و کتاب‌ها (۸۰ مورد، ۸/۲ درصد)، مقالات مروری یا غیر انگلیسی یا بدون متن کامل (۱۲۹ مورد، ۱۳/۳ درصد) و مقالات فاقد استفاده از روش‌های هوش مصنوعی یا خارج از هدف مطالعه (۴۴۴ مورد، ۴۵/۶ درصد) حذف شدند. فرایند غربالگری در دو مرحله (عنوان، چکیده و سپس متن کامل) توسط دو پژوهشگر مستقل انجام شد. اختلاف‌نظرها از طریق گفت‌وگو و در صورت نیاز با نفر سوم حل شد. در نهایت، ۲۰۸ مقاله (۲۱/۳ درصد) معیارهای ورود به مطالعه را داشتند و وارد تحلیل نهایی شدند.

هدف از بررسی این مقالات، تحلیل جامع ویژگی‌های مطالعاتی و فنی آنها در زمینه کاربرد الگوریتم‌های هوش مصنوعی در پیش‌بینی AKI بود. برای این منظور، از هر مقاله اطلاعات کلیدی استخراج شد؛ از جمله نوع الگوریتم به کاررفته (مانند جنگل تصادفی، XGBoost، شبکه‌های عصبی، ماشین بردار پشتیبان، درخت تصمیم و غیره)، حجم داده‌های مورد استفاده (تعداد بیماران یا نمونه‌ها)، تعداد ویژگی‌های ورودی (Variables) و شاخص‌های عملکرد مدل شامل AUC، دقت (Accuracy)، حساسیت (Sensitivity)، ویژگی (Specificity)، ارزش اخباری مثبت (Positive Predictive Value) و ارزش اخباری منفی (Negative Predictive Value) محاسبه شد.

رویکرد می‌شوند و نقش مهمی در ارزیابی قابلیت پیاده‌سازی عملی این مدل‌ها در محیط‌های بالینی خواهند داشت. در نهایت، این مطالعه زمینه‌ساز ارائه راهکارهایی جهت ادغام هوش مصنوعی در فرایندهای تصمیم‌گیری پزشکی شده و می‌تواند موجب ارتقاء کیفیت مراقبت‌های بالینی و بهبود نتایج درمانی بیماران شود.

یافته‌ها

در این مطالعه مرور نظام‌مند، تعداد ۲۰۸ مقاله که به پیش‌بینی آسیب حاد کلیوی با استفاده از الگوریتم‌های هوش مصنوعی و یادگیری ماشین پرداخته بودند، مورد ارزیابی قرار گرفت. از این میان، ۵۲ مقاله به کارگیری مدل‌های جعبه‌سفید را گزارش کرده بودند؛ مدلهایی که به دلیل قابلیت تبیین‌پذیری و شفافیت در تصمیم‌گیری، امکان درک

بهرتر توسط کاربران بالینی را فراهم می‌سازند. در مقابل، ۱۵۶ مقاله از مدل‌های جعبه‌سیاه استفاده کرده بودند که اگرچه معمولاً دقت بالاتری دارند، اما تفسیرپذیری آنها محدود است. در جدول شماره ۲، اطلاعات کلیدی مربوط به مقالات جعبه‌سفید شامل نوع الگوریتم، تعداد ویژگی‌های ورودی، اندازه مجموعه داده‌ها (تعداد بیماران)، و شاخص‌های ارزیابی عملکرد مانند دقت، حساسیت، ویژگی، ارزش اخباری، F-score، و AUROC گردآوری شده است. همچنین، مشخصات کامل ۱۵۶ مقاله مربوط به مدل‌های جعبه‌سیاه، به همراه الگوریتم‌های به کار رفته، حجم داده، تعداد ویژگی‌ها و ارزیابی عملکرد، در پیوست مقاله ارائه شده است. این دسته‌بندی و ارائه جامع داده‌ها امکان مقایسه عمیق بین ویژگی‌های فنی و بالینی مدل‌های جعبه‌سفید و جعبه‌سیاه را فراهم می‌سازد.

جدول ۲. مشخصات فنی و بالینی مدل‌های جعبه‌سفید استفاده‌شده در پیش‌بینی آسیب حاد کلیوی

ردیف مرجع	الگوریتم	متخصص بالینی در قوانین بالینی تحقیق	تعداد ویژگی	اندازه مجموعه داده	Accuracy	Sensitivity (recall)	Specificity	Precision (PPV)	NPV	F1-score	AUROC
۱ [۱۷]	Random Forest	✓	×	×	۸۵	۲۶۱	۸۰	۰/۸۷			
۲ [۱۸]	Random Forest	✓	×	×	۶۵۶۴	۲۶		۰/۸۳			
۳ [۱۹]	Decision Tree	✓	✓	×	۱/۱۶۲	۱۵	۹۷	۵۰	۹۹		
۴ [۲۰]	Causal Discovery Method	×	×	×	۹۵۷	۸۹۱	۷۶	۷۴	۷۵	۰/۸۱	
۵ [۲۱]	Random Forest	✓	×	×	۲۹۹۱	۲۸۵	۷۸	۸۱	۸۵	۰/۸۷	
۶ [۲۲]	Random Forest	✓	×	×	۲۵۲	۱۳			۹۷	۰/۸۹	
۷ [۲۳]	Random Forest	✓	×	×	۶۵۳۰	۱۹	۹۲	۶۸	۴۲	۰/۸۸	
۸ [۲۴]	Random Forest	✓	×	×	۲۱۲	۱۳				۰/۸۹	
۹ [۲۵]	Bayesian Network	×	✓	×	۵۵۳۳	۲۷	۶۹	۶۸	۶۸	۰/۷۳	
۱۰ [۲۶]	Bayesian Network	×	✓	×	۶۴۹۵	۱۱	۸۵	۲۴	۵۴	۰/۷۹	
۱۱ [۲۷]	Random Forest	×	×	×	۲۵۴۲	۱۷	۸۱			۸۳	۰/۸۰
۱۲ [۲۸]	Decision Tree	✓	✓	×	۷۸	۲۲				۰/۹۲	
۱۳ [۲۹]	Random Forest	✓	×	×	۲۵۲۹	۳۶۷	۷۹	۷۱	۴۴	۰/۸۴	
۱۴ [۳۰]	Random Forest	✓	✓	×	۴۹۳	۵۰	۷۹	۷۳	۸۱	۰/۸۵	
۱۵ [۳۱]	Decision Tree	✓	✓	×	۹۴۰	۱۵				۰/۷۹	
۱۶ [۳۲]	Random Forest	✓	×	×	۲۱۴	۱۵	۸۹	۹۷		۰/۹۷	
۱۷ [۳۳]	Random Forest	✓	✓	×	۲۴۵۰	۲۶				۰/۹۲	
۱۸ [۳۴]	Random Forest	✓	×	×	۴۲۷۸	۲۲	۷۳	۷۰		۶۴	۰/۸۱
۱۹ [۳۵]	Decision Tree	✓	✓	×	۱۰۹	۴۰				۰/۷۵	
۲۰ [۳۶]	Decision Tree	✓	✓	×	۳۷	۱۴	۸۱				
۲۱ [۳۷]	Random Forest	✓	×	×	۱۰۳۷	۲۳				۰/۸۶	
۲۲ [۳۸]	Random Forest	✓	×	×	۵۴۰۳	۲۰				۰/۸۸	
۲۳ [۳۹]	Decision Tree	✓	✓	×	۱۷۲	۱۸	۷۱				
۲۴ [۴۰]	Random Forest	×	×	×	۶۵۲۲	۳۰	۷۵	۶۷	۹۴	۸۷	۰/۷۵
۲۵ [۴۱]	Random Forest	✓	×	×	۷۲۳	۳۱	۷۵	۶۹	۶۹	۰/۷۵	
۲۶ [۴۲]	Random Forest	✓	×	×	۳۸۸۲	۳۷	۷۵	۷۹	۶۳	۰/۷۷	
۲۷ [۴۳]	Random Forest	✓	×	×	۱۴۳۵	۴۲	۶۷	۶۷	۷۴	۷۰	۰/۸۱
۲۸ [۴۴]	Random Forest	✓	×	×	۲۳۰۰	۱۲۱	۸۰	۷۹	۹۶	۰/۸۶	
۲۹ [۴۵]	Random Forest	✓	×	×	۴۲۹۵	۲۳				۰/۷۳	
۳۰ [۴۶]	Random Forest	✓	✓	×	۴۵۶	۴۶				۰/۸۱	
۳۱ [۴۷]	Random Forest	✓	×	×	۲۶۷۸	۳۹	۸۸	۹۶	۷۳	۰/۹۶	
۳۲ [۴۸]	Random Forest	✓	×	×	۲۱۱۷	۱۳۵	۷۵	۷۵	۳۷	۹۴	۰/۸۲
۳۳ [۴۹]	Random Forest	✓	×	×	۱۳۵	۳۴	۸۲	۵۹	۸۱	۶۸	۰/۸۵
۳۴ [۵۰]	Random Forest	✓	×	×	۱۶۹۴	۶۵				۰/۷۲	
۳۵ [۵۱]	Random Forest	✓	×	×	۸۰۸	۱۵	۹۵	۷۹	۷۶	۹۹	۰/۸۱
۳۶ [۵۲]	Random Forest	✓	×	×	۵۱۹۶	۱۲		۵۷	۸۳	۹۶	۰/۸۰
۳۷ [۵۳]	Random Forest	✓	×	×	۱۴۴۴	۳۰	۸۳	۶۷	۷۴	۰/۸۱	

ردیف مرجع	الگوریتم	متخصص گزارش ارزیابی بالینی در قوانین بالینی تحقیق	تعداد مجموعه داده	تعداد ویژگی	Accuracy	Sensitivity (recall)	Specificity	Precision (PPV)	NPV	F1-score	AUROC
۳۸ [۵۴]	Random Forest	✓	✓	×	۱۶۸۶	۱۲۶					۰/۸۲
۳۹ [۵۵]	Decision Tree	✓	✓	×	۲۴۳	۲۲					۰/۸۱
۴۰ [۵۶]	Decision Tree	✓	×	×	۵۸۲	۱۰	۹۱	۴۴	۸۶		۰/۸۵
۴۱ [۵۷]	Random Forest	×	×	×	۱۷۶۰	۹	۸۲	۸۳			۰/۹۱
۴۲ [۵۸]	Decision Tree	✓	×	×	۴۰۰	۱۸	۹۴	۶۳		۷۵	۰/۹۱
۴۳ [۵۹]	Random Forest	✓	×	×	۱۲۱۳	۳۱	۶۱	۶۹	۷۷	۶۷	۰/۶۹
۴۴ [۶۰]	Random Forest	✓	×	×	۱۲۶۴	۲۲	۶۶	۷۴	۴۹	۵۶	۰/۷۷
۴۵ [۶۱]	Random Forest	✓	✓	×	۳۸۸۲	۱۵	۷۳	۶۵	۴۸	۵۸	۰/۷۵
۴۶ [۶۲]	Random Forest	✓	×	×	۱۷۶۷	۵۶	۱۰۰		۴۶	۶۳	۰/۸۳
۴۷ [۶۳]	Random Forest	✓	×	×	۹۵۷	۳۳	۸۷	۹۰	۷۷	۹۵	۰/۹۱
۴۸ [۶۴]	Random Forest	✓	×	×	۱۷۴۱	۱۵	۸۶	۸۸	۸۸	۸۸	۰/۹۴
۴۹ [۶۵]	Random Forest	✓	×	×	۴۵۹۲	۲۷	۸۱	۶۰			۰/۷۷
۵۰ [۶۶]	Random Forest	✓	×	×	۳۳۶۶	۸۷	۶۷	۶۷	۳۲	۹۰	۰/۷۳
۵۱ [۶۷]	Decision Tree	✓	✓	×	۳۷۶	۴۷	۸۴	۸۲			۰/۸۹
۵۲ [۶۸]	Random Forest	×	×	×	۱۱۱۰	۳۹	۷۴		۸۰	۷۸	۰/۹۰

پراستفاده حفظ شود. در بین الگوریتم‌های جعبه‌سفید، الگوریتم جنگل تصادفی با سهم چشمگیر ۷۵ درصد (۳۹ مقاله از ۵۲ مقاله) رایج‌ترین روش مورد استفاده بوده و پس از آن، درخت تصمیم (Decision Tree) با ۱۹/۲۳ درصد و سایر الگوریتم‌ها با مجموع ۵/۷۷ درصد قرار دارند.

گزارش تفصیلی توزیع الگوریتم‌های جعبه‌سفید و جعبه‌سیاه در مطالعات بررسی شده در جدول شماره ۳ ارائه شده است. در این تحلیل، الگوریتم‌هایی که در کمتر از سه مقاله استفاده شده بودند، تحت عنوان «سایر (Other)» طبقه‌بندی شدند تا تمرکز بر الگوریتم‌های اصلی و

جدول ۳. توزیع فراوانی استفاده از الگوریتم‌های جعبه‌سفید و جعبه‌سیاه در مطالعات بررسی شده

ردیف	نام الگوریتم	تعداد (درصد)
جعبه‌سفید		
۱	Random Forest	۳۹ (۷۵)
۲	Decision Tree	۱۰ (۱۹/۲۳)
۳	Other	۳ (۵/۷۷)
جعبه‌سیاه		
۱	Extreme Gradient Boosting (XGBoost)	۴۳ (۲۷/۵۶)
۲	Deep Neural Network	۱۲ (۷/۶۹)
۳	Gradient Boosting Machine	۱۲ (۷/۶۹)
۴	LightGBM	۱۱ (۷/۰۵)
۵	Support Vector Machine	۱۰ (۶/۴۱)
۶	AdaBoost	۹ (۵/۷۷)
۷	Multi-Layer Perceptron	۸ (۵/۱۳)
۸	Recurrent Neural Network	۸ (۵/۱۳)
۹	Gradient Boosting Decision Tree	۶ (۳/۸۵)
۱۰	Convolutional Neural Network	۴ (۲/۵۶)
۱۱	Deep Forest	۴ (۲/۵۶)
۱۲	LSTM	۳ (۱/۹۲)
۱۳	Other	۲۶ (۱۶/۸)

دارند که عمدتاً برای دقت پیش‌بینی بالا طراحی شده‌اند، درحالی‌که در مدل‌های جعبه‌سفید، اولویت با الگوریتم‌های تبیین‌پذیرتر مانند جنگل تصادفی و درخت تصمیم است که برای استفاده بالینی، شفافیت تصمیم‌گیری بالاتری ارائه می‌دهند.

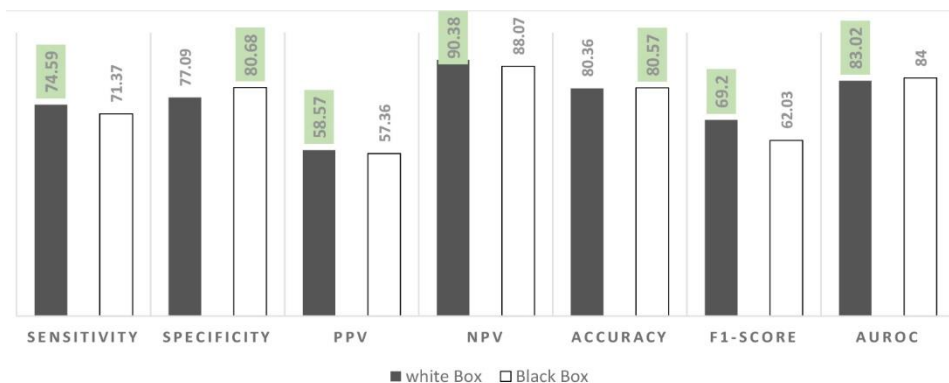
براساس شکل ۲، میانگین شاخص‌های عملکرد بین الگوریتم‌های جعبه‌سفید و جعبه‌سیاه تفاوت چشمگیری نشان نمی‌دهد. به‌طور مشخص، دقت الگوریتم‌های جعبه‌سیاه تنها ۰/۲۱ درصد بیشتر از جعبه‌سفید بوده است و مقدار AUROC نیز با اختلاف ناچیز ۹۸ درصد به‌نفع مدل‌های جعبه‌سیاه گزارش شده است. باین‌حال، الگوریتم‌های

در بخش الگوریتم‌های جعبه‌سیاه، بیشترین فراوانی مربوط به XGBoost با بیش از ۵۰ درصد بود که نشان‌دهنده محبوبیت بالای این الگوریتم در مدل‌سازی‌های دقیق و غیرقابل تبیین است. پس از آن، الگوریتم‌هایی همچون شبکه‌های عصبی عمیق، Gradient Boosting Machine، SVM، AdaBoost، و هر یک با فراوانی بین ۵ تا ۸ درصد، از الگوریتم‌های پرتکرار محسوب می‌شوند.

این توزیع نشان می‌دهد در کاربردهای مربوط به پیش‌بینی AKI، پژوهشگران گرایش به استفاده از الگوریتم‌های یادگیری ماشین (Machine Learning) و شبکه‌های عصبی در مدل‌های جعبه‌سیاه

مدل‌های جعبه‌سفید در معیار F1-Score با اختلاف ۷/۱۷ درصد برتری داشتند. این نتایج نشان می‌دهد که اگرچه عملکرد دو دسته مدل از نظر عددی نزدیک به هم است، اما مدل‌های جعبه‌سفید در برخی شاخص‌ها مانند حساسیت و F1-Score مزیت نسبی دارند.

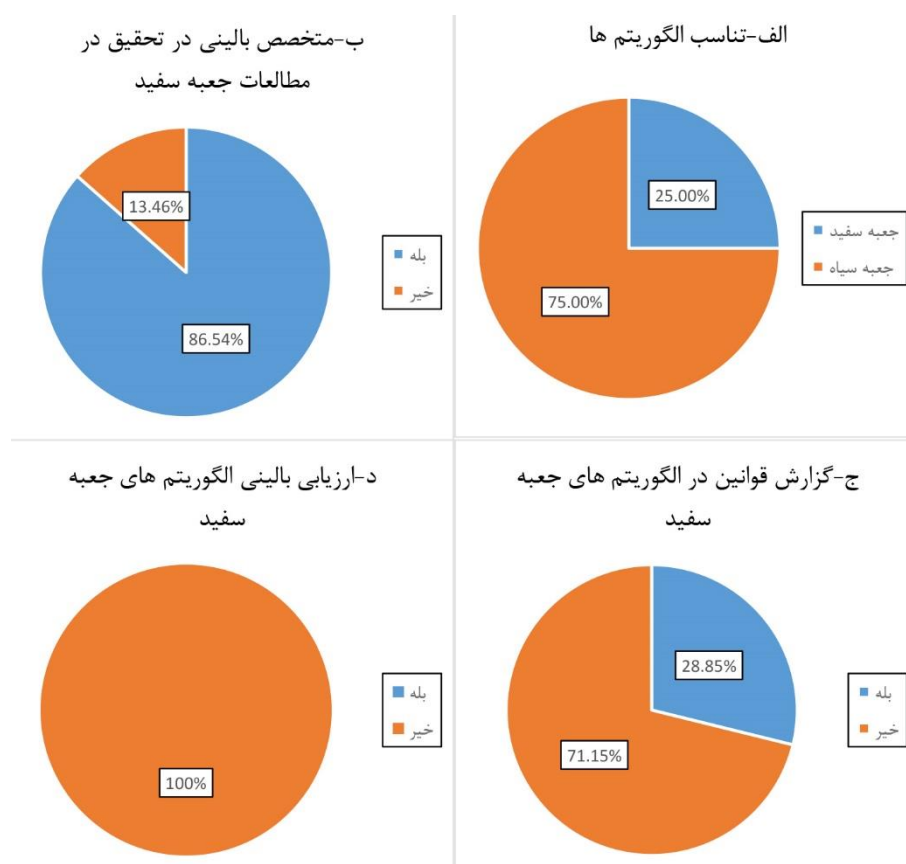
جعبه‌سفید در مقایسه با مدل‌های جعبه‌سیاه دارای میانگین حساسیت بالاتری به میزان ۳/۲۲ درصد و مقدار NPV بالاتر به میزان ۲/۳۱ درصد بوده‌اند. درمقابل، مدل‌های جعبه‌سیاه از نظر ویژگی با اختلاف ۳/۵۹ درصد عملکرد بهتری داشتند. همچنین، مدل‌های جعبه‌سفید با تنها ۱/۲۱ درصد اختلاف نسبت به جعبه‌سیاه بالاتر بود، درحالی‌که



شکل ۲. مقایسه میانگین شاخص‌های عملکرد بین الگوریتم‌های جعبه‌سفید و جعبه‌سیاه در پیش‌بینی آسیب حاد کلیوی

داشته‌اند (ب). باین حال، تنها نزدیک به یک‌سوم از این مطالعات، قوانین تصمیم‌گیری استخراج‌شده از مدل را به‌طور مستقیم گزارش کرده‌اند (ج) و نکته قابل توجه آن که هیچ‌یک از این مدل‌ها تاکنون تحت ارزیابی بالینی واقعی قرار نگرفته‌اند (د).

همان‌طور که در شکل ۳ مشاهده می‌شود، یک‌چهارم از مطالعات بررسی‌شده از الگوریتم‌های جعبه‌سفید استفاده کرده‌اند، درحالی‌که سه‌چهارم دیگر به مدل‌های جعبه‌سیاه اختصاص داشته‌اند (الف). در میان مطالعات مبتنی بر جعبه‌سفید، در بیش از چهارپنجم موارد، متخصصان بالینی در مراحل طراحی، اجرا یا تفسیر مدل‌ها مشارکت



شکل ۳. تحلیل نسبی ویژگی‌های مطالعات هوش مصنوعی در پیش‌بینی آسیب حاد کلیوی شامل نوع مدل، مشارکت متخصص بالینی، گزارش قوانین تبیینی و ارزیابی بالینی.

مجموعه داده (Sample Size) و تعداد ویژگی‌ها (Feature Count) به‌عنوان متغیرهای مستقل در نظر گرفته شدند. در جدول شماره ۴ خروجی رگرسیون برای هر دو مدل مشاهده شده است.

برای بررسی رابطه بین اندازه داده‌ها و تعداد ویژگی‌ها با عملکرد مدل‌های پیش‌بینی آسیب حاد کلیوی، یک تحلیل رگرسیون خطی چندمتغیره به‌طور جداگانه برای مدل‌های جعبه‌سفید و جعبه‌سیاه انجام شد. در این تحلیل‌ها، مقدار AUROC به‌عنوان متغیر وابسته، و اندازه

جدول ۴. نتایج رگرسیون اندازه داده‌ها و تعداد ویژگی با AUROC

مدل	متغیر	ضریب β	انحراف معیار	مقدار t	مقدار P	فاصله اطمینان ۹۵٪
جعبه‌سفید	مقدار ثابت	۰/۸۵۴۲	۰/۰۱۵	۵۶/۱۳	<۰/۰۰۱	[۰/۸۸۵ ... ۰/۸۲۴]
	اندازه داده‌ها	-۰/۰۰۰۱۰۱۱	۰/۰۰۰۰۴۷۷	-۲/۱۲	۰/۰۳۹	[-۰۰۰۰۰۰۵۱۴ ... -۰/۰۰۰۰۱۹۷]
	تعداد ویژگی‌ها	-۰/۰۰۰۰۲۳۹۷	۰/۰۰۰۰۶۸۱	-۰/۳۵	۰/۷۲۶	[-۰/۰۰۰۰۱۱۳ ... -۰/۰۰۰۰۱۶۱]
جعبه‌سیاه	مقدار ثابت	۰/۸۳۳۵	۰/۰۰۷	۱۱۱/۴۸	<۰/۰۰۱	[۰/۸۴۸ ... ۰/۸۱۹]
	اندازه داده‌ها	۰/۰۰۰۰۱۳۱	۰/۰۰۰۰۱۲۹	۱/۰۱	۰/۳۱۳	[-۰/۰۰۰۰۱۲۴ ... -۰/۰۰۰۰۳۸۵]
	تعداد ویژگی‌ها	۰/۰۰۰۰۲۴۶۸	۰/۰۰۰۰۳۴۶	۰/۷۱	۰/۴۷۷	[-۰/۰۰۰۰۴۳۷ ... -۰/۰۰۰۰۹۳۱]

شبکه‌های عصبی عمیق با دقت پیش‌بینی بالاتر برای کاربرد در محیط‌هایی که حساسیت پیش‌بینی اهمیت بیشتری دارد توصیه می‌شوند [۷۱].

باوجود مشارکت نسبتاً گسترده متخصصان بالینی در فرایند توسعه الگوریتم‌های تبیین‌پذیر، نتایج این مطالعه نشان داد که تنها حدود یک‌چهارم از مطالعات جعبه‌سفید توانستند قوانین استخراج‌شده از مدل را گزارش کنند و هیچ‌یک از آنها ارزیابی بالینی در محیط واقعی انجام ندادند. این شکاف در تفسیر خروجی مدل‌ها و ارزیابی عملیاتی می‌تواند یکی از موانع مهم در مسیر پیاده‌سازی بالینی این سامانه‌ها باشد [۷۲].

تحلیل آماری نشان داد در مدل‌های جعبه‌سفید، اندازه نمونه با AUROC رابطه معناداری دارد ($P=۰/۰۳۹$)، درحالی‌که در مدل‌های جعبه‌سیاه، هیچ‌یک از متغیرهای اندازه داده یا تعداد ویژگی‌ها تأثیر آماری معناداری بر عملکرد مدل نداشتند ($P>۰/۰۵$). این نتیجه نشان می‌دهد مدل‌های جعبه‌سیاه احتمالاً بیش از آنکه تحت تأثیر حجم داده باشند به عواملی مانند نوع الگوریتم، کیفیت داده‌ها و تنظیمات مدل وابسته‌اند [۷۳].

باین‌حال، باید توجه داشت که این معیارهای بالینی براساس سال‌ها استفاده در محیط‌های بالینی و داده‌های مکرر، عملکردی پایدار و قابل تکرار از خود نشان داده‌اند. درمقابل، مدل‌های هوش مصنوعی، باوجود دقت بالاتر، اغلب با چالش‌هایی مانند فقدان تکرارپذیری (Reproducibility)، وابستگی به تنظیمات خاص داده یا مرکز درمانی و فقدان ارزیابی مستقل مواجه‌اند. بنابراین، برای آنکه مدل‌های هوش مصنوعی بتوانند جایگزینی مطمئن و بالینی برای روش‌های سنتی باشند، طراحی آنها باید به‌گونه‌ای باشد که امکان تکرارپذیری نتایج در مراکز مختلف و بر روی مجموعه داده‌های مستقل فراهم شود. این امر مستلزم مستندسازی کامل مدل‌ها، شفافیت الگوریتمی و گزارش دقیق مراحل آموزش و ارزیابی است.

نتیجه‌گیری

این مرور نظام‌مند نشان می‌دهد مدل‌های هوش مصنوعی، چه از نوع جعبه‌سفید (مانند درخت تصمیم) و چه جعبه‌سیاه (مانند شبکه‌های عصبی)، قابلیت بالایی در پیش‌بینی آسیب حاد کلیوی با دقت و حساسیت مطلوب دارند. باین‌حال، باوجود مزیت تفسیرپذیری مدل‌های جعبه‌سفید که برای پذیرش بالینی حیاتی است، استفاده از آنها کمتر رواج داشته و شکاف عمده‌ای در زمینه ارزیابی بالینی واقعی،

نتایج تحلیل رگرسیون نشان می‌دهد که در مدل‌های جعبه‌سفید، اندازه داده‌ها دارای رابطه معکوس و معناداری با AUROC بود ($P=۰/۰۳۹$ ، $P=-۰/۰۰۰۰۱۰۱۱$)، به‌گونه‌ای که با افزایش حجم داده‌ها مقدار AUROC به‌طور جزئی کاهش می‌یابد. باین‌حال، تعداد ویژگی‌ها در این مدل‌ها تأثیر معناداری نداشت ($P=۰/۷۲۶$).

درمقابل، در مدل‌های جعبه‌سیاه، هیچ کدام از متغیرهای مستقل (اندازه داده یا تعداد ویژگی) ارتباط معناداری با AUROC نداشتند ($P>۰/۰۵$). همچنین، مقدار ضریب تعیین (R^2) برای مدل جعبه‌سفید برابر با ۰/۰۸۹ و برای مدل جعبه‌سیاه برابر با ۰/۰۱۱ بود، که نشان می‌دهد این متغیرها تنها بخش بسیار کوچکی از تغییرات عملکرد مدل را توضیح می‌دهند.

بحث

یافته‌های این مطالعه حاکی از آن است که مدل‌های هوش مصنوعی، به‌ویژه الگوریتم‌های جعبه‌سفید و جعبه‌سیاه، می‌توانند عملکرد مناسبی در پیش‌بینی آسیب حاد کلیوی داشته باشند. بررسی‌ها نشان داد AUROC گزارش‌شده برای الگوریتم‌های هوش مصنوعی به‌ترتیب تا ۰/۹۴ در مدل‌های جعبه‌سفید و ۰/۹۷ در مدل‌های جعبه‌سیاه می‌رسد، درحالی‌که این مقدار برای معیارهای سنتی اغلب کمتر از ۰/۸۴ بوده است. افزون‌براین، الگوریتم‌هایی نظیر Random Forest و XGBoost بیشترین کاربرد را در مطالعات داشتند که پژوهش‌والگلیانو (Vagliano) و همکاران [۱۴] این مسئله را تأیید کرده است. با این وجود، چالش‌هایی مانند عدم تکرارپذیری [۶۹]، فقدان ارزیابی بالینی [۷۰] و محدودیت در تبیین‌پذیری مدل‌ها، به‌ویژه در الگوریتم‌های جعبه‌سیاه، همچنان مطرح است. بنابراین، برای تحقق کاربرد بالینی مؤثر از این مدل‌ها، طراحی آنها باید به‌گونه‌ای باشد که هم از دقت کافی برخوردار باشند و هم امکان ارزیابی، شفافیت و تکرارپذیری در محیط‌های متنوع بالینی فراهم شود.

در این پژوهش، الگوریتم‌های استفاده‌شده برای پیش‌بینی آسیب حاد کلیوی به دو دسته کلی جعبه‌سفید و جعبه‌سیاه تقسیم شدند. الگوریتم‌های جعبه‌سفید ازجمله جنگل تصادفی و درخت تصمیم باوجود عملکرد قابل قبول، به‌دلیل قابلیت تفسیر بالا، گزینه‌ای مناسب برای استفاده در تصمیم‌گیری‌های پزشکی مبتنی بر شفافیت تلقی می‌شوند. درمقابل، الگوریتم‌های جعبه‌سیاه نظیر XGBoost و

تشکر و قدردانی

نویسندگان مراتب قدردانی و تشکر خود را از معاونت پژوهشی دانشگاه شهید چمران اهواز و همچنین معاونت پژوهشی دانشگاه علوم پزشکی شاهرود به دلیل حمایت‌های علمی و پژوهشی ابراز می‌کنند.

ملاحظات اخلاقی

این مطالعه از نوع مرور نظام‌مند بوده و هیچ‌گونه داده فردی یا اطلاعات مربوط به بیماران انسانی یا حیوانی در آن استفاده نشده است. بنابراین، براساس مقررات اخلاق در پژوهش، اخذ کد اخلاق ضروری نبوده و مطالعه حاضر مغایرتی با اصول اخلاق پژوهش ندارد.

تضاد منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافی وجود ندارد.

References

- Mehta RL, Cerdá J, Burdmann EA, Tonelli M, García-García G, Jha V, et al. International Society of Nephrology's 0by25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. *Lancet*. 2015;385(9987):2616-2643. [DOI: 10.1016/s0140-6736(15)60126-x] [PMID: 25777661]
- Susantitaphong P, Cruz DN, Cerdá J, Abulfaraj M, Alqahtani F, Koulouridis I, et al. World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol*. 2013;8(9):1482-1493. [DOI: 10.2215/cjn.00710113] [PMID: 23744003]
- Chawla LS, Eggers PW, Star RA, Kimmel PL. Acute kidney injury and chronic kidney disease as interconnected syndromes. *N Engl J Med*. 2014;371(1):58-66. [DOI: 10.1056/nejmra1214243] [PMID: 24988558]
- Kheirkhah JA, Salari Z, Dadkhah A, Modallalkar H, Hasandokht TS. Acute kidney injury after cardiac surgery in the North of Iran. *J Nephrophathol*. 2022;11(1):e1. [DOI: 10.34172/jnp.2022.05]
- Hoseini B, Jaafari MR, Golabpour A, Momtazi-Borojeni AA, Karimi M, Eslami S. Application of ensemble machine learning approach to assess the factors affecting size and polydispersity index of liposomal nanoparticles. *Sci Rep*. 2023;13(1):18012. [DOI: 10.1038/s41598-023-43689-4] [PMID: 37865639]
- Ahouz F, Golabpour A. Predicting the incidence of COVID-19 using data mining. *BMC Public Health*. 2021;21(1):1087. [DOI: 10.1186/s12889-021-11058-3] [PMID: 34098928]
- Agharezaei Z, Firouzi R, Hassanzadeh S, Zarei-Ghanavati S, Bahaadinbeigy K, Golabpour A, et al. Computer-aided diagnosis of keratoconus through VAE-augmented images using deep learning. *Sci Rep*. 2023;13(1):20586. [DOI: 10.1038/s41598-023-46903-5] [PMID: 37996439]
- Al-Absi DT, Simsekler MCE, Omar MA, Anwar S. Exploring the role of artificial intelligence in acute kidney injury management: a comprehensive review and future research agenda. *BMC Med Inform Decis Mak*. 2024;24(1):337. [DOI: 10.1186/s12911-024-02758-y] [PMID: 39543611]
- Liu X, Ai S, Yu R, Zhang C, Miao Q. Development and evaluation of a machine learning model for post-surgical acute kidney injury in active infective endocarditis. *Front Cardiovasc Med*. 2024;11:1425275. [DOI: 10.3389/fcvm.2024.1425275] [PMID: 39713217]
- Kaur I, Ali A. A complete study on machine learning algorithms for medical data analysis. *Fog Comput Intell Cloud IoT Syst*. 2024;137-172. [DOI: 10.1002/9781394175345.ch7]
- Luo XQ, Zhang NY, Deng YH, Wang HS, Kang YX, Duan SB. Major adverse kidney events in hospitalized older patients with acute kidney injury: machine learning-based model development and validation study. *J Med Internet Res*. 2025;27:e52786. [DOI: 10.2196/52786] [PMID: 39752664]
- Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. 2019;572(7767):116-119. [DOI: 10.1038/s41586-019-1390-1] [PMID: 31367026]
- Bacci MR, Minczuk CVB, Fonseca FLA. A systematic review of artificial intelligence algorithms for predicting acute kidney injury. *Eur Rev Med Pharmacol Sci*. 2023;27(20):9872-9879. [DOI: 10.26355/eurev.202310_34164] [PMID: 37916354]
- Vagliano I, Chesnaye NC, Leopold JH, Jager KJ, Abu-Hanna A, Schut MC. Machine learning models for predicting acute kidney injury: a systematic review and critical appraisal. *Clin Kidney J*. 2022;15(12):2266-2280. [DOI: 10.1093/ckj/sfac181] [PMID: 36381375]
- Hu C, Tan Q, Zhang Q, Li Y, Wang F, Zou X, et al. Application of interpretable machine learning for early prediction of prognosis in acute kidney injury. *Comput Struct Biotechnol J*. 2022;20:2861-2870. [DOI: 10.1016/j.csbj.2022.06.003] [PMID: 35765651]
- European Commission. European approach to artificial intelligence. 2024. [Link]
- Zacharias HU, Hochrein J, Vogl FC, Schley G, Mayer F, Jeleazcov C, et al. Identification of plasma metabolites prognostic of acute kidney injury after cardiac surgery with cardiopulmonary bypass. *J Proteome Res*. 2015;14(7):2897-2905. [DOI: 10.1021/acs.jproteome.5b00219] [PMID: 26036910]
- Sanchez-Pinto LN, Venable LR, Fahrenbach J, Churpek MM. Comparison of variable selection methods for clinical predictive modeling. *Int J Med Inform*. 2018;116:10-17. [DOI: 10.1016/j.ijmedinf.2018.05.006] [PMID: 29887230]
- Coeman L, Wyffels P, Van Limmen J, Wouters P. Predicting perioperative acute kidney injury in liver surgery. *Acta Anaesthesiol Belg*. 2021;72:63-70. [Link]
- Chen W, Hu Y, Zhang X, Wu L, Liu K, He J, et al. Causal risk factor discovery for severe acute kidney injury using electronic health records. *BMC Med Inform Decis Mak*. 2018;18:13. [DOI: 10.1186/s12911-018-0597-7] [PMID: 29589567]
- Adhikari L, Ozrazgat-Baslanti T, Ruppert M, Madushani R, Paliwal S, Hashemighouchani H, et al. Improved predictive models for acute kidney injury with IDEA: intraoperative data embedded analytics. *PLoS One*. 2019;14(4):e0215324. [DOI: 10.1371/journal.pone.0214904] [PMID: 30947282]
- Flechet M, Falini S, Bonetti C, Guiza F, Schetz M, Van den Berghe G, et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care*. 2019;23(1):282. [DOI: 10.1186/s13054-019-2563-x] [PMID: 31420056]

23. Chiofolo C, Chbat N, Ghosh E, Eshelman L, Kashani K. Automated continuous acute kidney injury prediction and surveillance: a random forest model. *Mayo Clin Proc.* 2019;94(5):783-792. [DOI: 10.1016/j.mayocp.2019.02.009] [PMID: 31054606]
24. Zhou CY, Wang R, Jiang WJ, Zhu JM, Liu YM, Zheng J, et al. Machine learning for the prediction of acute kidney injury and paraplegia after thoracoabdominal aortic aneurysm repair. *J Card Surg.* 2020;35(1):89-99. [DOI: 10.1111/jocs.14317] [PMID: 31765025]
25. Li Y, Xu JR, Wang YM, Zhang YL, Jiang WH, Shen B, et al. A novel machine learning algorithm, Bayesian networks model, to predict high-risk patients with cardiac surgery-associated acute kidney injury. *Clin Cardiol.* 2020;43(7):752-761. [DOI: 10.1002/clc.23377] [PMID: 32400109]
26. Li Y, Chen X, Shen Z, Wang Y, Hu J, Zhang Y, et al. Prediction models for acute kidney injury in patients with gastrointestinal cancers: a real-world study based on Bayesian networks. *Ren Fail.* 2020;42(1):869-876. [DOI: 10.1080/0886022x.2020.1810068] [PMID: 32838613]
27. Morid MA, Liu S, Del Fiore G, Facelli JC, Bray BE, Abdelrahman S. Temporal pattern detection to predict adverse events in critical care: a case study with acute kidney injury. *JMIR Med Inform.* 2020;8(3):e14272. [DOI: 10.2196/14272] [PMID: 32181753]
28. Jerin A, Mosa OF, Kališnik JM, Zibert J, Skitek M. Serum klotho as a marker for early diagnosis of acute kidney injury after cardiac surgery. *J Med Biochem.* 2020;39(2):233-240. [DOI: 10.2478/jomb-2019-0024] [PMID: 33033444]
29. Datta S, Loftus TJ, Ruppert MM, Giordano C, Upchurch GR, Rashidi P, et al. Added value of intraoperative data for predicting postoperative complications: the MySurgeryRisk PostOp extension. *J Surg Res.* 2020;254:350-363. [DOI: 10.1016/j.jss.2020.05.007] [PMID: 32531520]
30. He ZL, Zhou JB, Liu ZK, Dong SY, Zhang YT, Shen T, et al. Application of machine learning models for predicting acute kidney injury following donation after cardiac death liver transplantation. *Hepatobiliary Pancreat Dis Int.* 2021;20(3):222-231. [DOI: 10.1016/j.hbpd.2021.02.001] [PMID: 33726966]
31. Wong WEJ, Chan SP, Yong JK, Tham YYS, Lim JRG, Sim MA, et al. Assessment of acute kidney injury risk using a machine-learning guided generalized structural equation model: a cohort study. *BMC Nephrol.* 2021;22(1):63. [DOI: 10.1186/s12882-021-02238-9] [PMID: 33618695]
32. Cui H, Shu S, Li Y, Yan X, Chen X, Chen Z, et al. Plasma metabolites-based prediction in cardiac surgery-associated acute kidney injury. *J Am Heart Assoc.* 2021;10(22):e021825. [DOI: 10.1161/jaha.121.021825] [PMID: 34719239]
33. Dong JF, Xue Q, Chen T, Zhao YY, Fu H, Guo WY, et al. Machine learning approach to predict acute kidney injury after liver surgery. *World J Clin Cases.* 2021;9(36):11255-11264. [DOI: 10.12998/wjcc.v9.i36.11255] [PMID: 35071556]
34. Huang CT, Chang RC, Tsai YL, Pai KC, Wang TJ, Hsu CT, et al. Entropy-based time window features extraction for machine learning to predict acute kidney injury in ICU. *Appl Sci.* 2021;11(14):6523. [DOI: 10.3390/app11146364] [PMID: 34015009]
35. Xin W, Yi W, Liu H, Li H, Li D, Ma Y, et al. Early prediction of acute kidney injury after liver transplantation by scoring system and decision tree. *Ren Fail.* 2021;43(1):1137-1145. [DOI: 10.1080/0886022x.2021.1945462] [PMID: 34261422]
36. Herzog AL, von Jouanne-Diedrich HK, Wanner C, Weismann D, Schlesinger T, Meybohm P, et al. COVID-19 and the kidney: a retrospective analysis of 37 critically ill patients using machine learning. *PLoS One.* 2021;16(5):e0251930. [DOI: 10.1371/journal.pone.0251932] [PMID: 34015009]
37. Bajaj JS, Garcia-Tsao G, Reddy KR, O'Leary JG, Vargas HE, Lai JC, et al. Admission urinary and serum metabolites predict renal outcomes in hospitalized patients with cirrhosis. *Hepatology.* 2021;74(5):2699-2713. [DOI: 10.1002/hep.31907] [PMID: 34002868]
38. Scanlon LA, O'Hara C, Garbett A, Barker-Hewitt M, Barriuso J. Developing an agnostic risk prediction model for early AKI detection in cancer patients. *Cancers (Basel).* 2021;13(16):4182. [DOI: 10.3390/cancers13164182] [PMID: 34439336]
39. Karakaya E, Akdur A, Aydoğan C, Türk E, Sayin CB, Ayyavozlu Soy E, et al. A model for acute kidney injury in severe burn patients. *Burns.* 2022;48(1):69-77. [DOI: 10.1016/j.burns.2021.04.004] [PMID: 33879373]
40. Petrosyan Y, Mesana TG, Sun LY. Prediction of acute kidney injury risk after cardiac surgery using a hybrid machine learning algorithm. *BMC Med Inform Decis Mak.* 2022;22(1):137. [DOI: 10.1186/s12911-022-01859-w] [PMID: 35585624]
41. Lazebnik T, Bahouth Z, Bunimovich-Mendrazitsky S, Halachmi S. Predicting acute kidney injury following open partial nephrectomy treatment using SAT-pruned explainable machine learning model. *BMC Med Inform Decis Mak.* 2022;22(1):133. [DOI: 10.1186/s12911-022-01877-8] [PMID: 35578278]
42. Cai D, Xiao T, Zou A, Mao L, Chi B, Wang Y, et al. Predicting acute kidney injury risk in acute myocardial infarction patients: an artificial intelligence model using medical information mart for intensive care databases. *Front Cardiovasc Med.* 2022;9:964894. [DOI: 10.3389/fcvm.2022.964894] [PMID: 36158815]
43. Ejmalian A, Aghaei A, Nabavi S, Darabad MA, Tajbakhsh A, Abin AA, et al. Prediction of acute kidney injury after cardiac surgery using interpretable machine learning. *Anesthesiol Pain Med.* 2022;12(4):e127140. [DOI: 10.5812/aapm-127140] [PMID: 36937087]
44. Hu Y, Liu K, Ho K, Riviello D, Brown J, Chang AR, et al. A simpler machine learning model for acute kidney injury risk stratification in hospitalized patients. *J Clin Med.* 2022;11(19):5713688. [DOI: 10.3390/jcm11195688] [PMID: 36233556]
45. Li D, Jiang H, Yang X, Lin M, Gao M, Chen Z, et al. An online pre-procedural nomogram for the prediction of contrast-associated acute kidney injury in patients undergoing coronary angiography. *Front Med.* 2022;9:839856. [DOI: 10.3389/fmed.2022.839856] [PMID: 35360720]
46. Li XS, Wang ZY, Xuan H, Chu XQ, Kai P, Chen SS, et al. Prediction model of acute kidney injury after different types of acute aortic dissection based on machine learning. *Front Cardiovasc Med.* 2022;9:984772. [DOI: 10.3389/fcvm.2022.984772] [PMID: 36211563]
47. Liu WT, Liu XQ, Jiang TT, Wang MY, Huang Y, Huang YL, et al. Using a machine learning model to predict the development of acute kidney injury in patients with heart failure. *Front Cardiovasc Med.* 2022;9:911987. [DOI: 10.3389/fcvm.2022.911987] [PMID: 36176988]
48. Ren Y, Loftus TJ, Datta S, Ruppert MM, Guan Z, Miao S, et al. Performance of a machine learning algorithm using electronic health record data to predict postoperative complications and report on a mobile platform. *JAMA Netw Open.* 2022;5(4):e2211973. [DOI: 10.1001/jamanetworkopen.2022.11973] [PMID: 35576007]
49. Xue X, Liu ZY, Xue T, Chen W, Chen X. Machine learning for the prediction of acute kidney injury in patients after cardiac surgery. *Front Surg.* 2022;9:946610. [DOI: 10.3389/fsurg.2022.946610] [PMID: 36157418]
50. Yang L, Gabriel N, Hernandez I, Vouri SM, Kimmel SE, Bian J, et al. Identifying patients at risk of acute kidney injury among Medicare beneficiaries with type 2 diabetes initiating SGLT2 inhibitors: a machine learning approach. *Front Pharmacol.* 2022;13:834743. [DOI: 10.3389/fphar.2022.834743] [PMID: 35359843]
51. Peng C, Yang F, Li L, Peng L, Yu J, Wang P, et al. A machine learning approach for the prediction of severe acute kidney injury following traumatic brain injury. *Neurocrit Care.* 2023;38(2):335-344. [DOI: 10.1007/s12028-022-01606-z] [PMID: 36195818]
52. Huang S, Teng Y, Du J, Zhou X, Duan F, Feng C. Internal and external validation of machine learning-assisted prediction

- models for mechanical ventilation-associated severe acute kidney injury. *Aust Crit Care*. 2023;36(4):604-612. [DOI: [10.1016/j.aucc.2022.06.001](https://doi.org/10.1016/j.aucc.2022.06.001)] [PMID: [35842332](https://pubmed.ncbi.nlm.nih.gov/35842332/)]
53. Cox M, Panagides JC, Di Capua J, Dua A, Kalva S, Kalpathy-Cramer J, et al. An interpretable machine learning model for the prevention of contrast-induced nephropathy in patients undergoing lower extremity endovascular interventions for peripheral arterial disease. *Clin Imaging*. 2023;101:1-7. [DOI: [10.1016/j.clinimag.2023.05.011](https://doi.org/10.1016/j.clinimag.2023.05.011)] [PMID: [37247523](https://pubmed.ncbi.nlm.nih.gov/37247523/)]
 54. Shao J, Liu F, Ji S, Song C, Ma Y, Shen M, et al. Development, external validation, and visualization of machine learning models for predicting occurrence of acute kidney injury after cardiac surgery. *Rev Cardiovasc Med*. 2023;24(8):229. [DOI: [10.31083/j.rcm2408229](https://doi.org/10.31083/j.rcm2408229)] [PMID: [39076716](https://pubmed.ncbi.nlm.nih.gov/39076716/)]
 55. Komaru Y, Isshiki R, Matsuura R, Hamasaki Y, Nangaku M, Doi K. Application of urinary biomarkers for diagnosing acute kidney injury in critically ill patients without baseline renal function data. *J Crit Care*. 2023;77:154312. [DOI: [10.1016/j.jccrc.2023.154312](https://doi.org/10.1016/j.jccrc.2023.154312)] [PMID: [37058992](https://pubmed.ncbi.nlm.nih.gov/37058992/)]
 56. Wu M, Jiang X, Du K, Xu Y, Zhang W. Ensemble machine learning algorithm for predicting acute kidney injury in patients admitted to the neurointensive care unit following brain surgery. *Sci Rep*. 2023;13(1):6705. [DOI: [10.1038/s41598-023-33930-5](https://doi.org/10.1038/s41598-023-33930-5)] [PMID: [37185782](https://pubmed.ncbi.nlm.nih.gov/37185782/)]
 57. Alfieri F, Ancona A, Tripepi G, Rubeis A, Arjoldi N, Finazzi S, et al. Continuous and early prediction of future moderate and severe acute kidney injury in critically ill patients: development and multicentric multinational external validation of a machine learning model. *PLoS One*. 2023;18(3):e0287398. [DOI: [10.1371/journal.pone.0287398](https://doi.org/10.1371/journal.pone.0287398)]
 58. Choi H, Lee JY, Sul Y, Kim S, Ye JB, Lee JS, et al. Comparing machine learning and logistic regression for acute kidney injury prediction in trauma patients: a retrospective observational study at a single tertiary medical center. *Medicine (Baltimore)*. 2023;102(33):e34847. [DOI: [10.1097/md.00000000000034847](https://doi.org/10.1097/md.00000000000034847)] [PMID: [37603521](https://pubmed.ncbi.nlm.nih.gov/37603521/)]
 59. She S, Shen Y, Luo K, Zhang X, Luo C. Prediction of acute kidney injury in intracerebral hemorrhage patients using machine learning. *Neuropsychiatr Dis Treat*. 2023;19:2765-2773. [DOI: [10.2147/ndt.s439549](https://doi.org/10.2147/ndt.s439549)] [PMID: [38106359](https://pubmed.ncbi.nlm.nih.gov/38106359/)]
 60. Tong YF, Niu XG, Liu F. Prediction of acute kidney injury after extracorporeal cardiac surgery by machine learning algorithms. *Heart Surg Forum*. 2023;26(5):E537-E551. [DOI: [10.59958/hsf.5673](https://doi.org/10.59958/hsf.5673)] [PMID: [37920093](https://pubmed.ncbi.nlm.nih.gov/37920093/)]
 61. Wei J, Cai D, Xiao T, Chen Q, Zhu W, Gu Q, et al. Artificial intelligence algorithms permit rapid acute kidney injury risk classification of patients with acute myocardial infarction. *Heliyon*. 2024;10(16):e36051. [DOI: [10.1016/j.heliyon.2024.e36051](https://doi.org/10.1016/j.heliyon.2024.e36051)] [PMID: [39224361](https://pubmed.ncbi.nlm.nih.gov/39224361/)]
 62. Hou X, Zhang K, Liu T, Xu S, Zheng J, Li Y, et al. Prediction of acute kidney injury following isolated coronary artery bypass grafting in heart failure patients with preserved ejection fraction using machine learning with a novel nomogram. *Rev Cardiovasc Med*. 2024;25(2):43. [DOI: [10.31083/j.rcm2502043](https://doi.org/10.31083/j.rcm2502043)] [PMID: [39077338](https://pubmed.ncbi.nlm.nih.gov/39077338/)]
 63. Hu J, Xu J, Li M, Jiang Z, Mao J, Feng L, et al. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study. *eClinicalMedicine*. 2024;68:1023409. [DOI: [10.1016/j.eclinm.2023.102409](https://doi.org/10.1016/j.eclinm.2023.102409)] [PMID: [38273888](https://pubmed.ncbi.nlm.nih.gov/38273888/)]
 64. Omar ED, Mat H, Karim AZA, Sanaudi R, Ibrahim FH, Omar MA, et al. Comparative analysis of logistic regression, gradient boosted trees, support vector machine, and random forest algorithms for prediction of acute kidney injury requiring dialysis after cardiac surgery. *Int J Nephrol Renovasc Dis*. 2024;17:197-204. [DOI: [10.2147/ijnrd.s461028](https://doi.org/10.2147/ijnrd.s461028)] [PMID: [39070075](https://pubmed.ncbi.nlm.nih.gov/39070075/)]
 65. Behnouth AH, Shariatnia MM, Khalaji A, Asadi M, Yaghoobi A, Rezaee M, et al. Predictive modeling for acute kidney injury after percutaneous coronary intervention in patients with acute coronary syndrome: a machine learning approach. *Eur J Med Res*. 2024;29(1):76. [DOI: [10.1186/s40001-024-01675-0](https://doi.org/10.1186/s40001-024-01675-0)] [PMID: [38268045](https://pubmed.ncbi.nlm.nih.gov/38268045/)]
 66. Tschoellitsch T, Moser P, Maletzky A, Seidl P, Bock C, Roland T, et al. Potential predictors for deterioration of renal function after transfusion. *Anesth Analg*. 2024;138(3):645-654. [DOI: [10.1213/ane.0000000000006720](https://doi.org/10.1213/ane.0000000000006720)] [PMID: [38364244](https://pubmed.ncbi.nlm.nih.gov/38364244/)]
 67. Wang R, Zhang J, He M, Xu J. Classification and regression tree predictive model for acute kidney injury in traumatic brain injury patients. *Ther Clin Risk Manag*. 2024;20:139-149. [DOI: [10.2147/tcrm.s435281](https://doi.org/10.2147/tcrm.s435281)] [PMID: [38410117](https://pubmed.ncbi.nlm.nih.gov/38410117/)]
 68. Zeng Z, Tian X, Li L, Diao Y, Zhang T. An interpretable machine learning model to predict off-pump coronary artery bypass grafting-associated acute kidney injury. *Adv Clin Exp Med*. 2024;33(5):473-481. [DOI: [10.17219/acem/169609](https://doi.org/10.17219/acem/169609)] [PMID: [37593773](https://pubmed.ncbi.nlm.nih.gov/37593773/)]
 69. Ahmed H, Tchoua R, Lofstead J. Measuring reproducibility of machine learning methods for medical diagnosis. In: *Proc Fourth Int Conf Transdisciplinary AI (TransAI)*; 2022 Sep 19–21. p. 1-8. [DOI: [10.1109/TransAI54797.2022.00008](https://doi.org/10.1109/TransAI54797.2022.00008)]
 70. Park SH, Han K, Jang HY, Park JE, Lee JG, Kim DW, et al. Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis. *Radiology*. 2023;306(1):20-31. [DOI: [10.1148/radiol.220182](https://doi.org/10.1148/radiol.220182)] [PMID: [36346314](https://pubmed.ncbi.nlm.nih.gov/36346314/)]
 71. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017;30:4765-4774. [DOI: [10.48550/arXiv.1705.07874](https://doi.org/10.48550/arXiv.1705.07874)]
 72. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. *Proc Mach Learn Healthc Conf*. 2019:359-380. [DOI: [10.48550/arXiv.1905.05134](https://doi.org/10.48550/arXiv.1905.05134)]
 73. Van Calster B, Wynants L, Timmerman D, Steyerberg EW, Collins GS. Predictive analytics in health care: how can we know it works? *J Am Med Inform Assoc*. 2019;26(12):1651-1654. [DOI: [10.1093/jamia/ocz130](https://doi.org/10.1093/jamia/ocz130)] [PMID: [31373357](https://pubmed.ncbi.nlm.nih.gov/31373357/)]