

بررسی عملکرد هوش مصنوعی در تشخیص کرونا: یک مرور نظام مند

آزاده آبکار^۱، مهدی محرابی^۲، امین گلاب پور^۳، محمد امین شایگان^۴

مقاله مروری

چکیده

مقدمه: تشخیص زودهنگام بیماری کووید-۱۹ نقش مهمی در کاهش مرگ و میر بیماران دارد. با توجه به توانمندی های هوش مصنوعی، این فناوری می تواند به عنوان یکی از روش های مؤثر در تشخیص بیماری کووید-۱۹ مورد استفاده قرار گیرد. این مطالعه مروری نظام مند با هدف بررسی عملکرد هوش مصنوعی در این حوزه انجام شده است.

روش بررسی: با استفاده از ترکیب کلیدواژه های مرتبط، تمامی مقالات مربوط به کاربرد هوش مصنوعی در تشخیص بیماری کووید-۱۹ از پایگاه های داده پابمد، اسکوپوس و وب او ساینس استخراج شدند. این مقالات از جنبه های مختلفی از جمله نوع داده، نوع الگوریتم، پارامترهای ارزیابی، ارائه قوانین بالینی، و ارزیابی بالینی الگوریتم ها مورد بررسی قرار گرفتند. سپس با استفاده از مقالات مروری و متاآنالیز، مشخص شد که این الگوریتم ها تا چه حد در گایدلاین های بالینی مورد استفاده قرار گرفته اند.

یافته ها: بررسی ۱۴۳ مقاله وارد شده به مطالعه نشان داد که داده های ورودی این مقالات به سه دسته رکورد های اطلاعاتی، تصاویر بالینی، و پردازش صدا تقسیم می شوند. بیش از ۹۶ درصد از این الگوریتم ها به عنوان الگوریتم های جعبه سیاه شناخته می شوند و فاقد ارزیابی بالینی توسط متخصصان حوزه بالین هستند. با این حال، میانگین عملکرد تمامی الگوریتم ها بالای ۹۰ درصد گزارش شده است. همچنین، از بررسی ۴۵۱ مقاله مروری و متاآنالیز در این مطالعه مشخص شد که هیچ یک از الگوریتم ها در چارچوب بالینی مورد بررسی قرار نگرفته اند و تنها عملکرد الگوریتم ها ارزیابی شده است.

نتیجه گیری: مرور مقالات نشان داد هوش مصنوعی در تشخیص کووید-۱۹ عملکرد مطلوبی دارد، اما چالش هایی مانند عدم توضیح پذیری مدل ها، وابستگی به داده های آموزشی و نبود ارزیابی های بالینی گسترده وجود دارد. توسعه الگوریتم های جعبه سفید و بهینه سازی مدل ها برای پذیرش بالینی ضروری است.

واژه های کلیدی: COVID-19، تشخیص، داده کاوی، هوش مصنوعی

پیام کلیدی: هوش مصنوعی در تشخیص کووید-۱۹ دقت بالایی دارد، اما چالش هایی مانند عدم شفافیت مدل ها، وابستگی به داده های آموزشی، و نبود ارزیابی های بالینی گسترده، مانع از پذیرش کامل آن در محیط های درمانی شده است. توسعه الگوریتم های تفسیرپذیر (جعبه سفید) و انجام مطالعات بالینی معتبر، برای افزایش کاربرد عملی این فناوری ضروری است.

تاریخ انتشار: ۱۴۰۳/۴/۱۰

پذیرش مقاله: ۱۴۰۳/۳/۲۹

دریافت مقاله: ۱۴۰۳/۲/۱۲

ارجاع: آبکار، آزاده، محرابی، مهدی، گلاب پور، امین، شایگان، محمد امین. **بررسی عملکرد هوش مصنوعی در تشخیص کرونا: یک مرور نظام مند.** مدیریت اطلاعات سلامت ۲۱:۱۴۰۳ (۲): ۱۱۱-۱۰۳.

مقدمه

استفاده از هوش مصنوعی در تشخیص و مدیریت بیماری های واگیر، به ویژه کووید-۱۹، اهمیت بسیار بالایی دارد. این فناوری قادر است با سرعت و دقت تصاویر پرتوی، مانند اسکن های ریه، را تحلیل کرده و نشانه های التهاب ریوی را شناسایی کند (۱). همچنین، با تحلیل داده های اپیدمیولوژیکی، هوش مصنوعی می تواند شیوع بیماری را پیش بینی کرده و مقامات بهداشتی را در اتخاذ تدابیر مناسب یاری دهد. هوش مصنوعی، با تجزیه و تحلیل دقیق داده های پزشکی و علائم بالینی بیماران، می تواند به تشخیص دقیق تر بیماری کمک کند.

علاوه بر این، در تصمیم گیری های پیچیده مانند تعیین احتمال ابتلا به کووید-۱۹ بر اساس علائم بالینی و داده های پزشکی، نیز می تواند مورد استفاده قرار گیرد. همچنین، هوش مصنوعی با تحلیل داده های آزمایش های بالینی و مولکولی به توسعه داروها و واکسن ها علیه کووید-۱۹ کمک کند. به طور کلی، استفاده از هوش مصنوعی به عنوان ابزاری قدرتمند در تشخیص، پیش بینی و مدیریت کووید-۱۹ و دیگر بیماری های واگیر، پیشرفت های چشمگیری را در علم پزشکی به ارمغان می آورد و می تواند در کاهش شیوع و بهبود بهداشت جامعه نقش مهمی ایفا کند (۲، ۳).

استفاده از داده کاوی در تشخیص کووید-۱۹ به عنوان یکی از مهم ترین چالش های بهداشتی و پزشکی اهمیت بسیاری دارد. اطلاعات حجیم و پیچیده ای که درباره این بیماری جمع آوری می شود، با بهره گیری از تکنیک ها و الگوریتم های داده کاوی می تواند منجر به شناسایی الگوهای معنادار و مفاهیم کلیدی در میان داده ها شود. تحلیل علائم بالینی، نتایج آزمایشگاهی و تعاملات افراد، امکان کشف الگوهای مشخص در رفتار و شیوع بیماری را فراهم می کند. این اطلاعات نه تنها امکان پیش بینی شیوع، تشخیص زودهنگام، شناسایی مناطق پرخطر را فراهم می کنند، بلکه در ارزیابی تدابیر بهداشتی نیز مؤثر هستند.

۱. دانشجوی دکتری تخصصی گروه کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

۲. استادیار، گروه کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

۳. استادیار، گروه فناوری اطلاعات سلامت دانشگاه علوم پزشکی شاهرود، شاهرود، ایران

۴. دانشیار، گروه کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

نویسنده طرف مکاتبه: مهدی محرابی، استادیار، گروه کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

Email: mehdimehrabi1975@gmail.com

با بهره‌گیری از داده کاوی می‌توان تصمیم‌گیری‌های دقیق‌تری در زمینه مدیریت منابع پزشکی، تخصیص واکسن‌ها و تعیین اولویت‌های درمانی اتخاذ کرد. در نتیجه، ترکیب داده کاوی با تخصص‌های پزشکی می‌تواند تأثیر قابل توجهی در بهبود تشخیص، کنترل و مدیریت کوئید-۱۹ و بیماری‌های واگیر مشابه داشته باشد (۲، ۴).

در این پژوهش، مطالعات مربوط به داده کاوی برای تشخیص کووید-۱۹ بررسی می‌شوند تا میزان قابلیت عملیاتی آن‌ها ارزیابی گردد. داده کاوی به دو دسته الگوریتم‌های «جعبه سفید» و «جعبه سیاه» تقسیم می‌شود. الگوریتم‌های جعبه سفید به صورت شفاف عمل کرده و درک عمیقی از نحوه تصمیم‌گیری ارائه می‌دهند، در حالی که الگوریتم‌های جعبه سیاه، با وجود دقت بالا، عملکرد داخلی خود را پنهان می‌کنند. مشکلات اصلی الگوریتم‌های جعبه سیاه شامل عدم شفافیت، وابستگی به داده‌های آموزشی و عدم اطمینان از عملکرد در شرایط جدید است که می‌تواند منجر به تصمیم‌گیری‌های نادرست شود (۵، ۶).

الگوریتم‌های جعبه سفید به دلیل دقت و تفسیرپذیری بالا، به پزشکان امکان می‌دهند که ارتباط دقیق‌تری میان علائم، آزمایش‌ها و تصمیم‌گیری‌های پزشکی برقرار کنند و در تمامی مراحل تشخیص و درمان مشارکت فعال داشته باشند، در عین حال، این الگوریتم‌ها توانایی تنظیم درمان‌ها بر اساس شرایط خاص هر بیمار را فراهم می‌کنند. یکی از روش‌های مهم این الگوریتم‌ها، استخراج قوانین بالینی از داده‌ها است که نه تنها به پزشکان بلکه به افراد غیرپزشک نیز در اتخاذ تصمیمات بهداشتی و درمانی دقیق‌تر و پیش‌بینی تغییرات بیماری کمک می‌کند (۷، ۸).

این مطالعه یک مرور نظام‌مند است که تمامی پژوهش‌های مرتبط با کاربرد داده کاوی در تشخیص کووید-۱۹ را تحلیل می‌کند. هدف این پژوهش، ارزیابی میزان

کاربردپذیری این مطالعات در محیط‌های بالینی است. برای دستیابی به این هدف، ابتدا مطالعات از نظر نوع الگوریتم‌های به کار گرفته شده، شامل الگوریتم‌های «جعبه سیاه» و «جعبه سفید»، مورد بررسی قرار گرفت. سپس مشخص گردید که کدام یک از این الگوریتم‌ها در ارزیابی‌های بالینی نیز آزمون شده‌اند. در نهایت، پژوهش نشان می‌دهد که چه میزان از این مطالعات داده کاوی به طور عملی توسط تیم‌های درمانی به کار گرفته شده و چه تعداد تنها در سطح پژوهش باقی مانده‌اند. این تحلیل به شناسایی شکاف بین پژوهش‌های داده کاوی و کاربرد بالینی آن‌ها کمک کرده و راهکارهایی برای افزایش بهره‌گیری عملی از این تکنیک‌ها در تشخیص و مدیریت بیماری‌ها ارائه دهد.

روش بررسی

این پژوهش یک مطالعه مروری نظام‌مند است که به بررسی کاربرد هوش مصنوعی در پیش‌بینی و تشخیص کووید-۱۹ می‌پردازد. برای شناسایی مقالات مرتبط، جستجو در پایگاه‌های داده اسکوپوس، پایمد و وب او ساینس با استفاده از کلمات کلیدی «کرونا»، «هوش مصنوعی» و «تشخیص» انجام شده است. همچنین، کلمات مترادف این کلیدواژه‌ها با کمک منابع Mesh و دانش اعضای تیم پژوهش استخراج شده و در متن جستجو قرار گرفتند که در جدول ۱ این کلمات مترادف بیان شده است؛ به طوری که کلمات کلیدی اصلی با مترادف‌های خود به کمک عملگر OR و نتایج آن‌ها با عملگر AND ترکیب شدند.

پس از تجمیع نتایج جستجو از هر سه پایگاه داده، مقالات تکراری شناسایی و حذف شدند. تنها مقالاتی که معیارهای ورود را برآورده کرده و در زمینه پیش‌بینی و تشخیص کووید-۱۹ با استفاده از هوش مصنوعی بودند، برای تحلیل نهایی انتخاب شدند. متن جستجو در جدول ۱ مشاهده می‌شود.

جدول ۱: متن جستجو در سه پایگاه داده اسکوپوس، پایمد و WOS

ردیف	نام پایگاه داده	متن جستجو
1	Scopus	TITLE-ABS-KEY(("decision tree" OR "neural network" OR "Support vector machine" OR "Artificial intelligence" OR "Fuzzy" OR "machine learning" OR "Data mining" OR "genetic algorithm" OR metaheuristic OR "Meta heuristic" OR "meta-heuristic" OR "Cuckoo search" OR "Bees algorithm" OR "Artificial bee colony algorithm" OR "evolutionary algorithm" OR "Ant colony optimization" OR "Particle swarm optimization" OR "Firefly algorithm" OR "Memetic algorithm") AND ("COVID19" OR "Coronavirus Disease 2019" OR "COVID-19" OR "2019-nCoV") AND (diagnosis OR diagnostic OR diagnosing OR diagnosed OR diagnoses))
2	Pubmed	("decision tree"[Title/Abstract] OR "neural network"[Title/Abstract] OR "Support vector machine"[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 "Cuckoo search"[Title/Abstract] OR "Bees algorithm"[Title/Abstract] OR "Artificial bee colony algorithm"[Title/Abstract] OR "evolutionary algorithm"[Title/Abstract] OR "Ant colony optimization"[Title/Abstract] OR "Particle swarm optimization"[Title/Abstract] OR "Firefly algorithm"[Title/Abstract] OR "Memetic algorithm"[Title/Abstract]) AND ("COVID 19"[Title/Abstract] OR "Coronavirus Disease 2019"[Title/Abstract] OR "COVID-19"[Title/Abstract] OR "2019-nCoV"[Title/Abstract]) AND (diagnosis[Title/Abstract] OR diagnostic[Title/Abstract] OR diagnosing[Title/Abstract] OR diagnosed[Title/Abstract] OR diagnoses[Title/Abstract])
3	Web of Science (ISI)	TS=(("decision tree" OR "neural network" OR "Support vector machine" OR "Artificial intelligence" OR "Fuzzy" OR "machine learning" OR "Data mining" OR "genetic algorithm" OR metaheuristic OR "Meta heuristic" OR "meta-heuristic" OR "Cuckoo search" OR "Bees algorithm" OR "Artificial bee colony algorithm" OR "evolutionary algorithm" OR "Ant colony optimization" OR "Particle swarm optimization" OR "Firefly algorithm" OR "Memetic algorithm") AND ("COVID19" OR "Coronavirus Disease 2019" OR "COVID-19" OR "2019-nCoV") AND (diagnosis OR diagnostic OR diagnosing OR diagnosed OR diagnoses))

تشخیص کووید-۱۹ به کار گرفته شده و فراتر از آمار توصیفی، شامل تحلیل‌های آماری جامع و نتایج معتبر و قابل ارجاع باشند. همچنین در صورتی که الگوریتم مقاله در مقالات دیگر مرتبط با کووید-۱۹ تکرار شده باشند، قابل قبول نیستند. مقالات کنفرانسی یا

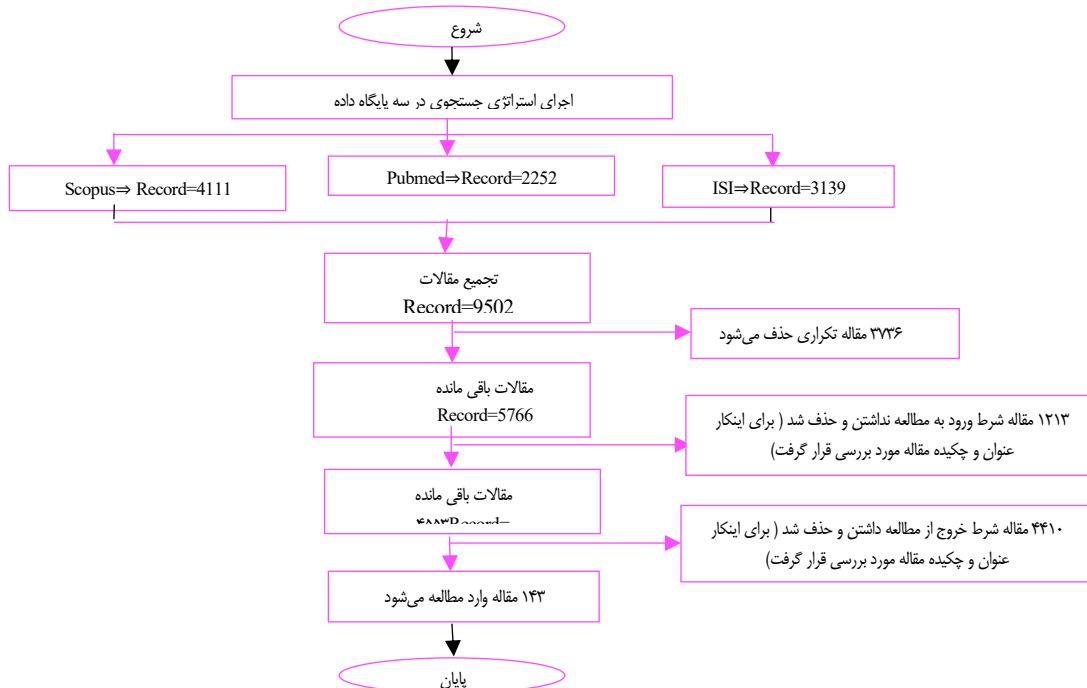
معیار ورود این مقالات عبارت بود از پژوهشی بوده، به زبان انگلیسی نوشته شده و به طور ویژه بر موضوع پیش‌بینی و تشخیص کووید-۱۹ با استفاده از الگوریتم‌های هوش مصنوعی تمرکز داشته و شامل مدلی پیش‌بینی‌گر باشند که به طور فعال در

استاندارد (مانند دقت، حساسیت و ویژگی) تضمین می‌کند که مدل‌ها قابلیت تشخیص دقیق و مؤثر را دارند که یکی از اصلی‌ترین معیارها برای تطابق با استانداردهای بالینی است (۱۰). همچنین تفسیرپذیری و شفافیت الگوریتم‌ها از آن جهت اهمیت دارد که می‌تواند نشان دهد الگوریتم‌ها قادر به توضیح و توجیه تصمیمات خود هستند، که این امر برای پذیرش در محیط‌های بالینی و همخوانی با فرآیندهای پزشکی ضروری است (۱۱) و در نهایت، تأیید یا ارزیابی قوانین استخراج شده توسط متخصصان حوزه پزشکی اطمینان می‌دهد که این قوانین با استانداردهای بالینی هماهنگ هستند و آماده به‌کارگیری در محیط‌های درمانی می‌باشند (۱۲).

در این مرحله، پارامترهای خروجی مقالات مرتبط با تشخیص کووید-۱۹ به‌طور توصیفی بیان شده و مشخص شد که کدام الگوریتم‌ها عملکرد بهتری دارند، چه درصدی از الگوریتم‌های مورد بررسی قادر به تولید قوانین قابل تفسیر بوده‌اند، و چند درصدی از این الگوریتم‌ها مورد ارزیابی بالینی قرار گرفته‌اند. برای ارزیابی عملیاتی شدن پژوهش‌ها در محیط‌های بالینی، لازم بود الگوریتم‌های مورد بررسی در گایدلاین‌های بالینی به‌عنوان یک روش شناخته‌شده گنجانده شده باشند. همه گایدلاین‌ها بر اساس مقالات مروری و متآنالیز طراحی می‌شوند (۱۳). بنابراین، در این مرحله، تمامی مقالات مروری و متا در زمینه کووید-۱۹ که به مقالات هوش مصنوعی در تشخیص این بیماری ارجاع داده‌اند، بررسی شدند تا مشخص شود آیا الگوریتم‌های به کار رفته در این مقالات مورد ارزیابی و تأیید بالینی قرار گرفته‌اند یا خیر.

یافته‌ها

جستجوی طراحی‌شده در بخش روش تحقیق، در تاریخ ۲۰۲۳/۰۵/۰۸ در سه پایگاه داده استنادی اسکوپوس، پابمد و وب او ساینس اجرا شد.



شکل ۱: فرآیند انتخاب مقالات مطالعه

آن‌هایی که صرفاً به توصیف آماری پرداخته‌اند و فاقد مدل‌های پیش‌بینی یا تحلیل‌های عملیاتی هستند، از فرآیند انتخاب حذف شده‌اند. هدف از این انتخاب دقیق، اطمینان از این است که نتایج به‌دست آمده از این مرور نظام‌مند مبتنی بر پژوهش‌های معتبر و با قابلیت به‌کارگیری در محیط‌های بالینی باشد و بتواند به ارزیابی عملیاتی بودن و میزان کاربرد هوش مصنوعی در تشخیص بیماری کمک کند.

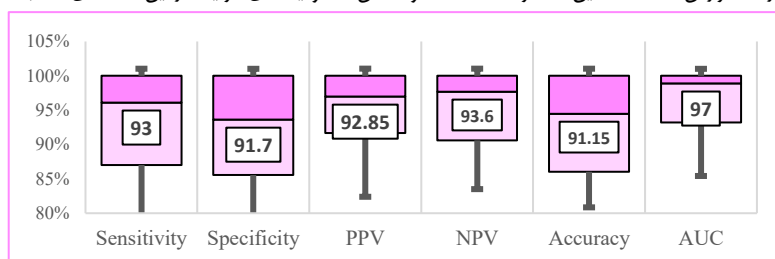
مقالات منتخب از شش جنبه کلیدی مورد بررسی قرار می‌گیرند این شش جنبه بر اساس بررسی‌های مقالات ورود به مطالعه و همچنین نظر تیم پژوهش انتخاب شده‌اند تا میزان عملیاتی شدن این پژوهش‌ها در محیط‌های بالینی به‌طور دقیق مشخص شود. این جنبه‌ها عبارتند از: نوع داده‌های به کار رفته در هر مطالعه، که اطلاعاتی درباره منابع و کیفیت داده‌ها ارائه می‌دهد؛ الگوریتم‌ها یا مجموعه الگوریتم‌های استفاده شده در هر مقاله، که به شناسایی روش‌های هوش مصنوعی به کار رفته کمک می‌کند؛ بررسی این که آیا الگوریتم‌های مورد استفاده به پردازش تصویر اختصاص دارند و نقش این الگوریتم‌ها در دقت تشخیصی؛ پارامترهای ارزیابی مدل پیشنهادی شامل معیارهای عملکردی مانند دقت، حساسیت، ویژگی و مقادیر مرتبط با هر پارامتر؛ قابلیت تفسیر الگوریتم در ایجاد قوانین تشخیصی، که نشان‌دهنده شفافیت و توضیح‌پذیری مدل است؛ و در نهایت، ارزیابی این که آیا قوانین تشخیصی استخراج شده (در صورت وجود) توسط متخصصان حوزه تأیید یا ارزیابی شده‌اند، تا اطمینان حاصل شود که این قوانین با استانداردهای بالینی همخوانی دارند و برای به‌کارگیری در محیط‌های درمانی مناسب هستند. زیرا نوع داده بررسی کیفیت و منابع داده‌ها این امکان را می‌دهد که اطمینان حاصل کنیم که داده‌ها از منابع معتبر و قابل اطمینانی استخراج شده‌اند، که در نتیجه می‌تواند به تطابق مدل‌ها با استانداردهای بالینی کمک کند (۹) معیارهای ارزیابی

پایین تر می‌تواند به محدودیت‌های ذاتی داده‌های صوتی و نیاز به توسعه بیشتر الگوریتم‌ها برگردد. با این حال، داده‌های صوتی به واسطه ماهیت غیرتجانسی و کم هزینه خود، پتانسیل بالایی برای بهبود و استفاده در تحقیقات آینده دارند. با توجه به خروجی شکل ۴ داده‌های تصویری، به‌ویژه تصاویر پزشکی، ابزار قدرتمندی برای تشخیص کووید-۱۹ هستند و می‌توانند به بهبود دقت تشخیص در محیط‌های بالینی کمک کنند.

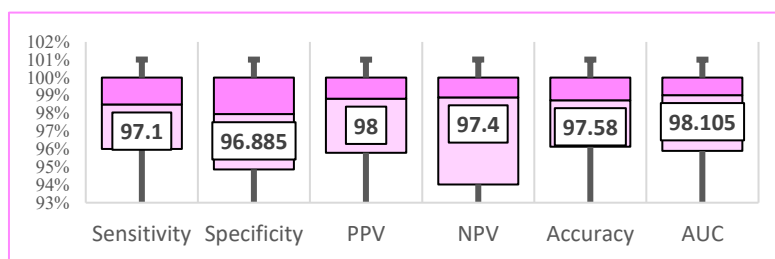
داده‌های صوتی و رکوردهای اطلاعاتی نیز می‌توانند به‌عنوان روش‌های مکمل یا جایگزین در محیط‌هایی که تجهیزات پیشرفته در دسترس نیست، مورد استفاده قرار گیرند. به‌کارگیری داده‌های ترکیبی می‌تواند رویکرد مناسبی برای افزایش دقت و کاهش محدودیت‌های هر یک از این دسته‌های داده باشد.

رکوردهای اطلاعاتی، به‌عنوان دومین گروه از داده‌های ورودی، عملکرد قابل‌قبولی داشته‌اند اما در مقایسه با تصاویر پزشکی در برخی پارامترها از دقت کمتری برخوردار بوده‌اند. حساسیت این الگوریتم‌ها در حدود ۹۳ درصد و ویژگی آن‌ها نزدیک به ۸۸ درصد است. این نتایج نشان می‌دهند که استفاده از رکوردهای اطلاعاتی به‌تنهایی نمی‌تواند به دقت الگوریتم‌های مبتنی بر تصاویر پزشکی برسد، اما همچنان به‌عنوان یک گزینه مقرون به صرفه و در دسترس در محیط‌های درمانی باقی می‌ماند.

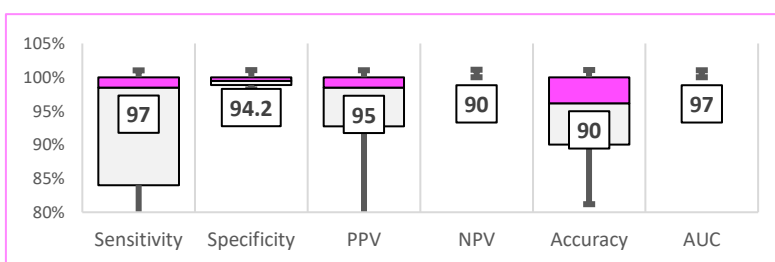
در مقابل، الگوریتم‌های مبتنی بر داده‌های صوتی کمترین عملکرد را در میان این سه دسته داده‌های ورودی داشته‌اند. میانگین AUC برای این الگوریتم‌ها حدود ۸۵ درصد و دقت کلی آن‌ها نزدیک به ۸۰ درصد گزارش شده است. این عملکرد



الف-خروجی الگوریتم‌های هوش مصنوعی تشخیص کرونا با استفاده از رکوردهای اطلاعاتی



ب-خروجی الگوریتم‌های هوش مصنوعی تشخیص کرونا با استفاده از تصاویر پزشکی

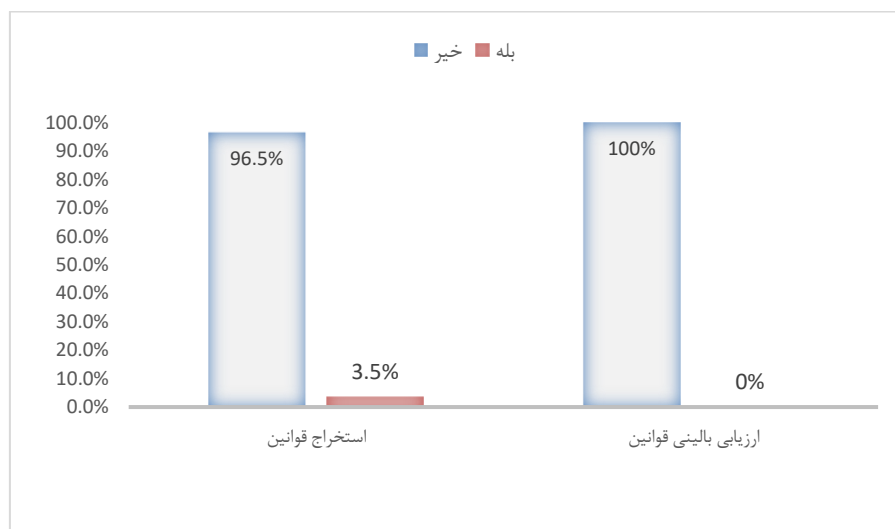


ج-خروجی الگوریتم‌های هوش مصنوعی تشخیص کرونا با استفاده از اطلاعات صوتی

شکل ۴: boxplot خروجی الگوریتم هوش مصنوعی در تشخیص کرونا

تنها ۳.۵ درصد مقالات موفق به استخراج قوانین بالینی شده‌اند و هیچ‌یک از این قوانین توسط متخصصان ارزیابی نشده است. این موضوع شکاف بزرگ بین تحقیقات هوش مصنوعی و کاربرد عملی آن‌ها در محیط‌های درمانی را نمایان می‌کند.

پارامتر دیگر مورد بررسی، تعداد مقالات دارای قوانین بالینی و تعداد الگوریتم‌هایی است که توسط متخصصان ارزیابی بالینی شده‌اند. نتایج این بررسی در شکل ۵ نمایش داده شده است. این شکل تحلیل میزان استخراج و ارزیابی قوانین بالینی در الگوریتم‌های هوش مصنوعی تشخیص کووید-۱۹ را نشان می‌دهند.



شکل ۵: بررسی عملکرد الگوریتم از نظر استخراج قوانین بالینی و ارزیابی آن قوانین توسط الگوریتم‌های هوش مصنوعی

بحث

در این پژوهش، کاربرد الگوریتم‌های هوش مصنوعی در تشخیص کووید ۱۹ از منظر عملیاتی شدن در محیط‌های بالینی بررسی شد. نتایج نشان داد که با وجود دقت بالای الگوریتم‌ها (میانگین بیش از ۹۰ درصد)، محدودیت‌هایی نظیر عدم شفافیت تصمیم‌گیری (جعبه سیاه)، فقدان قوانین بالینی استخراج شده و عدم ارزیابی این قوانین توسط متخصصان، مانع پذیرش و استفاده عملی آن‌ها در گایدلاین‌های بالینی شده است.

در این تحقیق، میزان کاربردپذیری هوش مصنوعی در تشخیص کووید-۱۹ از منظر استفاده بالینی بررسی شد. برای پاسخ به این پرسش، ابتدا معیارهای ورود یک روش به محیط بالینی تعریف شد؛ بدین معنا که روش تنها زمانی می‌تواند در بالین استفاده شود که در گایدلاین‌های پزشکی ثبت شده باشد. این ثبت نیز مستلزم تأیید عملکرد روش در مقالات مروری و متآنالیز است که خود از مقالات اصلی استخراج می‌شوند. بر اساس استراتژی جستجو، مقالات اصلی مرتبط با هوش مصنوعی در تشخیص کووید-۱۹ شناسایی و عملکرد آن‌ها ارزیابی شد. نتایج این بررسی نشان داد که هیچ‌کدام از این مقالات از نظر بالینی توسط متخصصان ارزیابی نشده‌اند و بیش از ۹۶ درصد از مطالعات از الگوریتم‌های جعبه سیاه استفاده کرده‌اند. همچنین، تمامی مقالات مروری و متآنالیز مرتبط تحلیل شدند و مشخص شد که هیچ‌کدام به‌طور رسمی تحت ارزیابی بالینی قرار نگرفته‌اند.

در این مطالعه، سه نوع داده برای تشخیص کووید-۱۹ شناسایی شد: رکوردهای اطلاعاتی، تصاویر پزشکی و داده‌های صوتی. در برخی مقالات، ترکیبی از دو نوع داده به کار رفته است، مانند مقاله (۱۴). نتایج تحلیل‌ها نشان داد که تصاویر پزشکی بهترین عملکرد را در میان این سه دسته دارند زیرا اطلاعات پیچیده‌تر و دقیق‌تری را در اختیار قرار می‌دهند. با این حال، الگوریتم‌های پردازش تصویر اغلب به صورت جعبه سیاه عمل می‌کنند و قادر به ارائه دلایل مشخص برای تصمیم‌گیری خود نیستند. در مقابل، پزشکان می‌توانند با تفسیر مستقیم تصاویر سی‌تی اسکن (CT) استدلال‌های بالینی خود را ارائه دهند که این موضوع یک تفاوت اساسی بین تصمیمات الگوریتم‌های پردازش تصویر و تشخیص‌های پزشکی ایجاد می‌کند.

به دلیل محدودیت‌های موجود در بررسی تمام مقالات مروری و متآنالیزهای مرتبط با کووید-۱۹، تصمیم گرفته شد که روند کار به صورت بالا به پایین انجام شود. بدین منظور، ابتدا در هر یک از ۱۴۳ مقاله بررسی‌شده، مقالات مروری و متآنالیز که به این مقالات ارجاع داده بودند، شناسایی شدند. سپس، این مقالات مروری و متآنالیز به طور دقیق‌تری مورد بررسی قرار گرفتند تا مشخص شود آیا در آن‌ها قوانین بالینی ارائه شده است یا خیر.

در این فرآیند، ۴۵۱ مقاله مروری و متآنالیز مرتبط با استفاده از هوش مصنوعی در تشخیص کووید-۱۹ استخراج و تحلیل شدند. این مقالات شامل مطالعات مختلفی از جمله مقالات تحلیلی، مرور سیستماتیک و متآنالیزهایی بودند که به بررسی روش‌ها و الگوریتم‌های مختلف هوش مصنوعی در زمینه تشخیص و پیش‌بینی این بیماری پرداخته بودند. پس از تحلیل دقیق این مقالات، نتیجه‌گیری‌های قابل‌توجهی حاصل شد که یکی از مهم‌ترین آن‌ها این بود که هیچ‌یک از این مقالات قوانین بالینی مشخصی را برای استفاده عملیاتی در بالین ارائه نکرده بودند. این موضوع به وضوح نشان می‌دهد که، علیرغم حجم بالای مقالات مروری و متآنالیز در این حوزه، هنوز هیچ تلاش جدی، منظم و سیستماتیک برای تدوین و ارائه قوانین بالینی بر اساس یافته‌های تحقیقاتی و تجربی موجود انجام نشده است.

علاوه بر این، بررسی‌ها نشان داد که اغلب مقالات به تحلیل‌های تئوریک و گزارش‌های آماری محدود شده‌اند و کمتر به ایجاد چارچوب‌های عملی و ارزیابی‌های دقیق بالینی پرداخته‌اند. این کمبود قوانین بالینی می‌تواند به‌طور جدی بر پذیرش و به‌کارگیری الگوریتم‌های هوش مصنوعی را در فرآیندهای تشخیصی و درمانی محدود کند، بنابراین ضروری است که پژوهش‌های آینده علاوه بر ارائه نتایج آماری و الگوریتمی، تمرکز بیشتری به طراحی و تدوین قوانین بالینی مبتنی بر این یافته‌ها داشته باشند تا این الگوریتم‌ها بتوانند به‌طور موثر و معتبر در محیط‌های درمانی مورد استفاده قرار گیرند.

گزارش این مقالات در پیوست ۲ ارائه شده است تا محققان و علاقه‌مندان بتوانند به تمامی جزئیات دسترسی داشته و از نتایج این بررسی‌ها بهره‌مند شوند. این گزارش می‌تواند به عنوان یک منبع ارزشمند برای تحقیقات آینده و تدوین استراتژی‌های بالینی مورد استفاده قرار گیرد.

برای جلب اعتماد پزشکان و پذیرش این فناوری‌ها در محیط‌های درمانی ضروری است.

اگرچه الگوریتم‌های هوش مصنوعی در تشخیص کووید-۱۹ دقت بالایی دارند، اما بدون شفاف‌سازی فرآیند تصمیم‌گیری و ارزیابی قوانین بالینی، کاربرد عملی آن‌ها همچنان محدود خواهد بود. تحقیقات آینده باید بر روی توسعه الگوریتم‌های شفاف (جعبه سفید) و ارزیابی گسترده‌تر قوانین استخراج‌شده توسط تیم‌های متخصص بالینی متمرکز شوند تا امکان ادغام مؤثر این الگوریتم‌ها در گایدلاین‌های بالینی فراهم گردد.

نتیجه‌گیری

این پژوهش بر ارزیابی قابلیت عملیاتی‌شدن این الگوریتم‌ها در محیط‌های بالینی تأکید دارد. یافته‌ها نشان می‌دهد که علاوه بر بهبود دقت، توجه به شفافیت، تفسیرپذیری و ارزیابی قوانین استخراج‌شده برای پذیرش این الگوریتم‌ها در محیط‌های درمانی ضروری است. همچنین پیشنهاد می‌شود که تحقیقات آینده به توسعه روش‌های ترکیبی بپردازند که داده‌های تصویری و رکوردهای اطلاعاتی را ادغام کنند، چرا که این رویکرد می‌تواند هم دقت و هم شفافیت الگوریتم‌ها را بهبود بخشد. در نهایت، ایجاد همکاری نزدیک‌تر بین محققان هوش مصنوعی و تیم‌های پزشکی برای توسعه و ارزیابی قوانین بالینی امری حیاتی است. چنین همکاری‌هایی می‌توانند به کاهش شکاف میان تحقیقات نظری و کاربردهای عملی در محیط‌های درمانی کمک کرده و پذیرش گسترده‌تر الگوریتم‌های هوش مصنوعی را تسهیل نمایند.

پیشنهادهای

۱. توسعه الگوریتم‌های تفسیرپذیر: طراحی مدل‌هایی که بتوانند فرآیند تصمیم‌گیری خود را توضیح دهند تا اعتماد پزشکان و بیماران جلب شود.
۲. ترکیب داده‌های چندمنظوره: استفاده همزمان از تصاویر پزشکی (مانند CT اسکن) و داده‌های بالینی برای بهبود دقت تشخیص.
۳. ارزیابی بالینی گسترده: انجام مطالعات بالینی با همکاری متخصصان پزشکی برای سنجش عملکرد الگوریتم‌ها در شرایط واقعی.
۴. تدوین دستورالعمل‌های بالینی: ایجاد راهنمای استاندارد برای استفاده از هوش مصنوعی در تشخیص کووید-۱۹ بر اساس شواهد علمی.
۵. تقویت همکاری بین‌رشته‌ای: افزایش تعامل بین پژوهشگران هوش مصنوعی و پزشکان برای اطمینان از انطباق فناوری با نیازهای بالینی.

تشکر و قدردانی

این پژوهش بخشی از رساله دکتری نویسنده اول مقاله با کد ۱۶۳۹۴۸۳۲۲۹۵۳۴۵۰۷۳۰۰۲۵۱۶۲۷۸۸۶۷۶ است. نویسندگان از معاونت پژوهشی دانشگاه آزاد شیراز به دلیل حمایت‌های ارزشمندشان قدردانی می‌کنند.

تضاد منافع

در انجام پژوهش حاضر، نویسندگان هیچ‌گونه تضاد منافی نداشتند.

References

1. Stokes K, Castaldo R, Federici C, Pagliara S, Maccaro A, Cappuccio F, et al. The use of artificial intelligence systems in diagnosis of pneumonia via signs and symptoms: A systematic review. *Biomedical Signal Processing and Control*. 2022;72:103325.
2. Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives. *Int J Biol Sci*. 2021;17(6):1581-7.

مطالعات مشابه در این حوزه نیز تأیید کرده‌اند که الگوریتم‌های هوش مصنوعی، به‌ویژه روش‌های یادگیری عمیق، در تشخیص بیماری‌هایی مانند کووید-۱۹ عملکرد بالایی داشته‌اند. به‌عنوان مثال، پژوهش (۱۵) نشان می‌دهد که مدل‌های یادگیری عمیق، به‌ویژه شبکه‌های عصبی کانولوشنی (CNN)، در تحلیل تصاویر CT دقتی بالاتر از ۹۵ درصد ارائه داده‌اند. این یافته‌ها با نتایج پژوهش حاضر همخوانی دارند و تأکید می‌کنند که تصاویر پزشکی منبع داده‌ای بسیار قدرتمند برای الگوریتم‌های تشخیصی محسوب می‌شوند. با این حال، مشابه این پژوهش حاضر، در پژوهش (۱۵) نیز گزارش شده است که بیشتر الگوریتم‌ها فاقد قابلیت تفسیر هستند و نتایج به‌دست‌آمده تاکنون توسط تیم‌های بالینی ارزیابی نشده‌اند.

در پژوهش‌های دیگری مانند (۱۶، ۱۷) اهمیت توسعه الگوریتم‌های جعبه سفید (توضیح‌پذیر) تأکید شده است. این مطالعات نشان داده‌اند که عدم شفافیت الگوریتم‌های جعبه سیاه باعث کاهش اعتماد پزشکان به این سیستم‌ها شده و مانعی جدی برای استفاده گسترده آن‌ها در محیط‌های بالینی ایجاد می‌کند. پژوهش حاضر نیز این مسئله را برجسته کرده و بر ضرورت توسعه الگوریتم‌هایی تأکید دارد که بتوانند قوانین بالینی شفاف و قابل تفسیر ارائه دهند. چنین الگوریتم‌ها می‌توانند نقش مهمی در افزایش پذیرش هوش مصنوعی در پزشکی ایفا کنند.

از سوی دیگر، برخی پژوهش‌ها مانند تحقیقات (۱۸، ۱۹) به بررسی مزایای الگوریتم‌های تفسیرپذیر پرداخته‌اند و نشان داده‌اند که این الگوریتم‌ها با ارائه قوانین واضح و ارزیابی‌های بالینی، می‌توانند به‌طور مؤثر در تصمیم‌گیری‌های درمانی مورد استفاده قرار گیرند. با این حال، نتایج پژوهش حاضر نشان می‌دهد که تنها ۳.۵ درصد از مقالات بررسی‌شده موفق به استخراج قوانین بالینی شده‌اند و حتی این قوانین نیز مورد ارزیابی بالینی قرار نگرفته‌اند. این یافته‌ها نشان‌دهنده یک شکاف قابل توجه میان پژوهش‌های انجام‌شده و کاربرد علمی الگوریتم‌های هوش مصنوعی در محیط درمانی است.

کمترین میزان استفاده به داده‌های صوتی تعلق دارد که تنها در ۵ درصد از مقالات بررسی‌شده‌اند. داده‌های صوتی، مانند صدای سرفه یا تنفس، به‌عنوان روش‌های غیرتهاجمی و کم‌هزینه پتانسیل بالایی دارند، اما به‌دلیل چالش‌هایی مانند نیاز به ابزارهای خاص برای جمع‌آوری و تحلیل صدا و عدم اطمینان در نتایج، کمتر مورد توجه قرار گرفته‌اند.

این تحلیل نشان می‌دهد که انتخاب نوع داده ورودی تأثیر زیادی بر عملکرد الگوریتم‌های هوش مصنوعی دارد. داده‌های تصویری به‌دلیل دقت بالا بیشترین کاربرد را داشته‌اند، در حالی که داده‌های صوتی و رکوردهای اطلاعاتی نیازمند توسعه بیشتر هستند. استفاده از داده‌های ترکیبی، مانند ترکیب تصاویر پزشکی و رکوردهای اطلاعاتی، می‌تواند به بهبود عملکرد مدل‌ها و افزایش کاربرد آن‌ها در محیط‌های بالینی کمک کند.

عدم وجود ارزیابی بالینی برای این قوانین نشان می‌دهد که تعامل بین محققان هوش مصنوعی و متخصصان بالینی در بسیاری از مطالعات کافی نبوده است. این موضوع همچنین نمایانگر نیاز به انجام تحقیقات بیشتری است که علاوه بر ارائه الگوریتم‌های پیشرفته، به ارزیابی بالینی آن‌ها نیز پرداخته شود. چنین ارزیابی‌هایی

3. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr*. 2020;14(4):337-9.
4. Ahouz F, Golabpour A. Predicting the incidence of COVID-19 using data mining. *BMC Public Health*. 2021;21(1):1087.
5. Azodi CB, Tang J, Shiu SH. Opening the Black Box: Interpretable Machine Learning for Geneticists. *Trends Genet*. 2020;36(6):442-55.
6. Loyola-Gonzalez O. Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view. *IEEE access*. 2019;7:154096-113.
7. Khan PF, Meehan K, editors. Diabetes prognosis using white-box machine learning framework for interpretability of results. 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC); 2021: IEEE.
8. Wang S, Wang Y, Wang D, Yin Y, Wang Y, Jin Y. An improved random forest-based rule extraction method for breast cancer diagnosis. *Applied Soft Computing*. 2020;86.
9. Pipino LL, Lee YW, Wang RY. Data quality assessment. *Communications of the ACM*. 2002;45(4):211-8.
10. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. *American Journal of Roentgenology*. 2005;184(1):14-9.
11. Shakir MA, Abass HK, Jelwy OF, Al-Bayati HNA, Salman SM, Mikhav V, et al., editors. Developing Interpretable Models for Complex Decision-Making. 2024 36th Conference of Open Innovations Association (FRUCT); 2024 30 Oct.-1 Nov. 2024.
12. Qiao H, Chen Y, Qian C, Guo Y. Clinical data mining: challenges, opportunities, and recommendations for translational applications. *Journal of Translational Medicine*. 2024;22(1):185.
13. Connor Gorber S, Singh H, Pottie K, Jaramillo A, Tonelli M. Process for guideline development by the reconstituted Canadian Task Force on Preventive Health Care. *CMAJ*. 2012;184(14):1575-81.
14. Yu L, Shi X, Liu X, Jin W, Jia X, Xi S, et al. Artificial Intelligence Systems for Diagnosis and Clinical Classification of COVID-19. *Front Microbiol*. 2021;12:729455.
15. Sekeroglu B, Ozsahin I. Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks. *SLAS Technol*. 2020;25(6):553-65.
16. Giannotti F, Naretto F, Bodria F. Explainable for Trustworthy AI. In: Chetouani M, Dignum V, Lukowicz P, Sierra C, editors. *Human-Centered Artificial Intelligence: Advanced Lectures*. Cham: Springer International Publishing; 2023. p. 175-95.
17. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information Fusion*. 2023;99:101805.
18. Vernon EM, Masuyama N, Nojima Y, editors. Integrating White and Black Box Techniques for Interpretable Machine Learning. *Proceedings of Ninth International Congress on Information and Communication Technology*; 2024 2024//; Singapore: Springer Nature Singapore.
19. Caniço A, Santos A. Witter: A Library for White-Box Testing of Introductory Programming Algorithms 2023. 69-74 p.

Evaluating AI Performance in COVID-19 Diagnosis: A Systematic Review

Azadeh Abkar¹, Mahdi Mehrabi², Amin Golabpour³, Mohammad Amin Shayegan⁴

Review Article

Abstract

Introduction: Early detection of COVID-19 plays a crucial role in reducing mortality. With its ability to analyze large volumes of medical data, artificial intelligence (AI) can aid in the rapid and accurate diagnosis of this disease. This review study aims to assess the performance of AI models in detecting COVID-19.

Methods: A systematic search was conducted in the PubMed, Scopus, and WOS databases, yielding 143 relevant articles. The studies were analyzed based on the type of input data, algorithms used, performance evaluation metrics, the presentation of clinical rules, and the clinical validation of the models. Additionally, 451 review and meta-analysis articles were examined to determine the extent to which these algorithms have been integrated into clinical guidelines.

Results: The input data in the articles consisted of clinical records, medical images, and audio processing. More than 96% of the algorithms were found to be black-box models, lacking clinical validation by specialists. However, the average performance of these models was reported to be above 90%. The review of meta-analysis articles revealed that none of the algorithms had undergone formal clinical evaluation; only their performance on available data was assessed.

Conclusion: The review of studies shows that AI performs well in COVID-19 detection but faces limitations, including lack of explainability, reliance on training data, and the absence of extensive clinical evaluations. The development of white-box models and conducting broad clinical studies are crucial for ensuring the widespread acceptance of these algorithms in clinical settings.

Keywords: COVID-19, diagnosis, data mining, artificial intelligence

Received: 1 May; 2024

Accepted: 18 Jun; 2024

Published: 21 Jun; 2024

Citation: Abkar A, Mehrabi M, Golabpour A, Shayegan MA. **Evaluating AI Performance in COVID-19 Diagnosis: A Systematic Review.** Health Inf Manage 2024; 21(2):103-111.

Article resulted from PhD thesis with No 163948322953450730025162788676

1. PhD student, Department of Computer Science, Shiraz Branch, Islamic Azad University, Shiraz, Iran

2. Assistant Professor, Department of Computer Science, Shiraz Branch, Islamic Azad University, Shiraz, Iran

3. Assistant Professor, Department of Health Information Technology, Shahrood University of Medical Sciences, Shahrood, Iran

4. Associate Professor, Department of Computer Science, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Corresponding author: Mehdi Mehrabi, Assistant Professor, Department of Computer Science, Shiraz Branch, Islamic Azad University, Shiraz, Iran
Email: mehdimehrabi1975@gmail.com