مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره پنجم شماره سوم. ص ۳۶۱-۳۷۲، ۱۳۹۷

بهینه‌سازی ماشین بردار پشتیبان با الگوریتم GBC جهت پیش‌بینی ارتباط آنفارکتوس قلبی و سرطان: مطالعه موردی

مهدی نوشیار[1*]، محمد مؤمنی[2]، ثریا غراوی[3]، فاطمه حورعلی[4]

• دریافت مقاله: ۹۷/۲/۲۸ • پذیرش مقاله: ۹۷/۷/۱۷

مقدمه: آگاهی از ابتلا به سرطان، احتمال بروز اختلالات عصبی و استرس در بیمار را افزایش می‌دهد. همچنین استرس خطر بروز آنفارکتوس قلبی را بیشتر می‌کند. مطالعه حاضر بر پایه الگوریتم GBC، به بررسی احتمال بروز سکته قلبی در بیماران سرطانی پرداخت.

روش: اطلاعات بیماران از پایگاه داده بیمارستان فوق تخصصی شهید صدوقی یزد جمع‌آوری شد. پرونده پزشکی ۱۶۷۹ بیمار مبتلا به آنفارکتوس قلبی مورد بررسی قرار گرفت که از این تعداد، ۸۱ بیمار مبتلا به سرطان بودند. در فرآیند انتخاب ویژگی توسط مدل پیشنهادی، اگر ابتلا به سرطان به عنوان ویژگی تأثیرگذار شناسایی شود، آنگاه ارتباط بین سرطان و آنفارکتوس قلبی معنادار خواهد بود.

نتایج: با استفاده از مدل پیشنهادی، ویژگی ابتلا به سرطان برای پیش‌بینی امکان وجود احتمال بروز سکته قلبی انتخاب شد که نشان‌دهنده وجود رابطه معنادار بین این دو ویژگی در بیماران مستعد عارضه قلبی می‌باشد. نتایج نشان داد با انتخاب ویژگی ابتلا به سرطان، دقت پیش‌بینی مدل پیشنهادی برابر با ۰/۹۱ بود.

نتیجه‌گیری: با استفاده از انتخاب ویژگی ابتلا به سرطان، مدل پیشنهادی نسبت به سایر مدل‌های مورد مقایسه دارای حداقل میزان خطا و بیش‌ترین دقت و صحت در پیش‌بینی آنفارکتوس قلبی است. روش نایو بیز، حداکثر میزان خطا و کم‌ترین دقت را دارا است. نتایج شبیه‌سازی حاکی از احتمال بروز سکته قلبی، در بیماران مستعد عارضه قلبی، پس از خبردار شدن از بیماری سرطان طی ماه‌های اولیه خود است.

کلید واژه‌ها: سرطان، آنفارکتوس قلبی، GBC، ماشین بردار پشتیبان، افزایش دقت پیش‌بینی

•ارجاع: نوشیار مهدی، مؤمنی محمد، غراوی ثریا، حورعلی فاطمه. بهینه‌سازی ماشین بردار پشتیبان با الگوریتم GBC جهت پیش‌بینی ارتباط آنفارکتوس قلبی و سرطان :مطالعه موردی، مجله انفورماتیک سلامت و زیست‌پزشکی ۱۳۹۷؛ ۵(۳): ۳۷۲-۳۶۱.

۱. دکتری تخصصی مهندسی برق، دانشیار، گروه کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه محقق اردبیلی، اردبیل، ایران
۲. دانشجوی دکتری مهندسی کامپیوتر، گروه کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران
۳. کارشناسی ارشد مهندسی کامپیوتر نرم افزار، مربی، دانشکده مهندسی کامپیوتر و برق، گروه کامپیوتر، مجتمع آموزش عالی اسفراین، خراسان شمالی، اسفراین، ایران
۴. کارشناسی ارشد مهندسی کامپیوتر نرم افزار، مربی، دانشکده مهندسی کامپیوتر و برق، گروه برق، مجتمع آموزش عالی اسفراین، خراسان شمالی، اسفراین، ایران
*** نویسنده مسئول:** اردبیل، انتهای خیابان دانشگاه، دانشگاه محقق اردبیلی
•شماره تماس: ۰۴۵۳۱۵۰۵۰۰۰
• **Email:** nooshyar@uma.ac.ir

[Downloaded from jhbmi.com on 2026-11-17]

مقدمه

سرطان، ناشی از رشد خارج از قاعده سلول‌های غیرطبیعی است. در هر دو نوع تومورهای خوش‌خیم و بدخیم، رشد سریع و زیاد سلول‌ها وجود دارد. روند زیاد شدن سلول‌ها در تومورهای خوش‌خیم در مرحله مشخصی متوقف می‌شود؛ اما در تومورهای بدخیم این رشد به صورت غیر قابل مهار ادامه می‌یابد تا حدی که در صورت عدم درمان، تمامی قسمت‌های بدن را تحت تأثیر قرار داده و از کار می‌اندازد [۱].

تردیدی نیست که تشخیص بیماری‌های تهدید کننده حیات چون سرطان، اثرات متعددی بر کیفیت زندگی فرد می‌گذارد. سرطان صرفاً یک رویداد با پایانی معین نیست، بلکه یک موقعیت دائمی مبهم است که با تأثیرات دیررس و تأخیری ناشی از بیماری، درمان و مسائل روان‌شناختی مربوط به آن‌ها مشخص می‌گردد [۲]. با توجه به این باور عمومی که خبر بد (مثلاً در مورد پیش‌آگهی بد بیماری) منجر به پریشانی روانی، غمگینی و اضطراب بیمار می‌شود [۳] ممکن است احتمال بروز سکته قلبی را افزایش دهد.

سکته قلبی یا آنفارکتوس توسط لخته درون رگ کرونری، جریان خون و اکسیژن‌رسانی به عضله قلبی را مختل می‌کند که سبب مرگ سلول‌های قلبی در آن ناحیه می‌شود. ماهیچه قلبی آسیب دیده، توانایی خود را برای انقباض از دست می‌دهد و عضله قلبی باقی‌مانده برای جبران منطقه آسیب دیده وارد عمل می‌شود. استرس شدید ناگهانی می‌تواند یک حمله قلبی را شروع کند. درد پشت جناق سینه (استخوان وسط سینه) یک علامت عمده حمله قلبی است که ممکن است دو دقیقه یا بیشتر طول بکشد؛ ولی در بسیاری افراد، به خصوص در افراد مسن و بیماران دیابتی، درد ممکن است ناچیز یا حتی اصلاً وجود نداشته باشد. اغلب درد از ناحیه سینه به بازوها یا شانه، گردن، فک، شکم یا پشت انتشار می‌یابد. گاهی اوقات درد تنها در یکی از این نواحی احساس می‌شود. درد ممکن است شدید و زیاد یا خفیف و نامشخص باشد یا به صورت فشار و سنگینی در قفسه سینه، احساس درد در معده، درد شدید در دست، تهوع و اضطراب باشد [۴،۵].

پیش‌بینی تقریبی آنفارکتوس قلبی با استفاده از ریسک فاکتورهای ماژور مثل استعمال سیگار، افزایش کلسترول خون، پرفشاری خون، دیابت و چاقی قابل انجام است و می‌توان سرعت پیشرفت بیماری یا علائم بیماری و عوارض آن را کنترل کرد. مسئله مهم در پیش‌بینی آنفارکتوس قلبی، تشابه علائم بالینی و آزمایشگاهی است که احتمال بروز خطا در تشخیص را افزایش می‌دهد؛ مثلاً آنفارکتوس قلبی فقط یکی از علل متعددی است که سبب بروز دردهای شدید قفسه سینه می‌شود. تشخیص این عارضه، به مدت زمان و شکل خاص درد بستگی کامل دارد. با این‌حال آزمایش‌های لازم برای تأیید تشخیص، ضروری است. کاهش بروز خطا در پیش‌بینی آنفارکتوس قلبی دارای اهمیت است [۶].

تشخیص و پیش‌بینی انواع بیماری‌ها با استفاده از ماشین بردار پشتیبان (Support Vector Machines) SVM امکان‌پذیر است. اساس ماشین‌های بردار پشتیبان، کمینه‌سازی خطای ساختاری طبق نظریه‌های یادگیری آماری است. یکی از موارد مهم در یادگیری، قابلیت تعمیم مناسب با استفاده از تعداد محدود داده آموزشی می‌باشد. در بیش‌تر بیماری‌ها داده‌های آموزشی کافی برای الگوریتم یادگیرنده وجود ندارند و یا فراهم نمودن آن زمان‌بر و مشکل است. فرآیند یادگیری برای پیش‌بینی بیماری‌ها با تعداد محدود داده آموزشی از طریق ماشین‌های بردار پشتیبان امکان‌پذیر است [۷].

زایری و همکاران [۸] مطالعه‌ای روی عوامل تأثیرگذار بر مرگ‌ومیر پس از عمل جراحی کرونری بای پاس در بیماران غیر وابسته به دیالیز انجام دادند. داده‌ها از یک مطالعه مقطعی از ۱۳۹۰ بیمار در مدت ۳ سال از بیمارستان قلب شریعتی جمع‌آوری شدند. در این تحقیق برای پیش‌بینی مرگ‌ومیر از مدل‌های رگرسیون لجستیک و رده‌بندی درختی با نرم‌افزارهای SPSS نسخه ۲۴ و CART استفاده شد و نتایج با یکدیگر مقایسه شده‌اند.

سجادی و همکاران [۹] با استفاده از روش‌های خوشه‌بندی در داده‌کاوی، عوامل خطرساز بروز عوارض قلبی در بیماران دیابتی را شناسایی کردند که ویژگی‌های افزایش فشارخون، افزایش شاخص توده بدنی و اختلالات چربی خون از عوامل تأثیرگذار در این بیماری شناخته شدند.

Parthiban و همکاران با استفاده از شبکه بیزین، احتمال بروز مشکلات قبلی را در بیماران پیش‌بینی کردند. برای این کار از مؤلفه‌هایی مانند سن، جنسیت، فشارخون و میزان قندخون استفاده کردند. دقت مدل ۷۴٪ بود [۱۰]. همچنین Vidic و همکاران با استفاده از ماشین بردار پشتیبان، مدلی برای پیش‌بینی احتمالات بروز مشکلات قلبی در بیماران دیابتی با دقت ۹۴٪ ایجاد نمودند [۱۱].

الگوریتم GBC(Genetic Bee Colony) رویکرد جدیدی است که توان افزایش دقت و کاهش خطا در پیش‌بینی را با محاسبات حداقل و مرتبه زمانی کمتر از کارهای مشابه دارد. این

[Downloaded from jhbmi.com on 2026-11-17]

الگوریتم روشی جدید با ترکیبی از الگوریتم ژنتیک و کلونی زنبور مصنوعی برای استفاده از مزایای هر دو الگوریتم است. به منظور دستیابی به یک تعادل بین بهره‌برداری و اکتشاف در الگوریتم ABC(Artificial bee colony algorithm) و بهبود توانایی‌های جستجوی محلی، از اپراتورهای جهش الگوریتم ژنتیک در طول فرآیند جایگزینی راه‌حل‌ها استفاده می‌شود [۱۲].

در این پژوهش برای پیش‌بینی آنفارکتوس قلبی در ابتدا پرونده‌های بیماران مبتلا به این بیماری جمع‌آوری و داده‌های موجود در پرونده‌ها مورد بررسی قرار گرفت. سپس متغیرهای مورد پردازش، انتخاب شدند. پس از پیش‌پردازش و آماده‌سازی مجموعه داده جهت پردازش نهایی، با به کارگیری الگوریتم GBC جهت بهینه‌سازی پارامترهای ماشین بردار پشتیبان، آنفارکتوس قلبی پیش‌بینی گردید. در این مقاله، مجموعه داده‌ها یک بار با به‌کارگیری ویژگی وضعیت ابتلا به سرطان و یک بار بدون احتساب این ویژگی شبیه‌سازی شدند تا ارتباط بین سرطان و سکته قلبی مشخص شود.

روش

گردآوری مجموعه داده‌ها: برای تهیه مجموعه داده‌های مربوط به ارتباط بین سرطان و آنفارکتوس قلبی، در ابتدا پرسشنامه استاندارد بر اساس ویژگی‌های بیماران مستعد عارضه قلبی طراحی شد، سپس در بیمارستان فوق تخصصی شهید صدوقی یزد تکمیل و جمع‌آوری گردید. داده‌ها مربوط به سال‌های ۱۳۹۳ تا ۱۳۹۵ است. این مجموعه داده شامل ۱۶۷۹ نمونه بوده است که از این تعداد، ۸۱ بیمار، مبتلا به سرطان بوده‌اند. برای هر بیمار تعداد ۲۵ ویژگی ثبت شده است. در این مقاله، داده‌ها یک بار با به‌کارگیری ویژگی وضعیت ابتلا به سرطان و یک بار بدون احتساب این ویژگی شبیه‌سازی شدند تا ارتباط بین سرطان و سکته قلبی مشخص شود. همچنین ویژگی‌های کمتر بررسی شده مانند افسردگی و سلامت روان برای پیش‌بینی ارتباط بین سرطان و آنفارکتوس قلبی به کار گرفته شد. بازه مقادیر اولیه ویژگی‌های بالینی بیمار در جدول ۱ نمایش داده شد. قالب مناسب داده‌ها به عنوان ورودی داده‌کاوی در نتایج و خروجی تأثیرگذار است. اگر مقادیر ویژگی‌های مجموعه داده در دامنه متفاوتی قرار داشته باشند، احتمال بروز خطا در یافته‌ها افزایش می‌یابد [۱۳]. در مدل پیشنهادی، نرمال‌سازی به روش Max/Min انجام شده و همه ویژگی‌ها در بازه [۱-۰] قرار دارند [۱۴].

ویژگی‌هایی مانند سابقه خانوادگی سکته قلبی، چربی خون، قند خون و کلسترول، احتمال ابتلا به سکته قلبی را نشان می‌دهد. در این پژوهش با استفاده از داده‌کاوی، ارتباط بین این ویژگی‌ها، بروز سکته قلبی و ابتلا به سرطان بررسی شد.

ماشین‌های بردار پشتیبان: اساس ماشین‌های بردار پشتیبان، کمینه‌سازی خطای ساختاری طبق نظریه‌های یادگیری آماری است. یکی از موارد مهم در یادگیری، قابلیت تعمیم مناسب با استفاده از تعداد محدود داده آموزشی است. در بیشتر بیماری‌ها داده‌های آموزشی کافی برای الگوریتم یادگیرنده وجود ندارد و یا فراهم نمودن آن زمان‌بر و مشکل است. فرآیند یادگیری برای پیش‌بینی بیماری‌ها با تعداد محدود داده آموزشی از طریق ماشین‌های بردار پشتیبان امکان‌پذیر است [۱۵].

در روش SVM برای کلاس‌بندی، ابتدا داده‌ها از فضای اولیه Rn با استفاده از تبدیل غیرخطی $\emptyset$، به فضای Rm که m>n است، منتقل شد. برای یافتن مرز تصمیم‌گیری بهینه از فرمول ۱ استفاده گردید.

$$max_{a_i,\dots,a_N}\left[\frac{-1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}a_i a_j y_i y_j\left(\emptyset(x_i).\emptyset(x_j)\right)+\sum_{i=1}^{N}a_i\right]$$
$$0 \le a_i \le C \quad i=1,\dots,N$$
$$0 = {}_i y_i a_{1=i}^{N} \quad \sum \quad \sum_{i=1}^{N} a_i y_i = 0 \qquad (۱)$$

که x نقطه‌ای روی مرز تصمیم‌گیری، a_i ضریب لاگرانژ و C عدد ثابت است. در رابطه ۱ از یک تابع هسته به جای $\emptyset$ به صورت رابطه ۲ استفاده شد.

$$k(x_i, x_j) = \emptyset(x_i)\emptyset(x_j) \qquad (۲)$$

با تعیین مقدار $k(x_i, x_j)$ مسئله بهینه‌سازی حل می‌شود. می‌توان تابع سیگموئید را به عنوان تابع هسته به صورت رابطه ۳ در نظر گرفت:

$$k(x_i, x_j) = e^{(-\gamma\|x-x_i\|^2)} \qquad (۳)$$

انتخاب مناسب پارامترهای C و γ در الگوریتم ماشین بردار پشتیبان دارای اهمیت است. پارامتر C میزان جریمه را نشان داده و در خروجی دسته‌بند تأثیرگذار است [۱۶].

الگوریتم GBC: الگوریتم GBC رویکرد جدیدی است که توان افزایش دقت و کاهش خطا در پیش‌بینی را با محاسبات

[Downloaded from jhbmi.com on 2026-11-17]

حداقل و مرتبه زمانی کمتر از کارهای مشابه دارد [۱۷]. روال بهینه‌یابی در الگوریتم GBC براساس یک روند تصادفی- هدایت شده استوار است. این روش، بر مبنای نظریه تکامل تدریجی و ایده‌های الگوریتم زنبورعسل پایه گذاری شده است. الگوریتم زنبور عسل هر نقطه را در فضای پارامتری متشکل از پاسخ‌های ممکن به عنوان منبع غذا تحت بررسی قرار می‌دهد. زنبورهای دیده‌بان–کارگزاران شبیه‌سازی شده – به صورت تصادفی فضای پاسخ‌ها را ساده می‌کنند و به وسیله تابع شایستگی، کیفیت موقعیت‌های بازدید شده را گزارش می‌دهند. جواب‌های ساده شده رتبه‌بندی می‌شوند و دیگر زنبورها نیروهای تازه‌ای هستند که فضای پاسخ‌ها را در پیرامون خود برای یافتن بالاترین رتبه محل‌ها جستجو می‌کنند که گلزار نامیده می‌شود. الگوریتم به صورت گزینشی دیگر گلزارها را برای یافتن نقطه بیشینه تابع شایستگی جستجو می‌کند.

در نهایت، به منظور دستیابی به یک تعادل بین بهره‌برداری و اکتشاف در الگوریتم ABC و بهبود توانایی‌های جستجوی محلی، از اپراتورهای جهش الگوریتم ژنتیک در طول فرآیند جایگزینی راه‌حل‌ها استفاده شد. الگوریتم GBC با دادن جمعیت تصادفی اولیه از فضای جستجو یعنی ویژگی‌های استخراج شده از تصاویر ماموگرافی شروع شد که به صورت رابطه ۴ است:

$$X_{mi} = X_i^{min} + rand * (X_i^{max} - X_i^{min} \quad (۴)$$

Xmi یک بردار راه‌حل (ویژگی‌های انتخاب شده از تصاویر ماموگرافی) برای مسئله بهینه‌سازی و i=1,…,n و m=1,…,SN می‌باشد. SN بیانگر تعداد جمعیت اولیه و هر Xi یک بردار n بعدی است. سپس تابع برازش هر راه‌حل به صورت رابطه ۵ محاسبه شد.

$$\vec{P}_m = \frac{fit_m(\vec{x}_m)}{\sum_{m=1}^{Sn} fit_m(\vec{x}_m)} \quad (۵)$$

fit(x_m) تابع برازش x_m می‌باشد.

مدل پیشنهادی: بیماری‌های قلبی و عروقی در رأس علت مرگ‌ومیر قرار می‌گیرند. شناخت عوامل خطر یا ریسک فاکتورهای ماژور بیماری قلبی و عروقی و لزوم پیشگیری و کنترل این عوامل خطر و به حداقل رساندن اثرات بد آن‌ها دارای اهمیت است. وجود هر کدام از ریسک فاکتورهای ماژور می‌تواند همراه با یا بدون یک زمینه فامیلی و خانوادگی و نوع رژیم غذایی و شیوه زندگی، ریسک ابتلای بیماری قلبی و عروقی را چندین برابر افزایش دهد یا سبب تشدید یا تسریع بیماری قلبی شود [۱۸]. شباهت در علائم ریسک فاکتورهای ماژور، احتمال بروز خطا در پیش‌بینی این بیماری توسط پزشکان را افزایش می‌دهد. استفاده از الگوریتم‌های داده‌کاوی مانند ماشین بردار پشتیبان باعث کاهش خطا در پیش‌بینی بیماری و افزایش شانس درمان موفقیت‌آمیز بیمار می‌شود. در ادامه، مراحل مدل پیشنهادی در تشخیص ارتباط بین ویژگی ابتلا به سرطان و آنفارکتوس قلبی معرفی شد.

پیش‌پردازش: در روش پیشنهادی برای مشخص کردن ارزش و نقش هر یک از ویژگی‌ها در تشخیص بیماری به طور تصادفی برای هر یک از ویژگی‌ها وزن صفر یا یک اختصاص داده شد که نشان دهنده انتخاب یا عدم انتخاب ویژگی است. مقادیر وزن‌دار ویژگی‌ها به عنوان ورودی به الگوریتم‌های کلاسه‌بندی داده شده و توسط الگوریتم GBC بهینه گردید. هر راه‌حل در جمعیت اولیه در الگوریتم GBC مجموعه‌ای از مقادیر تصادفی (صفر یا یک) است که بیانگر وزن هریک از ویژگی بود؛ بنابراین اگر وزن نهایی ویژگی ابتلا به سرطان صفر شود، آنگاه این ویژگی با آنفارکتوس قلبی رابطه معنادار ندارد. همچنین وزن یک در ویژگی ابتلا به سرطان نشان دهنده وجود ارتباط معنادار با آنفارکتوس قلبی است.

به‌کارگیری الگوریتم GBC: الگوریتم GBC ابتدا یک راه‌حل به صورت تصادفی یا جمعیت اولیه از اندازه SN که در آن SN نشان دهنده اندازه جمعیت و یا تعداد کل منابع غذایی است، تولید می‌کند. اگر i یک راه‌حل خاص فرض شود (i = 1, 2,..., SN) هر راه حل، یک بردار چندبعدی D به وجود می‌آورد (j = 1, 2, 3,... D) که در آن D نشان دهنده تعداد ویژگی‌های آموزنده در هر راه‌حل بهینه شده است. هر سلول X_{ij} نشان دهنده شاخص وزن مربوط به ویژگی‌ها است. در مسئله انتخاب ویژگی، هر راه‌حل همراه با دقت طبقه‌بندی که با استفاده از الگوریتم‌های دسته‌بندی کننده مانند شبکه عصبی و SVM به دست می‌آید، مورد پردازش قرار می‌گیرد. مقدار شهد در یک منبع غذایی به مقدار دقت از راه‌حل وابسته در الگوریتم GBC مرتبط است. اگر مقدار دقت جدید بهتر از ارزش دقتی باشد که قبلاً به دست آمده است، پس از آن زنبور راه‌حل قدیمی را رها می‌کند (منبع غذایی) و به سوی یکی از راه‌حل‌های جدید حرکت می‌کند. در غیر آن صورت آن راه‌حل‌های قدیمی را حفظ می‌کند.

[Downloaded from jhbmi.com on 2026-11-17]

بعد از مقداردهی اولیه از راه‌حل‌های تصادفی (جمعیت)، الگوریتم GBC شروع به جستجو برای راه حل مطلوب می‌کند.

مرحله زنبورعسل کارگر: در این مرحله، زنبور عسل کارگر به راه‌حل‌های کاندید شده (منابع غذایی) ارسال گردیده و برازندگی (مقدار شهد) آن‌ها با استفاده از دقت طبقه‌بندی ارزیابی شد؛ بنابراین زنبورعسل کارگر در سراسر فضای راه‌حل‌ها جستجو شد.

مرحله زنبورعسل ناظر: در این مطالعه، عملیات همبری برای به اشتراک‌گذاری اطلاعات بین زنبورهای عسل کارگر و ناظر در فضای جستجو بهینه‌سازی (کندو) استفاده شد. زنبورهای ناظر با تماشای رقص زنبورهای عسل کارگر محل، راه‌حل (منبع غذایی) را یاد می‌گیرند. در الگوریتم پیشنهادی GBC، زنبورهای عسل ناظر از محل بهترین منبع غذایی که راه‌حلی با بالاترین مقدار برازندگی دارد (به عنوان ملکه زنبور عسل در الگوریتم مطالعه حاضر مشخص شده‌اند) استفاده می‌کنند. شایان ذکر است که در الگوریتم ABC اصلی، زنبورهای ناظر از این اطلاعات برای انتخاب همسایه استفاده نمی‌کنند. هر زنبور عسل ناظر به طور تصادفی یک زنبور کارگر را به عنوان یک همسایه انتخاب می‌کند. در الگوریتم GBC، یک مدل مبتنی بر عملیات همبری یکسان برای انتخاب همسایه زنبور ناظر ارائه شد. در مدل پیشنهادی، بهترین نسل به دست آمده از نتیجه همبری، به عنوان یک خروجی در نظر گرفته شد.

مرحله زنبور دیده‌بان: به منظور دستیابی به تعادل بین قابلیت بهره‌برداری و اکتشاف الگوریتم ABC، حرکات زنبور عسل دیده‌بان در الگوریتم ABC اصلی اصلاح شد. در الگوریتم پیشنهادی به منظور بهبود فرآیند بهره‌برداری، از اپراتورهای جهش GA در طول روند جایگزینی راه‌حل‌های قدیمی استفاده شد.

اولین زنبورعسل دیده‌بان طبق الگوریتم پایه ABC فرآیند جستجو را بازنشانی کرده و به طور تصادفی فضای جدیدی از راه‌حل را کشف می‌کند و متعاقباً، یک عملگر جهش مکانی را برای جای‌گذاری زنبورعسل دیده‌بان دوم در نظر می‌گیرد. در زنبورعسل دیده‌بان دوم، به دنبال یک محل در سراسر فضای جستجو با توجه به راه‌حلی با بالاترین برازندگی که تاکنون تولید شده است هستیم. نرخ احتمال جهش ۰/۰۱ تنظیم شد.

پیاده‌سازی روش پیشنهادی: در این مطالعه از الگوریتم GBC برای بهینه‌سازی پارامترهای روش SVM به منظور افزایش دقت پیش‌بینی آنفارکتوس قلبی استفاده شد. دو پارامتر مهم C و γ در نتیجه پیش‌بینی سکته قلبی مؤثر هستند. مقدار کوچک برای پارامتر C منجر به کاهش نرخ دقت دسته‌بندی شد. اگر مقدار C بزرگ باشد، دقت دسته‌بندی در مرحله آموزش افزایش می‌یابد، ولی کاهش دقت در مرحله آزمایش را به دنبال خواهد داشت. تأثیر پارامتر γ در خروجی دسته‌بندی نسبت به C بیش‌تر است. مقدار بزرگ برای پارامتر γ باعث برازش بیش از حد و مقدار کوچک آن، کاهش دقت را در پی دارد.

تابع هزینه در دسته بند SVM به شکل فرمول ۶ تعریف شد:

$$Cost\ (x_i) = Miss\ (x_i) \qquad (۶)$$

در رابطه فوق، x_i یک راه‌حل برای مسئله بوده و $Miss\ (x_i)$ نشان‌دهنده تعداد نقاط آموزشی می‌باشد که توسط x_i به صورت نادرست دسته‌بندی شده‌اند. $Cost\ (x_i)$ مقدار هزینه x_i است.

تابع هدف برای الگوریتم GBC به صورت فرمول ۷ تعریف شد:

$$\text{Objective Function} = \text{Minimize (Error Rate)} \qquad (۷)$$

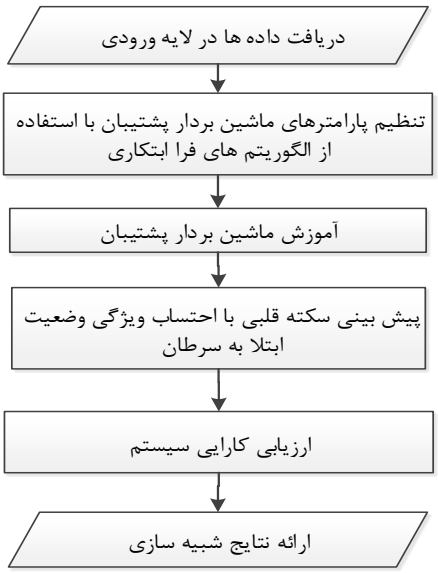
روش کار و نمودار جریان کاری الگوریتم GBC برای محاسبه C و γ بهینه به ترتیب در جدول ۲ و شکل ۱ نشان داده شد.

داده‌های موجود در پرونده بیماران با استفاده از نرم‌افزار Matlab توصیف، شبیه‌سازی و تحلیل شد.

جدول ۲: روش کار الگوریتم GBC برای محاسبه C و γ بهینه

۱- شروع
۲- مقداردهی اولیه به متغیرها
۳- برای هر دوره شبیه‌سازی
۳-۱- محاسبه‌ی تابع هدف
۳-۲- مقایسه با بهترین تابع‌های هدف قبلی و یافتن تابع هدف بزرگتر.
۳-۳- یافتن C و γ مطلوبی که منجر به بیشینه شدن تابع هدف شود.
۳-۴- به روز رسانی پارامترها
۴- بررسی شرط پایان تکرار الگوریتم
۵- پایان

[Downloaded from jhbmi.com on 2026-11-17]



شکل ۱: روندنمای رهیافت پیشنهادی

عمل طبقه‌بندی با مقادیر ویژگی‌های انتخاب شده که بر اساس هر یک از زنبورها محاسبه گردید، به روش ماشین بردار پشتیبان انجام شـــد. هر زنبور که کلاســـه‌بندی را با خطای پیش‌بینی کـم‌تری انـجام دهـد، شـایسته‌تر خـواهد بود. تخمـین خطا بر مبنای نمونه‌گیری به روش ۱۰-Fold Cross Validation [۱۹] انجام شد. تابع برازندگی نشان دهنده شایستگی یا توانایی هر کروموزوم اســـت. الگوریتم ۱ نمایش دهنده تابع برازندگی روش پیشنهادی است:

الگوریتم ۱: تابع برازندگی روش پیشنهادی

1) Read the data from a file
2) Normalize the attribute values in the range of 0 to 1.
3) Let $x_1, x_2 \ldots x_m$ denote the m instances from data set $\{f_1, f_2, \ldots, f_n\}$, n = Number of features
4) Assign weight w_i to each instance x_i in the training set
5) Predict the class value from weighted dataset by classifier.
6) Calculate the error rate as
Error Rate = 1- (# of correctly classified examples / All) * 100
8) Cost Function=Minimize (Error Rate), Objective function=maximize(Accuracy)

نتایج

در فرآیند انتخاب ویژگی توسط مدل پیشنهادی، ابتلا به سرطان به عنوان ویژگی تأثیرگذار در احتمال بروز سکته قلبی شناسایی شد. همچنین ارتباط بین ابتلا به سرطان و آنفارکتوس قلبی در بیماران مستعد عارضه قلبی، معنادار است. مدل پیشنهادی با سه روش نایو بیز [۱۹]، درخت‌تصمیم [۲۰] و نزدیک‌ترین همسایه [۲۱] مقایسه شد. ارتباط بین کلاس‌های واقعی و کلاس‌های پیش‌بینی شده با استفاده از ماتریس Confusion قابل محاسبه است. پارامترهای موردنیاز ماتریس Confusion (پارامترهای مورد نیاز برای ارتباط بین کلاس‌های واقعی و کلاس‌های پیش‌بینی شده) به صورت زیر ذکر شد:

برای مقایسه مدل پیشنهادی با سایر روش‌ها از معیارهای Accuracy، Sensitivity، Specificity، Precision و F-Measure با توجه به شکل ۵ طبق روابط زیر استفاده شد [۲۲].

[Downloaded from jhbmi.com on 2026-11-17]

$$Accuracy = (TP + TN) / All \quad (۸)$$

$$Sensitivity = TP / (TP + FN) \quad (۹)$$

$$Specificity = TN / (FP + TN) \quad (۱۰)$$

$$Precision = TP / (TP + FP) \quad (۱۱)$$

$$Recall = TP / (TP + FN) \quad (۱۲)$$

$$F\ Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (۱۳)$$

TP: تعداد رکوردهایی که به درستی، مثبت تشخیص داده شوند.

TN: تعداد رکوردهایی که به درستی، منفی تشخیص داده شوند.

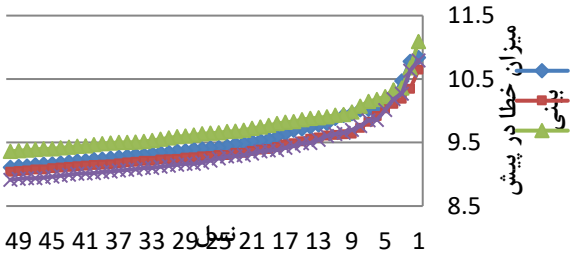
FP: تعداد رکوردهایی که به غلط، مثبت تشخیص داده شوند.

FN: تعداد رکوردهایی که به غلط، منفی تشخیص داده شوند.

در شکل ۲ مقایسه میزان خطا برای پیش‌بینی بیماری در الگوریتم پیشنهادی با روش‌های ژنتیک، بهینه‌سازی اجتماع ذرات و زنبور عسل مصنوعی در ۵۰ نسل نمایش داده شد.

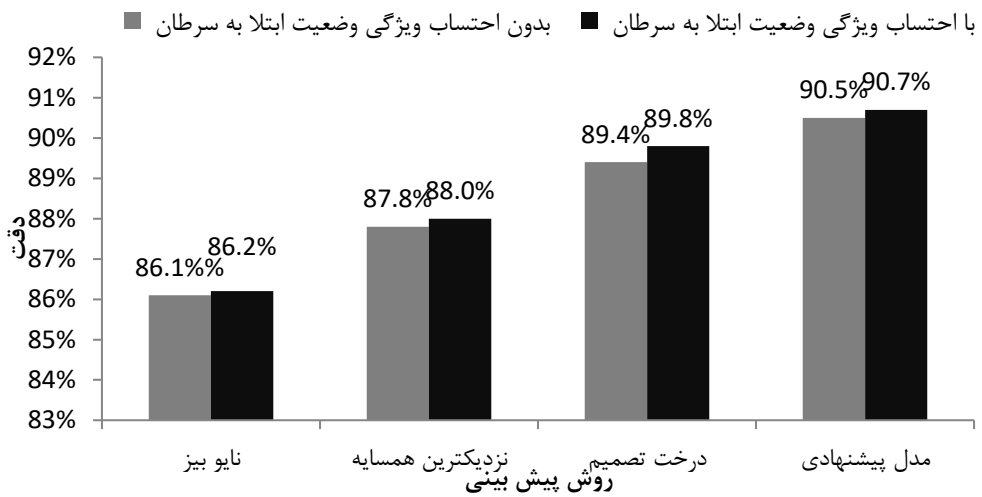
همان‌طور که مشاهده شد، دقت پیش‌بینی آنفارکتوس قلبی در هر مرحله افزایش یافت.

نتایج شبیه‌سازی، یک بار با به‌کارگیری ویژگی وضعیت ابتلا به سرطان و یک بار بدون احتساب این ویژگی نشان داده شد تا ارتباط بین سرطان و سکته قلبی مشخص شود. شکل ۳ نشان‌دهنده نمودار نتایج تشخیص سکته قلبی روش‌های مختلف با معیار Accuracy است. همان‌گونه که مشاهده شد مدل پیشنهادی دقت بیش‌تری نسبت به سایر روش‌ها دارد. هم‌چنین دقت نتایج با به‌کارگیری ویژگی وضعیت ابتلا افزایش یافت. مقایسه نتایج پیش‌بینی سکته قلبی یک بار با به‌کارگیری ویژگی وضعیت ابتلا به سرطان و یک بار بدون احتساب این ویژگی با معیارهای Sensitivity و Specificity به ترتیب در جدول‌های ۲ و ۳ نمایش داده شد.



شکل ۲: درصد خطا برای پیش‌بینی سکته قلبی در الگوریتم پیشنهادی در ۵۰ نسل

[Downloaded from jhbmi.com on 2026-11-17]



شکل ۳: نمودار نتایج پیش‌بینی سکته قلبی با معیار Accuracy یک بار با به‌کارگیری

جدول ۲: مقایسه نتایج نتایج پیش‌بینی سکته قلبی با معیار Sensitivity

		نایو بیز	نزدیک‌ترین همسایه	درخت تصمیم	مدل پیشنهادی
بدون احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۴۰	۰/۴۱	۰/۵۴	۰/۶۲
	عدم بروز سکته قلبی	۰/۹۸	۰/۹۷	۰/۹۷	۰/۹۸
با احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۴۱	۰/۴۳	۰/۵۶	۰/۶۳
	عدم بروز سکته قلبی	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹

جدول ۳: مقایسه نتایج پیش‌بینی سکته قلبی با معیار Specificity بدون احتساب ویژگی وضعیت ابتلا به سرطان

		نایو بیز	نزدیک‌ترین همسایه	درخت تصمیم	مدل پیشنهادی
بدون احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۹۸	۰/۹۷	۰/۹۷	۰/۹۸
	عدم بروز سکته قلبی	۰/۴۰	۰/۴۱	۰/۵۴	۰/۶۲
با احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹
	عدم بروز سکته قلبی	۰/۴۱	۰/۴۳	۰/۵۶	۰/۶۳

جدول ۴ و ۵ نتایج روش‌ها به ترتیب با معیار Precision و F-Measure مورد مقایسه قرار گرفت. نتایج مقایسه نشان دهنده برتری عملکرد مدل پیشنهادی است. در نهایت دقت نتایج با به‌کارگیری ویژگی وضعیت ابتلا افزایش پیدا نمود.

جدول ۴: مقایسه نتایج پیش‌بینی سکته قلبی با معیار Precision بدون احتساب ویژگی وضعیت ابتلا به سرطان

		نایو بیز	نزدیک‌ترین همسایه	درخت تصمیم	مدل پیشنهادی
بدون احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۹۳	۰/۹۳	۰/۹۳	۰/۹۴
	عدم بروز سکته قلبی	۰/۸۵	۰/۸۴	۰/۸۷	۰/۹۰
با احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۹۴	۰/۹۵	۰/۹۵	۰/۹۵
	عدم بروز سکته قلبی	۰/۸۶	۰/۸۶	۰/۸۹	۰/۹۰

[Downloaded from jhbmi.com on 2026-11-17]

جدول ۵: مقایسه نتایج پیش‌بینی سکته قلبی با معیار F-Measure بدون احتساب ویژگی وضعیت ابتلا به سرطان

		نایو بیز	نزدیک‌ترین همسایه	درخت تصمیم	مدل پیشنهادی
بدون احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۵۷	۰/۵۷	۰/۶۸	۰/۷۵
	عدم بروز سکته قلبی	۰/۹۲	۰/۹۲	۰/۹۲	۰/۹۴
با احتساب ویژگی وضعیت ابتلا به سرطان	بروز سکته قلبی	۰/۵۶	۰/۵۹	۰/۷۰	۰/۷۶
	عدم بروز سکته قلبی	۰/۹۱	۰/۹۲	۰/۹۴	۰/۹۵

بحث و نتیجه‌گیری

در این مطالعه، مجموعه‌ای از الگوهای کمتر شناخته شده و مؤثر در بروز سکته قلبی بر اساس یک پایگاه داده بومی مورد پردازش قرار گرفت. طراحی و تکمیل پرسش‌نامه‌ها، ثبت علائم بالینی و نتایج آزمایشگاهی در مجموعه داده‌های مورد استفاده، توسط نویسندگان این مقاله در بیمارستان فوق تخصصی شهید صدوقی یزد جمع‌آوری شد. پژوهش‌های مربوط به پیش‌بینی آنفارکتوس قلبی، محدودیت‌هایی از قبیل تعداد کم بیماران برای ایجاد مدل، داده‌های از دست رفته و متغیرهای ناقص را دارا می‌باشند. در این پژوهش تعداد قابل قبولی از بیماران با متغیرهای مناسب با حداقل داده‌های ازدست رفته به کار گرفته شد. مطالعه حاضر با به‌کارگیری ویژگی‌های بالینی و آزمایشگاهی بیماران با استفاده الگوریتم GBC و ماشین بردار پشتیبان، آنفارکتوس قلبی را پیش‌بینی نمود. مدل پیشنهادی این مطالعه شامل نرمال‌سازی ویژگی‌ها، انتخاب ویژگی‌ها، وزن‌دهی به ویژگی‌ها و معرفی روشی برای پیش‌بینی آنفارکتوس قلبی با استفاده از ویژگی‌های وزن‌دار است. در این مقاله، مجموعه داده‌ها یک بار با به‌کارگیری ویژگی وضعیت ابتلا به سرطان و یک بار بدون احتساب این ویژگی شبیه‌سازی شد تا ارتباط بین سرطان و سکته قلبی مشخص شود. نتایج شبیه‌سازی حاکی از احتمال بروز سکته قلبی، در بیماران مستعد عارضه قلبی، پس از خبردار شدن از بیماری سرطان طی ماه‌های اولیه خود است.

سابقه فشارخون بالا، ویژگی‌های اختلالات سندرم متابولیک، BMI(Body Mass Index) بالاترین کارایی در دسته‌بندی مدل پیشنهادی دارا بودند. همچنین ویژگی قرار گرفتن در معرض دود دخانیات، حداقل تأثیر در دسته‌بندی را داشت. متغیر تحصیلات هیچ تأثیری در تشخیص و پیش‌بینی آنفارکتوس قلبی نداشت. در این مطالعه علاوه بر شناسایی مهم‌ترین ویژگی‌ها، با استفاده از انتخاب ویژگی‌ها بر اساس الگوریتم ژنتیک به عملکردی بهتر بر حسب شاخص‌های دقت، حساسیت و ویژگی دست یافته شد.

Lavanya و Rani [۲۱] با استفاده از داده‌های پایگاه داده (Wisconsin Breast Cancer Dataset) WBCD و درخت تصمیم و دسته‌بندی دو مرحله‌ای به دقت ۹۴٪/۸۷ رسیدند. Kiyan و Yildirim [۲۲] با استفاده از روش‌های (Radial Basis Function) و MLP RBF(Multi-Layer Perceptron) برای پیش‌بینی سرطان پستان به دقت ۹۶/۱۷٪ و ۹۵/۷۴٪ رسیدند. Land و Verheggen [۲۳] برای پیش‌بینی مجموعه داده‌های سرطان پستان از ماشین بردار استفاده کرده‌اند. Soltani-Sarvestani و همکاران [۲۴] با مقایسه میانگین مربع خطا در شبکه‌های عصبی چندلایه، رقابتی و پایه شعاع برای پیش‌بینی درجه بدخیمی سرطان پستان پرداختند که بهترین دقت را شبکه عصبی پایه شعاعی داشت. Chaurasia و Chakrabarti و همکاران [۲۵] به دقت ۹۶٪/۴۶ با استفاده از SVM رسیدند. طلوعی اشلقی و همکاران [۲۶] پیش‌بینی عود مجدد سرطان پستان را به کمک سه تکنیک داده کاوی ارائه دادند. که دقت در سه الگوریتم داده‌کاوی، Artificial Neural Networks(ANN)، درخت تصمیم‌گیری و SVM به ترتیب برابر ۰/۹۳۶، ۰/۹۴۷ و ۰/۹۵۷ بوده است.

همان گونه که در شکل ۷ مشاهده شد، این پژوهش بر روی مجموعه داده بیماران بومی مبتلا به سکته قلبی بیمارستان شهید صدوقی یزد به‌دقت ۹۰/۸ ٪ رسید.

روش پیشنهادی با به‌کارگیری الگوریتم GBC برای افزایش دقت SVM با هدف طراحی و ارزیابی یک مدل پزشک‌یار در پیش‌بینی سکته قلبی ارائه شد. مدل پزشک‌یار طراحی شده در این پژوهش در تشخیص سکته قلبی موفق بوده است و طبقه‌بندی با دقت قابل قبولی انجام گرفت. آزمایش‌ها و شبیه‌سازی نشان داد سیستم پزشک‌یار معرفی شده

[Downloaded from jhbmi.com on 2026-11-17]

در این پژوهش بر روی مجموعه داده بیماران بومی مبتلا به سکته قلبی بیمارستان شهید صدوقی یزد به‌دقت ۹۰/۸٪ رسید. با استفاده از الگوریتم‌های انتخاب ویژگی مبتنی بر مدل پیشنهادی، امکان تأثیرگذاری ابتلا به سرطان در بروز آنفارکتوس قلبی در بیماران مستعد عارضه قلبی، مشخص شد.

جستجو در پایگاه داده‌های پزشکی برای رسیدن به دانش و اطلاعات جهت پیش‌بینی، تشخیص و تصمیم‌گیری از کاربردهای داده‌کاوی در پزشکی است. می‌توان از الگوریتم‌های وراثتی مانند الگوریتم GBC برای بهینه‌سازی تکنیک‌های داده‌کاوی استفاده کرد. پیش‌بینی و تشخیص صحیح آنفارکتوس قلبی با استفاده از هوش مصنوعی و یادگیری ماشین، احتمال درمان موفق را بالا می‌برد. مقایسه نتایج شبیه‌سازی نشان می‌دهد که مدل پیشنهادی از روش‌های نایو بیز، درخت تصمیم و نزدیک‌ترین همسایه دقت بیش‌تری دارد.

جمع‌آوری مجموعه داده‌ها به صورت دستورالعمل‌های آماری از مناطق مختلف، پیشنهادی است که برای بهبود این روش مطرح گردید. کمبود تعداد نمونه‌ها برای شبیه‌سازی مهم‌ترین محدودیت روش پیشنهادی می‌باشد که نیاز است تدابیری اندیشه شود تا این محدودیت رفع شده و نتایج دقیق‌تر شود.

تعارض منافع

نویسندگان بر خود لازم می دانند از مدیریت محترم بیمارستان فوق تخصصی مرتاض یزد به خاطر تأمین مجموعه داده‌های موردنیاز جهت انجام این پژوهش تشکر نمایند.

References

1. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. Cancer 2018;124(8):1682-90.

2. Klaus A, Fathi O, Tatjana TW, Bruno N, Oskar K. Expression of hypoxia-associated protein hif-1alpha in follicular thyroid cancer is associated with distant metastasis. Pathol Oncol Res 2018;24(2):289-96.

3. Bordoloi D, Ajaikumar AB. The potential of curcumin: a multitargeting agent in cancer cell chemosensitization. Role of Nutraceuticals in Cancer Chemosensitization 2018; 38: 31-60.

4. Bronson MR, Kapadia NS, Austin AM, Wang Q, Feskanich D, Bynum JPW, et al. Leveraging Linkage of Cohort Studies With Administrative Claims Data to Identify Individuals With Cancer. Med Care 2018;56(12):e83-e9.

5. Xiao R, Ward MC, Yang K, Adelstein DJ, Koyfman SA, Prendes BL, et al. Increased pathologic upstaging with rising time to treatment initiation for head and neck cancer: A mechanism for increased mortality. Cancer 2018;124(7):1400-14.

6. Romero I, Garrido C, Algarra I, Chamorro V, Collado A, Garrido F, et al. MHC Intratumoral heterogeneity may predict cancer progression and response to immunotherapy. Front Immunol 2018;9:102.

7. Wang J, Wu CJ, Bao ML, Zhang J, Shi HB, Zhang YD. Using support vector machine analysis to assess PartinMR: A new prediction model for organ-confined prostate cancer Journal of Magnetic Resonance Imaging 2018;48(2):499–506.

8. Zayeri F, Sadeghi Nejad R, Noorkojuri H, Bagheri J, Ghazanfari E. Application of classification tree model for determining the effective factors of mortality after coronary bypass surgery in dialysis-independent patients. Daneshvar Medicine 2012; 19(98):15-24. Persian

9. Sajjadi F, Mohammadifard N, Ghaderian N, Alikhasi H, Maghroon M. Clustering of cardiovascular risk factors in diabetic and glucose intolerant cases. J Qazvin Univ Med Sci 2005; 9(2):35-43. Persian

10. Parthiban G, Rajesh A, Srivatsa SK. Diagnosis of heart disease for diabetic patients using naive bayes method. International Journal of Computer Applications 2011; 24(3):7-11.

11. Vidic I, Egnell L, Jerome NP, Teruel JR, Sjobakk TE, Ostlie A, et al. Support vector machine for breast cancer classification using diffusion-weighted MRI histogram features: Preliminary study. J Magn Reson Imaging. 2018;47(5):1205-16.

12. Alshamlan HM, Badr GH, Alohali YA. Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification. Comput Biol Chem 2015;56:49-60.

13. Jerline Amutha A, Padmajavalli R, Prabhakar D. A novel approach for the prediction of treadmill test in cardiology using data mining algorithms implemented as a mobile application. Indian Heart Journal 2018;70(4):511-8.

14. Kaur A, Arora J. Heart disease prediction using data mining techniques: a survey. International Journal of Advanced Research in Computer Science 2018; 9(2): 569-72.

15. Bloodgood M. Support Vector Machine Active Learning Algorithms with Query-by-Committee versus Closest-to-Hyperplane Selection. Proceedings of the IEEE 12th International Conference on Semantic Computing; 2018 Jan; Laguna Hills, CA, USA: IEEE; 2018. p. 148-55.

16. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 2000; 16(10): 906-14.

[Downloaded from jhbmi.com on 2026-11-17]

17. Gong D, Han Y, Sun J. A novel hybrid multi-objective artificial bee colony algorithm for blocking lot-streaming flow shop scheduling problems. Knowledge-Based Systems 2018;148:115-30.
18. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97(18):1837-47.
19. Gao CZ, Cheng Q, He P, Susilo W, Li J. Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack. Information Sciences 2018;444:72-88.
20. Satapathy SC, Bhateja V, Das S. Smart Computing and Informatics. Sehgal R, Mehrotra D, Bala M, editors. A Decision Tree Approach to Identify the Factors Affecting Reliability for Component-Based System. Singapore: Springer; 2018. p. 237-43.
21. Lavanya D, Rani KU. Ensemble decision tree classifier for breast cancer data. International Journal of Information Technology Convergence and Services 2012; 2(1): 17-24.
22. Kiyan T, Yildirim T. Breast cancer diagnosis using statistical neural networks. IU-Journal of Electrical & Electronics Engineering 2004;4(2):1149-53.
23. Land WH, Verheggen EA. Multiclass primal support vector machines for breast density classification. Int J Comput Biol Drug Des 2009;2(1):21-57.
24. Soltani-Sarvestani A, Safavi AA, Parandeh MN, Salehi M. Predicting breast cancer survivability using data mining techniques. 2nd International Conference on Software Technology and Engineering; 2010 Oct 3-5; San Juan, PR, USA: IEEE; 2010.p. 227.
25. Chaurasia S, Chakrabarti P. An approach with Support Vector Machine using variable features selection on breast cancer prognosis. International Journal of Advanced Research in Artificial Intelligence 2013;2(9):38-42.
26. Toloee Ashlaghi A, Pour Ebrahimi A, Ebrahimi M, Ghasem Ahmad L. Using data mining techniques for prediction breast cancer recurrence. Iranian Quarterly Journal of Breast Diseases 2013;5(4):23-34. Persian

[Downloaded from jhbmi.com on 2026-11-17]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2018; 5(3): 361-372

Using GBC Algorithm to Optimize Support Vector Machine Parameters for Predicting the Relationship between Cancer and Cardiac Infarction: A Case Study

Nooshyar Mehdi[1]*, Mo'meni Mohammad[2], Gharravi Sorayya[3], Hourali Fatemeh[4]

• **Received:** 18 May, 2018 • **Accepted:** 9 Oct, 2018

Introduction: Awareness of cancer increases the probability of neurotic disorders and stress in the patient. Also, stress increases the risk of myocardial infarction. The present study aimed to determine the probability of a heart attack in cancer patients based on the GBC algorithm.
Method: In this study, data were collected from the database of Shahid Sadoughi subspecialty hospital in Yazd. The medical records of 1679 patients with heart attack were studied, of which 81 ones belonged to patients with cancer. In the process of selecting features by the proposed model, if cancer is identified as an effective feature, then the relationship between cancer and cardiac infarction will be meaningful.
Results: Using the proposed model, the cancer feature was selected to predict the probability of heart attack, which indicated a significant relationship between these two characteristics in patients who were vulnerable to cardiac disease. The predictive accuracy of the proposed model was 0.91
Conclusion: By choosing the cancer feature, the proposed model compared to other models has the least error rate and the most accuracy in predicting myocardial infarction. Naive bias method has maximum error rate and minimum accuracy. The simulation results indicate that in patients who are vulnerable to cardiac disease, after being diagnosed with cancer during the early months, heart attack is possible.

Keywords: Cancer, Heart attack, GBC algorithm, Support Vector Machine, Increase precision of prediction

• **Citation:** Nooshyar M, Mo'meni M, Gharravi S, Hourali F. Using GBC Algorithm to Optimize Support Vector Machine Parameters for Predicting the Relationship between Cancer and Cardiac Infarction: A Case Study. Journal of Health and Biomedical Informatics 2018; 5(3): 361-372.

1. Ph.D., in Electrical Engineering, Associate Professor, Computer Dept., School of Electrical and Computer Engineering, University of Mohaghegh Ardabili, Ardabil, Iran
2. PhD Student in Computer Engineering, Computer Dept., School of Electrical and Computer Engineering, Yazd University, Yazd, Iran
3. M.S.c, in Computer Engineering, Lecturer, Electrical and Computer Engineering Dept., Esfarayen University of Technology, North Khorasan, Esfarayen, Iran
4. M.S.c, in Computer Engineering, Lecturer, Electrical and Computer Engineering Dept., Esfarayen University of Technology, North Khorasan, Esfarayen, Iran.

***Correspondence**: End of Daneshgah Street, Mohaghegh Ardebil University, Ardebil
• **Tel:** 04531505000 • **Email:** nooshyar@uma.ac.ir

[Downloaded from jhbmi.com on 2026-11-17]