مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره پنجم شماره چهارم. ص ٤٦٨-٤٥٧، ١٣٩٧

مقایسه سه مدل تصمیم‌گیری در افتراق پنج نوع بیماری قلبی (مطالعه موردی: بیمارستان فوق تخصصی قائم کرج)

سیده راحیل موسوی[1]*، محمدمهدی سپهری[2]

• دریافت مقاله: ۹۷/۳/۲ • پذیرش مقاله: ۹۷/۶/۱۷

مقدمه: بیماری‌های قلبی- عروقی در حال تبدیل‌شدن به اصلی‌ترین عامل مرگ‌ومیر و ناتوانی بشر در اغلب کشورهای دنیا هستند. هدف از انجام این پژوهش، پیش‌بینی انواع بیماری‌های قلبی جهت تشخیص دقیق‌تر به وسیله تکنیک‌های داده‌کاوی و شبکه عصبی می‌باشد.

روش: این پژوهش به صورت کاربردی-پیمایشی انجام و پس از پیش‌پردازش داده‌ها از سه رویکرد شبکه عصبی، درخت تصمیم‌گیری و الگوریتم ساده بیزی برای پیش‌بینی و تشخیص در نرم افزار Rapid miner و از مدل شبکه عصبی مصنوعی برای پیش‌بینی در نرم‌افزار Matlab استفاده شد.

نتایج: از الگوریتم ژنتیک برای انتخاب متغیرهای مؤثر و برای پیش‌بینی انواع بیماری قلبی، در داده‌کاوی از مدل‌های شبکه عصبی مصنوعی، درخت تصمیم‌گیری و الگوریتم ساده بیزی استفاده شد. همچنین از مدل AHP برای تعیین مدل با بهترین عملکرد پیش‌بینی انواع بیماری‌های قلبی استفاده شد.

نتیجه‌گیری: شبکه عصبی عملکرد بسیار بهتری نسبت به مدل‌های داده‌کاوی دیگر ارائه شده در تشخیص انواع بیماری‌های قلبی در این پژوهش دارد. همچنین در تشخیص بیماری به وسیله شبکه عصبی مصنوعی، مدل با دقت بالای ۸۰ درصد، خوب و مورد قبول واقع شد.

کلید واژه‌ها: داده‌کاوی، شبکه عصبی، درخت تصمیم، بیماری قلبی، تشخیص

•ارجاع: موسوی سیده راحیل، سپهری محمدمهدی. مقایسه سه مدل تصمیم‌گیری در افتراق پنج نوع بیماری قلبی (مطالعه موردی: بیمارستان فوق تخصصی قائم کرج). مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۷؛ ۵(۴): ۴۵۷-۴۶۸.

۱. دانشجوی دکتری مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه علوم و تحقیقات، تهران، ایران
۲. دکترای تحقیق در عملیات، استاد گروه مهندسی صنایع، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران
*** نویسنده مسئول:** تهران، دانشگاه علوم و تحقیقات، دانشکده فنی و مهندسی، ایران، گروه مهندسی صنایع
• شماره تماس: ۰۹۳۷۱۰۳۶۹۹۷
• **Email:** Raahil.mousavi@yahoo.com

[Downloaded from jhbmi.com on 2026-11-15]

مقدمه

بیماری‌های قلبی –عروقی در حال تبدیل‌شدن بـه اصـلی‌تـرین عامل مرگ‌ومیر و ناتوانی بشر در اغلب کشورهای دنیا هسـتند. این بیماری‌ها در حال حاضر جـزء سـه علـت اول مرگ‌ومیـر و ناتوانی انسان در سراسر دنیـا هسـتند کـه سـالانه هزینـه‌هـای زیادی صرف این بیماری می‌شود. هر روز بخشی از افراد جامعه به عنوان بیماران قلبی به مراکز درمانی کشور مراجعه می‌نمایند که برخی سرپایی درمـان مـی‌شـوند و عـده‌ای دیگـر نیـاز بـه بستری‌شـدن در بیمارسـتان و انجـام اقـدامات درمـانی دارنـد. عـده‌ای از ایـن بیمـاران بـه انجـام درمـان‌هـای تهـاجمی و غیرتهاجمی متعدد از جملـه تسـت ورزش، آنژیـوگرافی، اعمـال جراحی باز و غیرباز قلبی و ... احتیاج دارند که علاوه بر ایـن‌هـا باید به مخارج آزمایش‌ها، هزینه بسـتری و سـایر مـوارد اشـاره نمود. [۱].

در کتاب دست‌نامه طب داخلی Harrison's [۲] بیمـاری‌هـای قلبی به ۴۱ نوع تقسیم‌بندی شده است؛ که در این پـژوهش بـا نظر چند متخصص قلب این تقسـیم‌بنـدی بـه ۵ نـوع کـاهش یافت. این پنج گروه عبارت‌اند از: بیماری‌های قلبـی مـادرزادی، دریچه‌ای قلب، نارسایی قلبی، آنژین و آریتمی‌ها.

یکی از مشکلات این حوزه، عدم تشخیص به موقـع بیمـاری می‌باشد. یک‌سری الگوریتم‌هـا بـرای پـیش‌بینـی و کمـک بـه تشخیص، در ریاضی وجود دارند که می‌توانند خطاهای پزشکی را به شدت کـاهش دهنـد. الگـوریتم‌هـای داده‌کـاوی و شـبکه عصبی، دو ابزار کاربردی در این امـر مـی‌باشـند کـه پزشـکان می‌توانند از آن‌ها استفاده کنند. استخراج قواعد طبقه‌بندی نوعی داده‌کاوی است که در آن دانشی به شکل چندین قانون ساده و فهم‌پذیر از داده کشف‌شـده و در آینـده بـرای تصـمیم‌گیـری و پیشـگویی بـه‌کار بـرده مـی‌شـود. از مهم‌تـرین زمینـه‌هـای کـاربـردی استخـراج قـواعـد طبقـه‌بنـدی در حیطـه علـم پـزشکی است. در سیستم‌های پزشکی حجم زیادی داده شامل اعداد، متون و عکس‌ها ایجاد و ذخیره می‌شوند. متأسفانه ایـن داده‌ها به ندرت در تصمیم‌سازی‌های پزشکی مورد استفاده قرار می‌گیرند. روش‌های داده‌کاوی می‌تواننـد پزشـکان و مـدیران بهداشت و درمان را در تصـمیم‌گیـری‌هـا از جملـه تشـخیص بیماری‌ها، با استفاده از این داده‌ها یاری کنند [۳].

فرآینـد داده‌کـاوی شـامل ۳ مرحلـه اسـت: آمـاده‌سـازی داده، یادگیری مدل، ارزیابی و تفسیر مدل. اولین و مهم‌ترین مرحلـه در فرآیند داده‌کاوی آماده‌سازی داده مـی‌باشـد. هـدف در ایـن مرحله تأمین ورودی مناسب برای مرحله حیاتی یادگیری مـدل است. در این مرحله داده پـردازش نشـده از کـل منـابع داده‌ای موجود (که ممکن است توزیع شده نیز باشـد) اسـتخراج شـده، سپس در مرحله‌ای مستقل مورد پردازش اولیه قـرار مـی‌گیـرد. خروجـی در مرحلـه آمـاده‌سـازی داده عبـارت اسـت از داده پیش‌پردازش‌شده که امکان یـادگیری مـدل از روش آن وجـود دارد. در این مرحله با استفاده از الگوریتم‌های متنوع و با توجـه به ماهیت داده، سعی بر این است که نظم‌های مختلف موجـود در داده را شناسایی و در شکلی مشخص به‌عنوان دانش نهفتـه در داده ارائه شود. برای یادگیری مدل می‌بایست روش‌هـای آن را به درستی شناخت تا بتوان در جای مناسـب، روش درسـت را انتخاب نمود و به کار بست. در این مرحله دانـش تولیدشـده در مرحله قبل ارزیابی‌شده و مورد تفسیر قرار مـی‌گیـرد. منظـور از ارزیابی دانش آن است کـه مـی‌بایسـت میـزان صـحت دانـش تولیدشده مشخص شود تا بتوان به آن اعتماد نمود و به صورت عملی از آن استفاده کرد [۴].

مشهورترین روش‌های یادگیری مدل در داده‌کاوی؛ روش‌های پـیش‌بینـی و روش‌هـای توصـیفی مـی‌باشـد. در روش‌هـای پیش‌بینی مقادیر بعضی از ویژگی‌ها برای پیش‌بینی کردن مقدار یک ویژگی مشخص استفاده می‌کنند. روش‌هـای دسـته‌بنـدی، رگرسـیون و تشـخیص انحـراف سـه روش یـادگیری مـدل در داده‌کاوی با ماهیـت پـیش‌بینـی هسـتند. از کاربردهـای روش دسته‌بندی می‌تـوان اسـتفاده از آن در پزشـکی را نـام بـرد. در الگوریتم‌های دسته‌بندی مجموعه داده اولیه به دو مجموعه داده با عنوان مجموعه داده‌های آموزشی و مجموعـه داده آزمایشـی تقسیم می‌شود. با استفاده از مجموعه داده‌های آموزشـی مـدل ساخته می‌شود و از مجموعه داده آزمایشی برای اعتبارسنجی و محاسبه دقت مدل ساخته‌شده استفاده می‌شود [۳،۴].

مدل‌های پیش‌بینی زیادی در داده‌کاوی وجود دارد، سه مـورد از مهم‌ترین آن‌ها که در این پژوهش مـورد اسـتفاده قـرار گرفتـه عبارت‌اند از ۱– شبکه عصبی، ۲– درخت تصمیم و ۳– الگوریتم ساده بیزی.

شبکه عصبی: شبکه‌های عصبی مصنوعی در برخی از عملیات مانند پیش‌بینی و دسته‌بندی در مقایسه با سـایر روش‌هـا دارای مزایای نسبی بوده و معمـولاً در کارهـای اجرایـی تـرجیح داده می‌شوند. هر شبکه عصبی مصنوعی، مجموعـه‌ای از ورودی‌هـا است و بر اساس یک تابع فعال‌سازی مقدار خروجی آن محاسبه می‌شود [۴].

درخت تصمیم: یکی از مشهورترین و قدیمی‌ترین روش‌هـای ساخت مدل دسته‌بنـدی اسـت. در الگـوریتم‌هـای دسـته‌بنـدی

[Downloaded from jhbmi.com on 2026-11-15]

مبتنی بر درخت تصمیم دانش خروجی به صورت یک درخت از حالات مختلف مقادیر ویژگی‌ها ارائه می‌شود. نمایش دانش به شکل درخت سبب شده است که دسته‌بندی‌های مبتنی بر درخت تصمیم کاملاً قابل تفسیر باشند.

الگوریتم ساده بیزی: دسته‌بندی مبتنی بر رابطه نظریه بیز از یک چهارچوب احتمالی برای حل مسائل دسته‌بندی استفاده می‌کند. در نظریه بیز از احتمالات شرطی استفاده و در نهایت به قاعده بیز می‌رسد [۳].

در پژوهش Tsipouras و همکاران یک سیستم پشتیبانی مبتنی بر قاعده فازی (Design Support System) برای تشخیص بیماری عروق کرونر (Coronary CAD(Artery Disease ارائه شده است. مجموعه داده‌ای که برای تولید و ارزیابی DSS مورد استفاده قرار می‌گیرد شامل ۱۹۹ نفر است که هر کدام ۱۹ ویژگی را شامل داده‌های جمعیت‌شناختی و تاریخی و آزمایش‌های آزمایشگاهی دارند. سیستم از یک روش چهار مرحله‌ای ایجاد می‌شود: ۱) القاء یک درخت تصمیم‌گیری از داده‌ها ۲) استخراج تعدادی از قواعد از درخت تصمیم‌گیری، در قالب عادی دیوانه و فرموله کردن یک مدل واضح ۳) تبدیل مجموعه‌ای از قوانین به یک مدل فازی و ۴) بهینه‌سازی پارامترهای مدل فازی. اعتبارسنجی متقاطع ده مرتبه مورد استفاده قرار می‌گیرد و حساسیت و خصوصیات متوسط به دست آمده، به ترتیب ۶۲٪ و ۵۴٪، با استفاده از مجموعه قوانینی که از درخت تصمیم‌گیری (مراحل اول و دوم) استخراج می‌شوند، درحالی‌که حساسیت و خصوصیات متوسط به ۸۰٪ افزایش می‌یابد و به ترتیب ۶۵٪، زمانی که مراحل فازسازی و بهینه‌سازی استفاده می‌شود. این سیستم مزایای متعددی را ارائه می‌دهد. از آنجایی که به طور خودکار تولید می‌شود، تشخیص CAD را بر اساس ویژگی‌هایی به راحتی فراهم می‌کند و می‌تواند تفسیر مطلوبی را برای تصمیمات ارائه دهد [۵]. Srinivas و همکاران بیان داشتند، صنعت مراقبت‌های بهداشتی مقدار زیادی از داده‌های مراقبت‌های بهداشتی را جمع‌آوری می‌کند که متأسفانه، برای استخراج اطلاعات پنهان منحصر به فرد نیست. در این مطالعه، به طور خلاصه با استفاده بالقوه از تکنیک‌های داده‌کاوی مبتنی بر طبقه‌بندی قوانین، درخت تصمیم‌گیری، بیز ساده و شبکه عصبی مصنوعی، حجم گسترده‌ای از داده‌های مراقبت‌های بهداشتی جهت تشخیص حمله قلبی، بررسی شد و به این نتیجه حاصل شد که الگوریتم ساده بیزی عملکرد بهتری برای تشخیص حمله قلبی دارد [۶]. اسماعیلی و همکاران با کمک شبکه عصبی مصنوعی به ارزیابی بیماری عروق کرونری قلب پرداختند [۷]. زمان‌پور و شمسی برای یافتن روشی جهت تشخیص و پیش‌بینی بیماری‌های قلبی از روش‌های درخت تصمیم‌گیری، بیز ساده، شبکه عصبی، K نزدیک‌ترین همسایه و طبقه‌بندی کننده انجمنی وزنی (Weighted Associative Classifier) استفاده کردند و یافتند که الگوریتم‌های بیز ساده و درخت تصمیم‌گیری مدل بهتری هستند [۸]. Lahsasna و همکاران به تشخیص بیماری عروق کرونری پرداختند. آن‌ها یک سیستم مبتنی بر قاعده فازی (Fuzzy Rule-Based System) طراحی نمودند که به عنوان یک سیستم پشتیبانی تصمیم برای تشخیص بیماری قلب و عروق کرونری (Coronary Heart Disease) طراحی شده است که نه تنها دقت تصمیم‌گیری قوانین را در نظر می‌گیرد، بلکه شفافیت آن‌ها را در همان زمان نیز در نظر می‌گیرد. از الگوریتم ژنتیک چند هدفه برای بهینه‌سازی آن‌ها استفاده شد. علاوه بر این، یک روش جدید برای استفاده از روش استراتژی طبقه‌بندی گروهی (Group Classification Strategy) برای ارتقای توانایی طبقه‌بندی FRBS پیشنهاد شده است. نتایج نشان می‌دهد که قوانین تولیدشده از لحاظ انسانی قابل فهم هستند، درحالی‌که دقت آن‌ها با سایر روش‌های طبقه‌بندی معیار مناسب است. علاوه بر این، FRBS تولیدشده قادر به شناسایی موارد عدم اطمینان است تا پزشک بتواند با آن‌ها مقابله کند و این امر منجر به مدیریت بهتر تلاش‌ها و وظایف خواهد شد. علاوه بر این، استفاده از ECS به طور خاص توانایی FRBS را برای تشخیص بیماران مبتلا به CHD بهبود بخشید که ویژگی مطلوب برای هر سیستم تشخیص CHD است و در نهایت با استفاده از قواعد فازی، تشخیص دقیق‌تر انجام شد [۹]. Anooj یک سیستم حمایت از تصمیم‌گیری بالینی مبتنی بر قاعده وزن فازی (CDSS) برای تشخیص بیماری قلبی ارائه کرد. سیستم تصمیم پیشگیری بالینی برای پیش‌بینی خطر ابتلا به بیماران قلب شامل دو مرحله است: (۱) رویکرد خودکار برای تولید قوانین فازی وزن و (۲) ایجاد یک سیستم حمایت از تصمیم‌گیری مبتنی بر قاعده فازی. در مرحله اول، برای استخراج قوانین فازی وزن، از تکنیک استخراج، انتخاب صفت و روش وزن‌دهی استفاده شد. سپس سیستم فازی با توجه به قوانین فازی وزن و ویژگی‌های انتخاب‌شده ساخته شد [۱۰]. اخیل جبار (Akhil Jbbar) بیان داشت؛ بیماری قلبی بزرگ‌ترین علت مرگ در کشورهای توسعه یافته است. از این‌رو

[Downloaded from jhbmi.com on 2026-11-15]

نیاز به ایجاد یک سیستم پشتیبانی تصمیم برای پیش‌بینی بیماری قلبی بیمار وجود دارد. در این پژوهش، الگوریتم طبقه‌بندی کارایی سازمانی با استفاده از رویکرد ژنتیکی برای پیش‌بینی بیماری‌های قلبی پیشنهاد شد. انگیزه اصلی برای استفاده از الگوریتم ژنتیک در کشف قوانین پیش‌بینی سطح بالا این است که قوانین کشف‌شده، با دقت پیش‌بینی شده، بسیار قابل فهم هستند و در این پژوهش سیستم پیشگیری بیماری‌های قلبی با استفاده از تکنیک‌های داده‌کاوی پیشنهاد شد. با استفاده از الگوریتم ژنتیک، با کاهش ویژگی‌ها و تعیین ویژگی‌ای که به تشخیص بیماری کمک می‌کند به پیش‌بینی پرداخته شد [۱۱]. طایفی تنها از درخت تصمیم برای پیش‌بینی بیماری عروق کرونری استفاده کرد که از نتیجه راضی بود [۱۲]. لطفی‌نژاد و همکاران با هدف تفکیک سیگنال‌های EEG (Electroencephalography) فرد بیمار صرع از فرد سالم از روش شبکه‌های عصبی پرسپترون چند لایه (Multi Layer Perceptron) آموزش دیده با الگوریتم کلاسیک و شبکه عصبیMLP آموزش دیده با الگوریتم هوشمند ژنتیک و همچنین شبکه عصبی با توابع شعاعی استفاده کردند و به این نتیجه رسیدند که شبکه عصبی MLP آموزش دیده با الگوریتم ژنتیک دارای بالاترین صحت است [۱۳]. صفدری و همکاران با استفاده از درخت تصمیم و شبکه عصبی به پیش‌بینی احتمال ابتلا افراد به آنفارکتوس قلبی پرداختند که درخت تصمیم مدل بهتری برای پیش‌بینی شد [۱۴]. کلاهی و رافع از روش نزدیک‌ترین همسایه مجاور به همراه انتخاب ویژگی برای تشخیص و پیش‌بینی سرطان سینه استفاده کردند و مدل پیشنهادی آن‌ها دارای عملکرد مناسبی بود [۱۵].

Kaan Uyar در مطالعه خود یک الگوریتم ژنتیک (Genetic Algorithm) مبتنی بر شبکه‌های عصبی مصنوعی فازی را برای تشخیص بیماری‌های قلبی ارائه داد. در این مطالعه، داده‌ها از دانشگاه کالیفرنیا ایروین (UCI) گرفته شده است. از مجموع ۲۹۷ نمونه از اطلاعات بیمار، ۲۵۲ مورد برای آموزش استفاده شده و ۴۵ نفر از آن‌ها برای آزمایش انتخاب شده‌اند. علاوه بر مشخصه دقت، مشخصه‌های دیگری نیز محاسبه می‌شوند، ازجمله خطای متوسط مربع ریشه، احتمال خطای اشتباه طبقه‌بندی، خاصیت، حساسیت، دقت و F-score [۱۶].

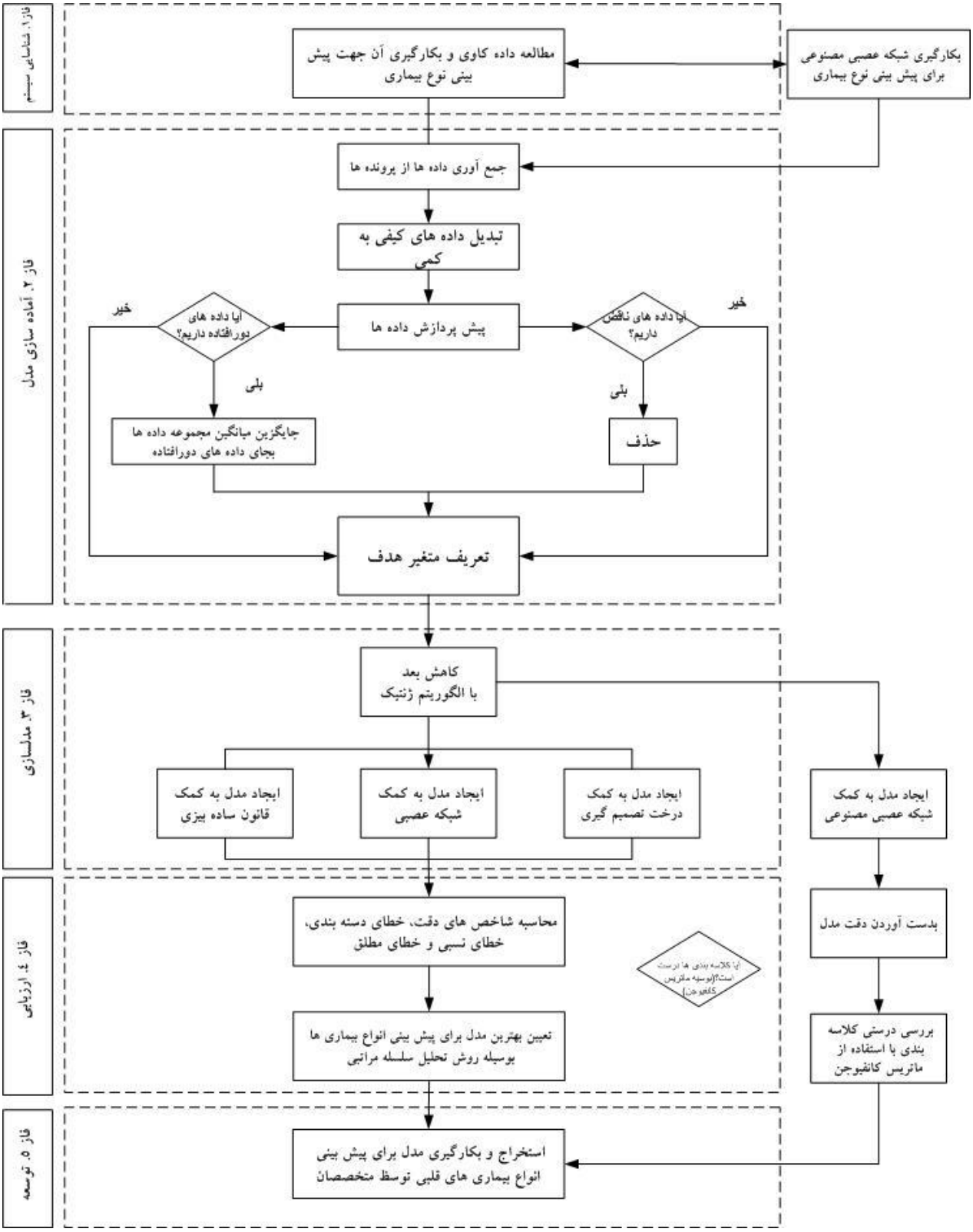
در تمامی پژوهش‌ها به تشخیص بیماری قلبی پرداخته شد، درصورتی‌که تشخیص نوع بیماری قلبی نیز از اهمیت بسیار بالایی برخوردار است و پیش‌بینی به موقع آن‌ها روند درمان بیمار را تسریع می‌بخشد. مدلی برای پیش‌بینی و تشخیص انواع بیماری‌های قلبی وجود ندارد، به همین دلیل ضرورت انجام این کار احساس شد. به کمک متخصصین قلب انواع بیماری‌های قلبی در ۵ دسته تقسیم‌بندی و پیش‌بینی با استفاده از این تقسیم‌بندی صورت گرفت و از داده‌های بالینی، آزمایشگاهی و دموگرافیک یک مدل طراحی شد.

روش

در این پژوهش برای تکنیک‌های داده‌کاوی از نرم‌افزار Rapidminer استفاده شد. این نرم‌افزار یک ابزار داده‌کاوی متن‌باز است که با زبان جاوا نوشته شده و از سال ۲۰۰۱ تا به حال توسعه داده شده است. در این نرم‌افزار سعی تیم توسعه‌دهنده، بر آن بوده است که تا حد امکان تمامی الگوریتم‌های رایج داده‌کاوی و همچنین یادگیری ماشین پوشش داده شوند. همچنین برای شبکه عصبی مصنوعی از نرم‌افزار Matlab نسخه ۲۰۱۶ استفاده شد که امکان آسان عملیات ماتریسی، محاسباتی و توابعی، استفاده از الگوریتم‌های مختلف و همچنین امکان ارتباط آسان با زبان‌های مختلف برنامه‌نویسی را به کاربر می‌دهد و دارای طیف کاربردی گسترده‌ای است، از جمله می‌توان به سیگنال و پردازش تصویر، ارتباطات، طراحی کنترلر، تست و اندازه‌گیری، مدل‌سازی مالی و تجزیه‌وتحلیل و زیست‌شناسی محاسباتی اشاره نمود [۳].

یکی از روش‌های بسیار قوی برای پیاده‌سازی و اجرای پروژه‌های داده‌کاوی روش (Standard Industry Cross-For Data Mining Process) CRISP است [۴،۱۷]. در مطالعه حاضر مدل پیشنهادی بر اساس CRISP که شامل پنج فاز است، ارائه شد. هریک از این فازها خود شامل زیر بخش‌هایی می‌شوند و ورودی هر فاز به خروجی فاز مرحله قبل وابسته است.

[Downloaded from jhbmi.com on 2026-11-15]



شکل ۱: مدل مفهومی پژوهش و ۵ فاز استاندارد CRISP در روش داده‌کاوی

چارچوب انجام پژوهش در شکل ۱ نشان داده شد. اطلاعات دموگرافیک، آزمایشگاهی و بالینی بیماران از بخش قلب بیمارستان فوق تخصصی قائم کرج جمع‌آوری گردید. اطلاعات تمام بیماران پیش‌پردازش شده و مقادیر آن‌ها جهت استفاده

[Downloaded from jhbmi.com on 2026-11-15]

مدل‌ها آماده‌سازی شده سپس در نرم‌افزار Rapidminer برای مدل‌سازی مورداستفاده قرار گرفت.

با استفاده از مطالعات و پژوهش‌های قبلی و نظر متخصصین قلب، متغیرها، ویژگی‌ها و عوامل مؤثر بر تصمیم‌گیری برای تشخیص بیماری‌های قلبی انتخاب شد. اطلاعات موجود در بیمارستان، به صورت نتایج آزمایشگاهی از قبیل آزمایش خون و نوار قلب و .... مورداستفاده قرار گرفت و از تکنیک‌های شبکه‌های عصبی مصنوعی در فضای Matlab و الگوریتم‌های داده‌کاوی در نرم‌افزار Rapidminer استفاده شد.

متغیرهای مدل عبارت‌اند از:

متغیرهای مستقل: متغیرهای بالینی، متغیرهای آزمایشگاهی، متغیرهای دموگرافیک

متغیرهای وابسته: نوع بیماری

ابتدا داده‌های توصیفی به کمی تبدیل شد. هر مجموعه داده‌ای دارای یکسری داده‌های دورافتاده است، بدین معنا که فاصله این داده‌ها از کل مجموعه داده‌ها بیشتر است و در یک دسته قرار نمی‌گیرند. داده‌های دورافتاده می‌توانند در پیش‌بینی نادرست پژوهش تأثیرگذار باشند. سپس نقاط دورافتاده از مجموعه داده حذف و با متغیرهای دیگر به جزء متغیر دور افتاده ادامه داده شد. در بین مجموعه داده، بعضی از مقادیر از دست‌رفته است، یعنی آن مقدار وجود ندارد. در اینجا سه راه‌حل موجود می‌باشد؛ جایگزین مقدار از دست‌رفته با مقدار دیگری (مثل میانگین، میانه، ماکزیمم، مینیمم و ...)، حذف کل متغیر و یا چشم‌پوشی کردن از آن که از روش اول در این پژوهش استفاده شد.

برای تشخیص نوع بیماری قلبی از بین ۵ نوع بیماری موجود در این پژوهش، از مدل‌های شبکه عصبی مصنوعی، درخت تصمیم‌گیری، الگوریتم ساده بیزی و روش انتخاب ویژگی الگوریتم ژنتیک استفاده شد. هر سه مدل مقادیر پارامتر دقت، خطای دسته‌بندی، خطای مطلق (میانگین انحراف مطلق از پیش‌بینی مقدار واقعی را نمایش می‌دهد) و خطای نسبی (میانگین انحراف مطلق از پیش‌بینی مقدار واقعی، تقسیم بر مقدار واقعی) به دست آمده و در آخر باهم مقایسه شد.

در پژوهش‌هایی که تعداد متغیرها زیاد است یا به عبارتی بُعدهای موردبررسی زیاد است، کاهش بعد انجام می‌شود. یکی از روش‌های کاهش بعد، الگوریتم ژنتیک است. در این پژوهش از الگوریتم ژنتیک برای انتخاب متغیرهای مهم و تأثیرگذار استفاده شد. پس از کاهش بُعد، مدل‌سازی شد. برای مدل‌سازی روی داده‌ها از عملگر اعتبارسنجی استفاده شد که شامل زیر فرایند است. مدل‌ها در زیر فرآیند این عملگر اعمال می‌شود.

عملگر شبکه عصبی، مدلی را با استفاده از شبکه عصبی آموزش‌یافته به‌وسیله الگوریتم پس انتشار خطا، یاد می‌گیرد. معماری این نوع شبکه عصبی مصنوعی اصطلاحاً پرسپترون چندلایه نامیده می‌شود. برای این مدل متغیرها باید عددی باشند پس توسط یک عملگر تبدیل اسمی به عددی، متغیرها تبدیل به متغیر عددی شدند. این مدل در پژوهش حاضر دارای؛ تعداد لایه‌های مخفی: ۹، تعداد چرخه‌های مورد استفاده برای **آموزش شبکه عصبی:** ۵۰۰ و نرخ آموزش: ۰/۳ می‌باشد.

عملگر درخت تصمیم، با قابلیت یادگیری از انواع داده ورودی (هم داده‌های اسمی و هم عددی)، یک درخت تصمیم را به‌عنوان مدل خروجی می‌سازد. این مدل در این پژوهش دارای ویژگی‌های زیر می‌باشد؛

حداقل اندازه یک گره برای آن که بتوان در آن شکست ایجاد کرد: ۴، حداقل تعداد رکورد در هر برگ: ۲، حداقل بهره: ۰/۱، حداکثر عمق درخت: ۲۰، درجه اطمینان: ۰/۲۵ می‌باشد.

این عملگر، با استفاده از روش توزیع‌های نرمال تخمینی و بر اساس نظریه بیز، مدلی را برای دسته‌بندی رکوردهای آتی در اختیار قرار می‌دهد. این عملگر در نرم‌افزار Rapidminer به «بیز ساده» معروف است.

و در آخر خروجی‌های مدل با استفاده از روش تحلیل سلسله مراتبی، رتبه‌بندی و اولویت مدل‌ها جهت پیش‌بینی بیماری مشخص شد.

علاوه بر الگوریتم داده‌کاوی، به طور جداگانه از روش شبکه عصبی مصنوعی نیز استفاده گردید. شبکه عصبی موردنظر دارای یک لایه پنهان و یک لایه خروجی می‌باشد. نوع شبکه عصبی پیش‌خور، نوع تابع شبکه عصبی الگوریتم لونبرگ مارکورت و سیستم یادگیری آن، پس انتشار خطا می‌باشد. داده‌ها به دو گروه آموزش و آزمایش تقسیم شد؛ به طوری که ۸۰٪ داده‌ها، در گروه آموزش و ۲۰٪ در گروه آزمایش جای گرفتند. پس از نوشتن مدل شبکه عصبی در نرم‌افزار Matlab، مدل را اجرا نموده و خروجی واقعی با خروجی شبکه عصبی مصنوعی مقایسه شد. سپس دقت مدل به دست آمد. سپس با استفاده از ماتریس کانفیوجن، درستی روش بررسی شد.

نتایج

برای جمع‌آوری داده‌های موردنیاز برای انجام این پژوهش، به بیمارستان فوق تخصصی قائم کرج مراجعه شد. در یک دوره

[Downloaded from jhbmi.com on 2026-11-15]

شش‌ماهه پرونده ۷۰۰ بیمار موردبررسی قرار گرفت. برای روش داده‌کاوی از اطلاعات تمامی پرونده‌ها و برای روش شبکه عصبی با استفاده از جدول مورگان و فرمول کوکران حجم نمونه برابر با ۲۴۸ به دست آمد که از همین تعداد برای شبکه عصبی استفاده شد. ۲۴ متغیر آزمایشگاهی، دموگرافیک و بالینی معرفی و بررسی شد، (جدول ۱). به دلیل زیاد بودن با استفاده از الگوریتم ژنتیک کاهش بعد انجام شد تا نتایج بررسی دقیق‌تر شود و تعداد متغیرهای پژوهش از ۲۴ به ۶ کاهش یافت. شش متغیر مهم و تأثیرگذار در این پژوهش عبارت است از: کلسترول، (Creatinine Kinase Muscle Brain) CPKMB، کلسترول با چگالی پایین (Low density Lipoprotein)LDL، فشارخون دیاستولیک، سیگار و سابقه فشارخون.

بررسی اطلاعات بیماران نشان می‌دهد؛ زنان بیشتر از مردان مبتلا به بیماری قلبی هستند. میانگین سنی بیماران مورد مطالعه ۶۲ سال بود. حدود ۷۵٪ از کسانی که مبتلا به تنگی نفس بودند بالای ۶۰ سال سن داشتند؛ بنابراین احتمال ابتلا به تنگی نفس در بین افراد بالای ۶۰ سال بیشتر است. در این پژوهش ۵۸/۳٪ زنان بیمار مبتلا به دیابت و ۴۱/۷٪ از مردان بیمار مبتلا به دیابت هستند. با توجه به داده‌های موجود در پژوهش حاضر، ۳۰٪ درصد از کسانی که مبتلا به بیماری قلبی هستند، سیگار مصرف می‌کنند. ۵۷/۲ درصد از بیماران بالای ۵۹ سال دارای فشارخون هستند و ۸۳/۳۳ درصد از بیماران بالای ۵۰ سال سیگار مصرف می‌کنند. جدول ۲ فراوانی بیماران را بر اساس نوع بیماری نشان می‌دهد.

جدول ۱: متغیرهای پژوهش

متغیرهای دموگرافیک	جنس	متغیرهای بالینی	تعداد روزهای بستری
متغیرهای بالینی	فشارخون سیستولیک	**متغیرهای آزمایشگاهی**	FBS
	فشارخون دیاستولیک		CHOL
	نبض		TG
	دیابت		HDL
	سیگار		LDL
	درد در قفسه سینه		CPK
	سابقه تنگی نفس		CPKMB
	سابقه سکته قلبی		Hb
	سابقه بیماری فامیلی		PT
	سابقه بیماری قلبی		PTT
	سابقه فشارخون		Trop i

جدول ۲: فراوانی انواع بیماری‌ها

نوع بیماری	بیماری‌های مادرزادی قلب در بزرگ‌سالان	بیماری‌های دریچه‌ای قلب	بیماری‌های نارسایی قلبی	بیماری‌های آنژین	بیماری آریتمی‌ها
فراوانی	۰/۱	۰/۲۴	۰/۱۶	۰/۱۵	۰/۳۵

مطابق با فاز سوم الگوریتم شکل ۱ پارامترهای دقت، خطای دسته‌بندی، خطای مطلق و خطای نسبی از حل الگوریتم‌های داده‌کاوی به دست می‌آید. جدول ۳ مقادیر پارامترها را نشان می‌دهد.

[Downloaded from jhbmi.com on 2026-11-15]

جدول ۳: نتایج حاصل از مدل‌های داده‌کاوی

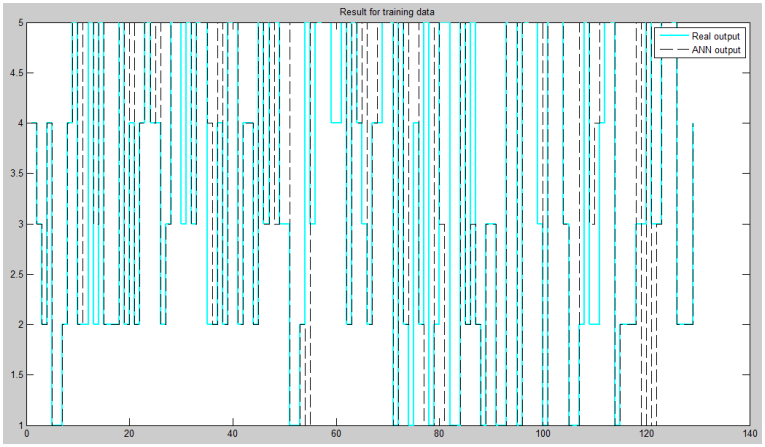
	دقت	خطای دسته‌بندی	خطای مطلق	خطای نسبی
شبکه عصبی	۲۸٪/۴۱	۷۱٪/۵۹	۰٪/۷۹	۷۹٪/۰۳
درخت تصمیم‌گیری	۲۸٪/۴۱	۷۱٪/۵۹	۰٪/۷۹	۷۹٪/۰۳
الگوریتم ساده بیزی	۲۳٪/۶۵	۷۶٪/۳۵	۰٪/۷۹۶	۷۹٪/۵۶

با روش تحلیل سلسله مراتبی، مدل‌های پیش‌بینی بیماری رتبه‌بندی شد. در هر جدول مقایسه‌های زوجی، نرخ سازگاری بررسی شد و مقدار آن کمتر از ۰/۱ و مورد قبول بود. وزن نهایی گزینه‌ها به ترتیب عبارت‌اند از؛ شبکه عصبی ۰/۴۲، درخت تصمیم ۰/۱۷ و الگوریتم ساده بیزی ۰/۳۲.

در این قسمت مطابق با الگوریتم بیان شده شکل ۱ با استفاده از شبکه عصبی به پیش‌بینی انواع بیماری‌های قلبی پرداخته شد. از پرونده ۲۴۸ بیمار استفاده شد که از طریق روش نمونه‌گیری کوکران به دست آمد. $n = \frac{z^2pq}{d^2}/1 + \frac{1}{N}(\frac{z^2pq}{d^2} - 1)$. در این فرمول p و q نسبت موفقیت و شکست هستند که ۰/۵ در نظر گرفته می‌شوند. مقدار $Z_{\alpha/2}$ در سطح خطای ۰/۰۵ برابر ۱/۹۶ است. مقدار خطای d نیز ۰/۰۵ در نظر گرفته می‌شود. مقدار N معرف حجم جامعه موردنظر است.

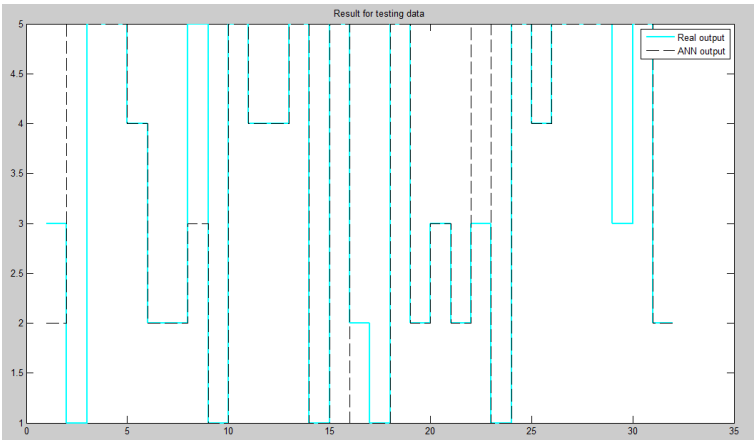
با استفاده از جدول مورگان و فرمول کوکران حجم نمونه برابر است با: ۲۴۸/۲۶. (این محاسبه با سطح خطای ۵ درصد صورت گرفته است).

شبکه عصبی مورداستفاده دارای یک لایه پنهان و یک لایه خروجی می‌باشد. نوع شبکه عصبی پیش‌خور، نوع تابع شبکه عصبی الگوریتم لونبرگ مارکورت و سیستم یادگیری آن، پس انتشار خطا می‌باشد. داده‌ها را به دو گروه آموزش و آزمایش تقسیم شدند. به طوری که ۸۰٪ داده‌ها، در گروه آموزش و ۲۰٪ در گروه آزمایش جای گرفتند. شکل ۲، نتیجه اجرای مدل شبکه عصبی در نرم‌افزار Matlab بر روی داده‌های آموزش و شکل ۳ نتیجه اجرای مدل شبکه عصبی بر روی داده‌های آزمایش را نشان می‌دهند. خطوط آبی خروجی واقعی و خطچین‌ها خروجی شبکه عصبی هستند.



شکل ۲: نتیجه شبکه عصبی مصنوعی بر داده‌های آموزش برای تشخیص بیماری

[Downloaded from jhbmi.com on 2026-11-15]



شکل ۳: نتیجه شبکه عصبی مصنوعی بر داده‌های آزمایش برای تشخیص بیماری

در قسمت‌هایی از شکل که خطوط آبی روی خط‌چین قرار گرفته است، نشان‌دهنده این است که پیش‌بینی درست است؛ اما برای محاسبه دقیق‌تر دقت مدل محاسبه شد. با توجه به جدول ۵، نتایج به دست‌آمده مورد تأیید است. در قسمت‌هایی از شکل که خطوط آبی روی خط‌چین قرار گرفته است، نشان‌دهنده این است که پیش‌بینی درست است؛ اما برای محاسبه دقیق‌تر دقت مدل محاسبه شد. با توجه به جدول ۴، نتایج به دست آمده مورد تأیید می‌باشد.

جدول ۴: دقت‌های به دست آمده از شبکه عصبی

correctnessPercentageTest	correctnessPercentageTrain
۸۱/۲۵۰۰	۸۱/۳۹۵۳

ماتریس کانفیوجن (Confusionmat) ماتریسی که نشان می‌دهد کلاسه‌بندی‌ها تا چه حد درست یا نادرست انجام شده است.

در شکل ۴، برای هر دو داده آموزش و آزمایش این ماتریس به دست می‌آید.

conf_Mat_tst =

4	0	0	0	1
1	6	0	0	0
0	1	1	0	2
0	0	0	4	0
0	0	1	0	11

conf_Mat_trn =

13	0	0	0	1
3	27	0	2	2
0	0	8	0	10
0	0	0	17	0
0	0	4	2	40

شکل ۴: ماتریس کانفیوجن

حالت بهینه این ماتریس زمانی است که همه اعداد صفر باشند، جزء اعداد روی قطر اصلی. مثلاً در ماتریس داده‌های آموزش، ۱۶ رکورد دارای بیماری نوع یک هستند که ۱۳ رکورد درست تشخیص داده شده است و ۳ رکورد اشتباه. ۲۷ رکورد دارای بیماری نوع دوم هستند که همه آن‌ها درست تشخیص داده شده‌اند.

بحث و نتیجه‌گیری

در این مطالعه، مجموعه‌ای از متغیرهای مؤثر در بروز بیماری‌های قلبی بر اساس پایگاه داده بیمارستان مورد پردازش قرار گرفت. استخراج اطلاعات و مقادیر متغیرها از پرونده بیماران قلبی در بیمارستان فوق تخصصی قائم کرج جمع‌آوری شد. پژوهش‌های مربوط محدودیت‌هایی از قبیل تعداد کم بیماران برای ایجاد مدل، داده‌های از دست‌رفته و متغیرهای

[Downloaded from jhbmi.com on 2026-11-15]

ناقص را دارا می‌باشند. در این پژوهش تعداد قابل قبولی از بیماران با متغیرهای مناسب و کافی با حداقل داده‌های از دست‌رفته به کار گرفته شد. مطالعه حاضر با به کارگیری ویژگی‌های بالینی و آزمایشگاهی بیماران با استفاده از الگوریتم‌های داده‌کاوی و شبکه عصبی، نوع بیماری قلبی را پیش‌بینی نمود. این مطالعه شامل پیش‌پردازش داده‌ها، کاهش بعد و انتخاب متغیرهای تأثیرگذار و معرفی مدل بهتر برای پیش‌بینی نوع بیماری قلبی است. از بین متغیرها، CHOL، LDL، CPKMB، فشار خود دیاستولیک، سیگار و سابقه فشارخون دارای تأثیرگذاری بیشتری بودند و متغیرهایی مثل سن از تأثیر کمتری برخوردار بودند. در این مطالعه علاوه بر شناسایی مهم‌ترین متغیرها، با استفاده از انتخاب ویژگی‌ها بر اساس الگوریتم ژنتیک به عملکردی بهتر بر حسب شاخص‌های دقت دست یافته شد.

در پژوهش حاضر یک بار توسط الگوریتم‌های داده‌کاوی طراحی مدل انجام شد که در این قسمت جامعه آماری ۷۰۰ بیمار بود. یک بار هم به صورت جداگانه از شبکه عصبی برای پیش‌بینی نوع بیماری استفاده شد که نمونه آماری شامل ۲۴۸ بیمار بود.

در این پژوهش، مدل بر اساس داده‌های آموزشی که ۸۰ درصد کل داده‌ها را تشکیل داده‌اند، ساخته شد، سپس توسط داده‌های آزمایشی که ۲۰ درصد داده‌های بودند آزمایش شد و الگوریتم‌های داده‌کاوی با هم مقایسه شدند. نتایج حاکی از این بود که از میان الگوریتم‌های داده‌کاوی، شبکه عصبی مصنوعی مدل بهتری برای پیش‌بینی نوع بیماری‌های قلبی می‌باشد. همچنین در تحلیل شبکه عصبی به طور جداگانه، مدل دارای دقت ۸۱ درصد برای پیش‌بینی بود.

پژوهشی که مشابه مطالعه حاضر، نوع بیماری قلبی را پیش‌بینی کرده باشد یافت نشد؛ اما Kaan Uyar در پژوهش خود یک الگوریتم ژنتیک (GA) مبتنی بر شبکه‌های عصبی مصنوعی فازی را برای تشخیص بیماری‌های قلبی ارائه داد که از مجموع ۲۹۷ نمونه از اطلاعات بیمار، ۲۵۲ مورد برای آموزش استفاده شده و ۴۵ نفر از آن‌ها برای آزمایش انتخاب شده‌اند. مدل پیشنهادی آن با دقت ۹۷/۷۸ درصد پیش‌بینی انجام داد [۱۶]. Anooj یک سیستم حمایت از تصمیم‌گیری بالینی مبتنی بر قاعده وزن فازی (CDSS) برای تشخیص بیماری قلبی ارائه کرد. در نهایت، آزمایش بر روی سیستم پیشنهاد شده با استفاده از مجموعه داده‌های به دست‌آمده از مخزن UCI انجام شد و عملکرد سیستم با سیستم مبتنی بر شبکه عصبی با استفاده از دقت، حساسیت و خاصیت مقایسه شد که شبکه عصبی با دقت ۵۲/۴۷ درصد عملکرد بهتری برای تشخیص داشت [۱۰]. Srinivas و همکاران، با استفاده بالقوه از تکنیک‌های داده‌کاوی مبتنی بر طبقه‌بندی قوانین، درخت تصمیم‌گیری، بیز ساده و شبکه عصبی مصنوعی، حجم گسترده‌ای از داده‌های مراقبت‌های بهداشتی جهت تشخیص حمله قلبی را بررسی نمودند و نتایج نشان داد الگوریتم ساده بیزی با دقت بالای ۸۳ درصد عملکرد بهتری برای تشخیص حمله قلبی دارد [۶].

در هر سه این پژوهش‌ها تنها به پیش‌بینی بیماری قلبی از بین تمام بیماری‌ها پرداخته شد در صورتی که در پژوهش حاضر تنها بیماران قلبی مورد بررسی قرار گرفته‌اند و در تشخیص نوع بیماری قلبی، شبکه عصبی به عنوان مناسب‌ترین مدل پیش‌بینی شناخته شد. تفاوت دیگر در تعداد متغیرهای مورد بررسی است. در پژوهش‌های بیان شده حداکثر تا ۱۳ متغیر بررسی شد در صورتی که در این مقاله ۲۴ متغیر شناسایی و بررسی شد.

در این پژوهش، برای مقایسه مدل‌های استفاده شده در داده‌کاوی علاوه بر شاخص دقت، سه شاخص خطای دسته‌بندی، خطای مطلق و خطای نسبی نیز اندازه‌گیری شد و مقایسه دقیق‌تری صورت گرفت و در نهایت شبکه عصبی به عنوان بهترین مدل انتخاب شد. اگر فقط شاخص دقت در نظر گرفته شود؛ مدل شبکه عصبی دارای دقت ۲۸/۴۱ درصد می‌باشد که در مقایسه با دقت‌های به دست‌آمده در مقالات مرتبط به ظاهر پایین است؛ اما باید توجه داشت که مقالات دیگر تنها به بررسی اینکه بیمار، بیماری قلبی دارد یا خیر پرداختند، یعنی دو حالت داشتند؛ اما در مقاله حاضر نتیجه ۵ حالت دارد؛ این که بیمار قلبی کدام پنج حالت را دارا است. به همین خاطر مقدار عددی دقت این مطالعه پایین است؛ اما این دقت بنا به شرایط این پژوهش، دقت خوبی است.

از کاستی‌های این پژوهش می‌توان به عدم در نظر گرفتن همه ۴۱ نوع بیماری قلبی اشاره کرد. به علت سخت بودن پیش‌بینی و به دست آمدن دقت پایین، این امر امکان‌پذیر نبود. پیشنهاد می‌شود در پژوهش‌های آتی، علاوه بر بررسی متغیرهای احتمالی مؤثر بر بیماری‌های قلبی خاص، روشی برای افتراق ۴۱ نوع بیماری قلبی نیز ارائه گردد.

تشکر و قدردانی

بدین‌وسیله از پزشکان، کارکنان و مدیریت محترم بیمارستان فوق تخصصی قائم کرج که با همکاری خود، ما را در این

[Downloaded from jhbmi.com on 2026-11-15]

پژوهش یاری کردند، تشکر و قدردانی می‌کنیم.

تعارض منافع

نویسندگان اعلام می‌دارند که هیچ‌گونه تعارض منافعی ندارند.

References

1. Azizi F, Esmaeilzadeh A, Mirmiran P. Obesity and cardiovascular risk factors: an epidemiological study in Tehran. Iranian Journal of Endocrinology and Metabolism 2004; 5(4): 389-97. Persian

2. Kasper DL, Fauci AS, Hauser SL, Longo DL, Larry Jameson J, Loscalzo J. Harrison's Internal Medicine Handbook. Translyted by Abdolreza Mansouri Rad, Mahshid Nik Pour, Tehran: Arjmand; 2013. Persian

3. Saniee Abadeh M, Mahmoodi S, Taherparvar M. Applied Data Mining. Tehran: Niaze-Danesh; 2014. Persian

4. Ghazanfari M, Alizadeh S, Timurpour B. Data Mining and Discovery. Tehran: Science and Technology Publication; 2013. Persian

5. Tsipouras MG, Exarchos TP, Fotiadis DI, Kotsia AP, Vakalis KV, Naka KK, et al. Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. IEEE Trans Inf Technol Biomed 2008;12(4):447-58.

6. Srinivas K, Kavihta B, Govrdhan A. Applications of data mining techniques in healthcare and prediction of heart attacks. International Journal on Computer Science and Engineering 2010; 2(2): 250-5.

7. Mohammadpour RA, Esmaeili MH, Ghaemian A, Esmaeili J. Application of Artificial Neural Network for Assessing Coronary Artery Disease. J Mazand Univ Med Sci 2012; 22(86): 9-17. Persian

8. Zamanpour S, Shamsi M. Study and comparison of the accuracy of data mining algorithms in the prediction of heart disease. 4th Iranian Conference on Electrical and Electronics Engineering; 2012 Aug 28-30; Gonabad: Islamic Azad University of Gonabad; Persian

9. Lahsasna A, Ainon RN, Zainuddin R, Bulgiba A. Design of a fuzzy-based decision support system for coronary heart disease diagnosis. J Med Syst 2012;36(5):3293-306.

10. Anooj PK. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. Journal of King Saud University - Computer and Information Sciences 2012;24(1):27-40.

11. Akhil jabbar M, Chandra P, Deekshatulu BL. Heart disease prediction system using associative classificationand genetic algorithm. International Conference on Emerging Trends in Electrical, Electronics and CommunicationTechnologies-ICECIT 2012; 1: 183-92.

12. Tayefi M, Tajfard M, Saffar S, Hanachi P, Amirabadizadeh AR, Esmaeily H, et al. hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm. Comput Methods Programs Biomed 2017;141:105-9.

13. Lotfali Nejad H, Keshvari R, Shabroo S, Moghaddas H. Comparison of MLP and RBF Neural Networks in Epilepsy. 8th Symposium on Advances in Science and Technology Commission-IV, 2014 Feb 6; Mashhad: Khavaran Institute of Higher Education; 2014. Persian

14. Safdari R Ghazi saeedi M, Gharooni M, Nasiri M Argi G. Comparing performance of decision tree and neural network in predicting myocardial infarction. Journal of Paramedical Sciences & Rehabilitation 2014; 3(2): 26-35. Persian

15. Kalahi S, Rafa V. A solution to diagnosis of disease using data mining techniques, 1st National Conference on Electrical and Computer; 2013 May 13; Khormooj: Islamic Azad University, Khormoj Branch; 2013. Persian

16. Uyar K, İlhan A. Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks. Procedia Computer Science 2017;120:588-93.

17. Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, et al. CRISP-DM 1.0 Step-by-step data mining guide. Translayted by Esmaeeli S, Basiri M. Available from: http://www.farabar.net/wp-content/uploads/2016/12/@Farabar_BI-CRISP-DM.pdf

[Downloaded from jhbmi.com on 2026-11-15]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2019; 5(4): 457-468

Comparison of Three Decision-Making Models in Differentiating Five Types of Heart Disease: A Case Study in Ghaem Sub-Specialty Hospital

Seyedeh Raahil Mousavi[1*], Mohammad Mehdi Sepehri[2]

• **Received:** 23 May, 2018 • **Accepted:** 8 Sep, 2018

Introduction: cardiovascular diseases are becoming the main cause of mortality and morbidity in most countries. This research goal was to predict the types of heart diseases for more accurate diagnosis by data mining and neural network technics.
Method: This research was an applied-survey study and after data preprocessing, three approaches of neural network, decision making tree and Bayes simple algorithm were used to predict and recognize in Rapidminer software and neural artificial network model was used for prediction in Matlab software.
Results: Genetic algorithm was used for selection of effective variables and neural artificial network models, decision making tree and Bayes simple algorithm were used to predict types of heart diseases in data mining. AHP model was used to determine a model with the best performance for predicting types of heart diseases.
Conclusion: Neural network had much better performance than other data mining models used to diagnose types of heart diseases in this research. Also, in detecting disease by artificial neural network, the model with accuracy of more than 80 percent was verified as good and acceptable.

Keywords: Data mining, Neural network, Decision tree, Heart disease, Diagnosis

• **Citation:** Mousavi SR, Sepehri MM. Comparison of Three Decision-Making Models in Differentiating Five Types of Heart Disease: A Case Study in Ghaem Sub-Specialty Hospital. Journal of Health and Biomedical Informatics 2019; 5(4): 457-468.

1. PhD. Student in Industrial Engineering, Faculty of Technical and Engineering, Research and Science University, Tehran, Iran
2. Ph.D. in Operations Research, Professor, Industrial Engineering Dep., Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran

*Correspondence: Industrial Engineering, Faculty of Technical and Engineering, University of Science and Research, Tehran, Iran
• **Tel:** 09371036997 • **Email:** raahil.mousavi@yahoo.com

[Downloaded from jhbmi.com on 2026-11-15]