ORIGINAL ARTICLE

Iranian Quarterly Journal of Breast Disease 2018; 11(3):71.

Improvement of Genetic Bee Algorithm to Select the Effective Features in Predicting Breast Cancer from Among Dietary Habits, Cultural Factors, Clinical Signs, and Laboratory Results

Nooshyar M: Department of Computer, University of Mohaghegh Ardabili, Ardebil, Iran
Momeni M: Ph.D. Student, Computer Department, Yazd University, Yazd, Iran
Gharravi S: Electrical and Computer Department, Esfarayne Higher Education Complex, Northern Khorasan, Esfarayne, Iran

Corresponding Author: Mehdi Nooshyar, nooshyar@uma.ac.ir

[Downloaded from ijbd.ir on 2026-11-13]

[DOR: 20.1001.1.17359406.1397.11.3.2.2]

[DOI: 10.30699/acadpub.ijbd..11.3.71]

Abstract

Introduction: The identification of factors influencing the incidence of breast cancer bears great importance. The wide range of symptoms of the disease makes the diagnosis difficult for doctors. Preventing breast cancer could be achieved via a knowledge of the factors affecting the incidence of the disease. The purpose of this paper was to identify variables related to dietary habits, cultural factors, and laboratory results that could contribute to the effective prediction of breast cancer. For this purpose, an optimal model based on genetic bee colony (GBC) algorithm was developed to increase machine learning accuracy.

Methods: In this study, patient information was collected from the database of Mortaaz subspecialty hospital in Yazd. Medical records of 711 breast cancer patients were screened for 63 variables. Patients had been followed up for at least two years. Variables most affecting the incidence of breast cancer were identified using the GBC algorithm and backup vector machine.

Results: Among diet- and culture-related factors, smoking and hookah use, physical inactivity, nighttime employment, and being single and cultural factors related to breast cancer were selected. Also, based on the proposed model, factors such as comparing the results of previous and current mammograms, the duration of taking contraceptive pills, hysterectomy, hormone replacement therapy, nipple retraction and pain, type of discharge, and detecting a mass in mammograms were effective in detecting the disease. No significant relationship was found between breast cancer and factors such as using the microwave in the kitchen, or the type of rice or oil consumed.

Conclusion: Using dietary habits and cultural factors in the prediction of breast cancer, the proposed model has the least error rate and the most accuracy compared with other models such as fuzzy and neural networks.

Keywords: Genetic Bee Colony Algorithm, Feature Selection, Breast Cancer, Dietary Habits, Cultural Factors, Laboratory Results

مقاله پژوهشی

فصلنامهٔ بیماری‌های پستان ایران، سال یازدهم، شماره سوم، پاییز ۱۳۹۷؛ (۸۲-۷۲)

تاریخ ارسال: ۹۷/۰۳/۲۳ | تاریخ پذیرش: ۹۷/۰۷/۰۳

بهبود الگوریتم زنبور عسل ژنتیکی برای انتخاب ویژگی‌های موثر در پیش‌بینی سرطان پستان از بین عادات غذایی، عوامل فرهنگی، علایم بالینی و نتایج آزمایشگاهی

مهدی نوشیار [*]: گروه کامپیوتر، دانشگاه محقق اردبیلی، اردبیل، ایران

محمد مومنی: دانشجوی دکتری، گروه کامپیوتر، دانشگاه یزد، یزد، ایران

ثریا غراوی: گروه برق و کامپیوتر، مجتمع آموزش عالی اسفراین، خراسان شمالی، اسفراین، ایران

[Downloaded from ijbd.ir on 2026-11-13]

[DOR: 20.1001.1.17359406.1397.11.3.2.2]

[DOI: 10.30699/acadpub.ijbd..11.3.71]

چکیده

مقدمه: کشف ویژگی‌های موثر در بروز سرطان پستان دارای اهمیت است. وجود علایم مختلف این بیماری، تشخیص را برای پزشکان دشوار می‌کند. پیشگیری از ابتلا به سرطان پستان با آگاهی از عوامل تاثیرگذار در بروز بیماری، میسر می‌گردد. هدف این مقاله، انتخاب ویژگی‌های موثر در پیش‌بینی سرطان پستان از بین عادات غذایی، عوامل فرهنگی و نتایج آزمایشگاهی است. برای این کار یک مدل بهینه مبتنی الگوریتم زنبور عسل ژنتیکی برای افزایش دقت یادگیری ماشین معرفی می‌شود.

روش بررسی: در این مطالعه، اطلاعات بیماران از پایگاه داده بیمارستان فوق تخصصی مرتاض یزد جمع‌آوری شده است. پرونده پزشکی ۷۱۱ بیمار مبتلا به سرطان پستان با تعداد ۶۳ ویژگی مورد بررسی قرار گرفته است. هر یک از بیماران حداقل به مدت دو سال تحت پیگیری بوده‌اند. ویژگی‌های تاثیرگذار در ابتلا و تشخیص سرطان پستان از بین عادات غذایی، عوامل فرهنگی، علایم بالینی و نتایج آزمایشگاهی، با استفاده از الگوریتم GBC و ماشین بردار پشتیبان انتخاب شد.

یافته‌ها: ویژگی‌های استعمال سیگار و قلیان، عدم فعالیت ورزشی، اشتغال در شیف شب و تجرد از بین عادات غذایی و عوامل فرهنگی در ارتباط با تشخیص سرطان پستان انتخاب شدند. همچنین مدل پیشنهادی ویژگی‌هایی مانند: مقایسه نتیجه ماموگرافی قبلی و فعلی، مدت مصرف قرص ضدبارداری، هیستروکتومی، جایگزینی هورمون، میزان تو رفتگی نوک پستان، میزان درد، نوع ترشح و توده موجود در تصاویر ماموگرافی را موثر در تشخیص این بیماری دانست. هیچ ارتباط معناداری بین سرطان پستان و ویژگی‌هایی مانند استفاده از ماکروفر در آشپزخانه، نوع برنج و روغن مصرفی یافت نشد.

نتیجه‌گیری: با استفاده از عادات غذایی و عوامل فرهنگی در پیش‌بینی سرطان پستان، مدل پیشنهادی نسبت به سایر مدل‌های مورد مقایسه مانند فازی و شبکه عصبی، دارای حداقل میزان خطا و بیش‌ترین دقت و صحت است.

واژه‌های کلیدی: الگوریتم زنبور عسل ژنتیکی، انتخاب ویژگی، سرطان پستان، عادات غذایی، عوامل فرهنگی، نتایج آزمایشگاهی

[*] نشانی نویسنده مسئول: اردبیل، دانشکده مهندسی برق و کامپیوتر، گروه کامپیوتر، دانشگاه محقق اردبیلی، مهدی نوشیار.

نشانی الکترونیک: nooshyar@uma.ac.ir

مقدمه

رشد خارج از قاعده سلول‌های غیرطبیعی در پستان از علل سرطان پستان است (۱). سرطان پستان قابل پیشگیری و تشخیص زودرس است و با ارایه راه‌کارهای ویژه می‌توان مراجعه دیرهنگام بیمار را کاهش داد. درمان موثر سرطان پستان به منظور افزایش بقا، کاهش مرگ و ارتقا کیفیت زندگی دارای اهمیت است. این بیماری تهدید بزرگی بر سلامت زنان بوده و از عوامل شایع در کاهش عمر زنان به شمار می‌رود (۲).

برای پیش‌بینی، تشخیص و درمان این بیماری عوامل بسیاری از قبیل وجود تومور، درگیری غدد لنفاوی، تورفتگی نوک پستان و بروز ترشح در پستان استفاده می‌شود (۳). وجود شباهت زیاد در علایم بالینی و آزمایشگاهی سرطان پستان احتمال تشخیص نادرست را افزایش می‌دهد (۴). توده، شایع‌ترین علامت سرطان پستان می باشد که در اغلب موارد توسط خود بیمار بصورت اتفاقی کشف می شود و در بقیه موارد توسط پزشک در معاینه بالینی مشخص می شود. این توده ممکن است دردناک باشد؛ ولی در اغلب موارد بدون درد است. در بعضی موارد سرطان پستان به‌صورت توده‌های متعدد بروز می‌کند (۸ -۵).

تشخیص و پیش‌بینی انواع بیماری‌ها با استفاده از ماشین بردار پشتیبان[1] یا SVM امکان‌پذیر است. تشابه علایم بالینی و آزمایشگاهی سرطان پستان، احتمال بروز خطا در تشخیص را افزایش می‌دهد. کشف الگوهای مفید بین بیماری و علایم بالینی و آزمایشگاهی بیمار از کاربردهای ماشین بردار پشتیبان در پزشکی می‌باشد. منظور از الگوی مفید، مدلی در داده‌ها است که ارتباط میان یک زیرمجموعه از داده‌های بیمار و تشخیص بیماری را بیان می‌کند. در سال‌های اخیر با پیشرفت‌هایی که در تشخیص زودرس این بیماری به وجود آمده است، درمان آن نیز با موفقیت بیش‌تری همراه شده است. اگر توده‌های پستان در اندازه کوچک کشف شوند به خوبی قابل درمان هستند. ماشین بردار پشتیبان از روش‌های تشخیص زودرس سرطان پستان است (۹).

استر و همکاران روش تجزیه و تحلیل گسسته خطی در تشخیص سرطان پستان با استفاده از بانک اطلاعاتی WDBC با دقت ۹۶/۸٪ ارایه دادند (۱۰). شیخ‌پور و همکاران تشخیص سرطان پستان با استفاده از کاهش دو مرحله‌ای ویژگی‌های استخراج شده آسپیراسیون سوزنی و الگوریتم‌های داده‌کاوی را معرفی کردند (۱۱).

الگوریتم زنبور (Bee Algorithm) یک الگوریتم گروهی مبتنی بر جستجو است که در سال ۲۰۰۵ میلادی ابداع شده است. این الگوریتم شبیه‌سازی رفتار جستجوی غذای گروه‌های زنبور عسل است. در نسخه ابتدایی این الگوریتم، الگوریتم نوعی از جستجوی محلی انجام می‌دهد که با جستجوی تصادفی ترکیب شده و می‌تواند برای بهینه‌سازی ترکیبی یا بهینه‌سازی تابعی استفاده شود. این الگوریتم نیز مانند سایر الگوریتم‌های هوش ازدحامی از دو روش اکتشاف و استخراج استفاده می‌کند. زنبورهای کارگر وظیفه استخراج و زنبورهای ناظر وظیفه اکتشاف را به عهده دارند. زنبورهای کارگر در اطراف یک منطقه (گل‌های پیدا شده یا منطقه‌ای که شامل جواب مسئله است) به دنبال جواب بهینه می‌گردند و زنبورهای ناظر با رفتار تصادفی به دنبال پیدا کردن مناطق جدید هستند.

استفاده از سه تکنیک داده‌کاوی برای تشخیص سرطان پستان توسط دلن و همکاران معرفی شد که پیش‌بینی سرطان توسط درخت تصمیم (C5) با دقت ۹۳/۶٪، SVM مصنوعی با دقت ۹۱/۲٪ و مدل رگرسیون لجستیک با دقت ۸۹/۲٪ شبیه‌سازی گردید. آرونا و همکاران با استفاده از بانک اطلاعاتی WDBC[2] توسط مدل‌های داده‌کاوی از قبیل درخت تصمیم و بیز ساده و SVM به ترتیب با دقت ۹۲/۹۷٪، ۹۲/۶۱٪ و ۹۳/۶۷ به پیش‌بینی سرطان پستان رسیدند (۱۲).

دقت تشخیص و پیش‌بینی ماشین بردار پشتیبان را می‌توان از طریق تنظیم پارامترهای ورودی با استفاده از الگوریتم‌های هوش جمعی افزایش داد. هوش جمعی در واقع بر پایه شبیه‌سازی رفتار جمعی گروهی از حیوانات مانند زنبورها شکل گرفته‌است (۱۳). در این مطالعه با طراحی پرسش‌نامه، تکمیل و گردآوری مجموعه داده‌های موبوط به پرونده‌های بیماران در بیمارستان فوق تخصصی مرتاض یزد، مدلی برای افزایش دقت تشخیص و پیش‌بینی ماشین بردار پشتیبان مبتنی بر الگوریتم کلونی زنبور عسل معرفی می‌گردد.

[1] Support vector machines

[2] Wisconsin Diagnostic Breast Cancer

[Downloaded from ijbd.ir on 2026-11-13]

[DOR: 20.1001.1.17359406.1397.11.3.2.2]

DOI: 10.30699/acadpub.ijbd..11.3.71]

مواد و روش‌ها

انتظار مراجعه‌کنندگان در مطب ماموگرافیست جهت نوبت‌دهی برای برداشت تصاویر ماموگرافی، فرصتی را فراهم کرد تا مجموعه داده‌ها به طور کامل جمع‌آوری گردد. تکمیل بانک اطلاعاتی با اخذ رضایت‌نامه از پزشک، ماموگرافیست و مراجعه‌کنندگان انجام گرفت. پرسشنامه‌های مربوط به سرطان پستان، تحت نظر پزشکان متخصص در بیمارستان فوق تخصصی مرتاض یزد بر اساس ویژگی‌های جدید طراحی، تکمیل و جمع‌آوری گردیده است. داده‌ها مربوط به سال‌های ۱۳۹۳ تا ۱۳۹۵ است. مجموعه داده‌ها در این مقاله شامل سه دسته است:

۱. پرسشنامه‌های جمع‌آوری شده از مراجعه‌کنندگان (جدول پ.۱ و پ.۲ در پیوست)
۲. فرم‌های تکمیل‌شده توسط تکنسین و ماموگرافیست (جدول پ. ۳ در پیوست)
۳. نظرات نهایی پزشکان در خصوص نتایج تشخیص سرطان پستان (جدول ۱)

جدول ۱: فرم نظرات پزشک معالج در خصوص تشخیص سرطان پستان

یافته‌های مهم	نظر پزشک
یافته‌های بالینی مهم	شرح نکات مهم یافته‌های بالینی
یافته‌های ماموگرافی	نتایج بررسی تصویر ماموگرافی
یافته‌های سونوگرافی	نتایج بررسی تصویر سونوگرافی
یافته‌های MRI	نتایج بررسی تصویر MRI
نتایج پاتولوژی قبلی	شرح مقایسه پاتولوژی قبلی و فعلی

اطلاعات پرسشنامه‌های عادات غذایی از مراجعه‌کنندگان جمع‌آوری شده است، در طراحی این پرسشنامه از ویژگی‌های جدید مانند استفاده از ماکروفر در آشپزخانه، مصرف برنج، استعمال قلیان استفاده شده است (جدول پ.۱ در پیوست). اطلاعات پرسشنامه عوامل فرهنگی نیز از مراجعه‌کنندگان جمع‌آوری شده است، سوالات این پرسشنامه از ویژگی‌هایی مانند جنس لباس‌های مورد استفاده، نحوه رفتار با خانواده و درآمد ماهیانه است (جدول پ.۲ در پیوست). اطلاعات کلیدی مراجعه‌کنندگان توسط تکنسین و ماموگرافیست تکمیل شده است (جدول پ.۳ در پیوست). پزشک معالج در طول فرایند تشخیص سرطان پستان، پس از بررسی تصاویر ماموگرافی، علایم بالینی و نتایج آزمایشگاهی نظر خود را اعلام می‌نماید. در تشخیص سرطان پستان یافته‌های بالینی مهم، ماموگرافی، سونوگرافی، MRI و نتایج پاتولوژی برای پزشک تعیین کننده است. جدول ۱ نشان‌دهنده نظرات نهایی پزشکان در خصوص نتایج تشخیص سرطان پستان است. مجموعه داده شامل ۷۱۱ نمونه و ۶۳ ویژگی می‌باشد.

در این مقاله اطلاعات موجود در پرونده شامل نام، نام‌خانوادگی، شماره پرونده بیمار و نشانی حذف شدند. پرونده بیمارانی را که فقط یک بار مراجعه داشته‌اند، کنار گذاشته شد؛ زیرا اطلاعات کامل از آزمایش‌ها در دسترس نبود. نمونه‌هایی که برای ویژگی‌های یافته‌های بالینی مهم، ماموگرافی، سونوگرافی، MRI و نتایج پاتولوژی قبلی، مقدار ناقص داشتند، به دلیل اهمیتی که این ویژگی‌ها در نتیجه نهایی دارند، حذف شدند. هونگجو لو ثابت کرده است که حذف عاقلانه، بهتر از جایگزین کردن ویژگی‌های پر اهمیت با تکنیک‌های مانند میانگین، انتساب تصادفی، انتساب رگرسیون و مدل‌های بیزی است. (۱۴) نمونه‌های که دارای چند ویژگی مفقوده بودند نیز حذف گردیدند. قد و وزن نیز به تنهایی اهمیتی ندارند، بلکه شاخص توده بدنی آن‌ها تاثیرگذار است. در نتیجه این ویژگی‌ها حذف شدند و به‌جای آن‌ها از شاخص‌های مرتبط استفاده شده است. اسامی کلاس‌ها و تعداد نمونه‌های موجود در هر کلاس در جدول ۲ نشان داده‌شده است.

جدول ۲: اسامی و تعداد نمونه‌های کلاس‌ها

نام کلاس	تعداد نمونه‌های موجود در کلاس
سالم	۵۵۲
خوش‌خیم	۱۳۳
بدخیم	۲۶

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 می‌باشد. در بیش‌تر بیماری‌ها داده‌های آموزشی کافی برای الگوریتم یادگیرنده وجود ندارند و یا فراهم نمودن آن زمان‌بر و مشکل است. فرایند یادگیری برای

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]

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 R^n با استفاده از تبدیل غیرخطی Ø، به فضای R^m که m>n است منتقل می‌شود. برای یافتن مرز تصمیم‌گیری بهینه از رابطه زیر استفاده می‌گردد:

$$\max_{av\ldots an} = \frac{-1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} ai\ aj\ yi\ yj((Ø(xi).Ø(xj)) + \sum_{i=1}^{n} ai$$

$$0 \le ai \le C \quad i=1, \ldots, N$$

$$\Sigma_{i=1}^{N}\ ai\ yi=o$$

(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 Ø به‌صورت زیر استفاده می‌گردد:

$$K(xi,xj)=Ø(xi)\ Ø(xj)$$

(۲)

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

$$K(x_i,x_j)=e^{(-¥|x-x_i|^2)}$$

(۳)

انتخاب مناسب پارامترهای C و ¥ در الگوریتم ماشین بردار پشتیبان دارای اهمیت است. پارامتر C میزان جریمه را نشان داده و در خروجی دسته‌بند تاثیرگذار است (۱۶).

الگوریتم زنبور عسل ژنتیکی: الگوریتم GBC[3] رویکرد جدیدی است که توان افزایش دقت و کاهش خطا در پیش‌بینی را با محاسبات حداقل و مرتبه زمانی کمتر از کارهای مشابه دارد (۱۷). روال بهینه‌یابی در الگوریتم GBC براساس یک روند تصادفی- هدایت شده استوار می‌باشد. این روش بر مبنای نظریه تکامل تدریجی و ایده‌های الگوریتم زنبور عسل پایه‌گذاری 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 به صورت تصادفی فضای پاسخ‌ها را ایجاد می‌کنند و به وسیله تابع شایستگی کیفیت موقعیت‌های بازدید شده را گزارش می‌دهند. جواب‌های ایجاد 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 نقطه‌ی بیشینه تابع شایستگی جستجو می‌کند.

در نهایت، به منظور دستیابی به یک تعادل بین بهره‌برداری و اکتشاف در الگوریتم ABC (۱۸) و بهبود توانایی‌های جستجوی محلی، از اپراتورهای جهش الگوریتم ژنتیک در طول فرآیند جایگزینی راه حل‌ها استفاده می‌شود. الگوریتم GBC با دادن جمعیت تصادفی اولیه از فضای جستجو یعنی ویژگی‌های استخراج شده شروع می‌شود:

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

(۴)

که در رابطه بالا X_{mi} یک بردار راه‌حل برای مساله بهینه‌سازی و $m=1,\ldots,S_N$ و $i=1,\ldots,n$ می‌باشد. S_N بیان گر تعداد جمعیت اولیه و هر X_i یک بردار n بعدی است. سپس تابع برازش هر راه حل محاسبه می‌شود.

$$P_m = \frac{Fit_m(x_m)}{\Sigma^{sn}_{m=1}\ fit_m(x_m)}$$

(۵)

و $fit(x_m)$ تابع برازش x_m می‌باشد.

مدل پیشنهادی

پیش‌پردازش: در روش پیشنهادی، علاوه بر انتخاب ویژگی‌های موثر در پیش‌بینی سرطان پستان، وزن هر یک از ویژگی‌ها یافته می‌شود. انتخاب ویژگی به منظور حذف ویژگی‌های نامرتبط و بی‌تاثیر در مجموعه داده‌هاست. ویژگی‌های نامرتبط باعث بروز خطا در کلاسه‌بندی می‌شود. در انتخاب ویژگی‌ها، زیرمجموعه‌ای از کل ویژگی‌ها، انتخاب می‌گردد. استفاده از روش انتخاب ویژگی باعث کاهش زمان آموزش، کاهش زمان محاسباتی و افزایش قابلیت تعمیم کلاسه‌بند می‌شود. یک الگوریتم انتخاب ویژگی از یک روش جستجو برای انتخاب زیرمجموعه‌ای از ویژگی‌ها و یک معیار ارزیابی برای امتیازدهی به این زیرمجموعه استفاده می‌کند. در ساده‌ترین الگوریتم، تمام زیرمجموعه‌های ممکن ویژگی‌ها مورد بررسی قرار گرفته و زیر مجموعه‌ای با کمترین مقدار

[3] Genetic Bee Colony

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]

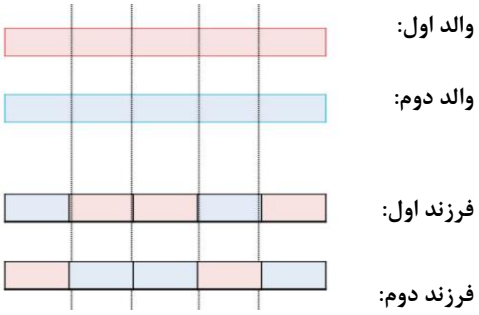
نرخ خطای کلاسه‌بند انتخاب می‌شود. جستجوی کامل فضای ویژگی‌ها دارای بار محاسباتی بالایی است. در این تحقیق از الگوریتم‌های تکاملی برای انتخاب ویژگی استفاده شده است.

در روش پیشنهادی برای مشخص کردن ارزش و نقش هر یک از ویژگی‌ها در تشخیص بیماری به طور تصادفی برای هر ویژگی‌ها وزن صفر یا یک اختصاص داده می‌شود که نشان‌دهنده انتخاب یا عدم انتخاب ویژگی است. مقادیر وزن‌دار ویژگی‌ها به عنوان ورودی به الگوریتم‌های کلاسه‌بندی داده شده و توسط الگوریتم GBC بهینه می‌گردد. هر راه‌حل در جمعیت اولیه در الگوریتم GBC مجموعه ای از مقادیر تصادفی (صفر یا یک) است که بیان‌گر وزن هریک از ویژگی می‌باشد.

به‌کارگیری الگوریتم GBC: الگوریتم GBC ابتدا یک راه حل به صورت تصادفی یا جمعیت اولیه از اندازه SN ، که در آن SN نشان‌دهنده اندازه جمعیت و یا تعداد کل منابع غذایی است، تولید می‌کند. اگر i یک راه حل خاص فرض شود($i = 1, 2, \ldots SN$)، هر راه حل یک بردار چندبعدی D به وجود می‌آورد ($j = 1, 2, 3, \ldots D$) که در آن D نشان‌دهنده تعداد ویژگی‌های آموزنده در هر راه حل بهینه شده است. هر سلول X_{ij} نشان‌دهنده 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 راه حل وابسته در الگوریتم GBC مرتبط است. اگر مقدار دقت جدید بهتر از ارزش دقتی است که قبلا به دست آمده است، پس از آن زنبور راه حل قدیمی را رها می‌کند (منبع غذایی)، و به سوی یکی از راه حل‌های جدید حرکت می‌کند. درغیر آن صورت آن راه حل‌های قدیمی را حفظ می‌کند. بعد از مقداردهی اولیه از راه حل‌های تصادفی (جمعیت)، الگوریتم GBC شروع به جستجو برای راه حل مطلوب می‌کند.

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

مرحله زنبورعسل ناظر: در این مطالعه، عملیات همبری برای به اشتراک‌گذاری اطلاعات بین زنبورهای عسل کارگر و ناظر در فضای جستجو بهینه‌سازی (کندو) استفاده می شود. زنبورهای ناظر با تماشای رقص زنبورهای عسل کارگر محل راه حل (منبع غذایی) را یاد می‌گیرند. در الگوریتم پیشنهادی GBC، زنبورهای عسل ناظر از محل بهترین منبع غذایی، که راه حلی با بالاترین مقدار برازندگی دارد (به‌عنوان ملکه زنبور عسل در الگوریتم ما 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 ارایه شده است. در مدل پیشنهادی، بهترین نسل به‌دست آمده از نتیجه همبری، به عنوان یک خروجی در نظر گرفته می شود (شکل ۱).



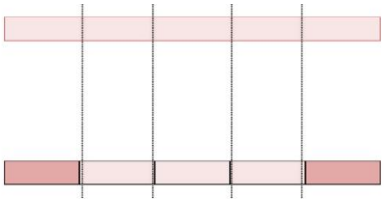
شکل۱ : همبری یکنواخت

مرحله زنبور دیده‌بان: به منظور دستیابی به تعادل بین قابلیت بهره برداری و اکتشاف الگوریتم ABC ، حرکات زنبور عسل دیده‌بان را در الگوریتم ABC اصلی اصلاح شده است. در الگوریتم پیشنهادی به منظور بهبود فرآیند بهره‌برداری، از اپراتورهای جهش GA در طول روند جایگزینی راه حل‌های قدیمی استفاده شده است.

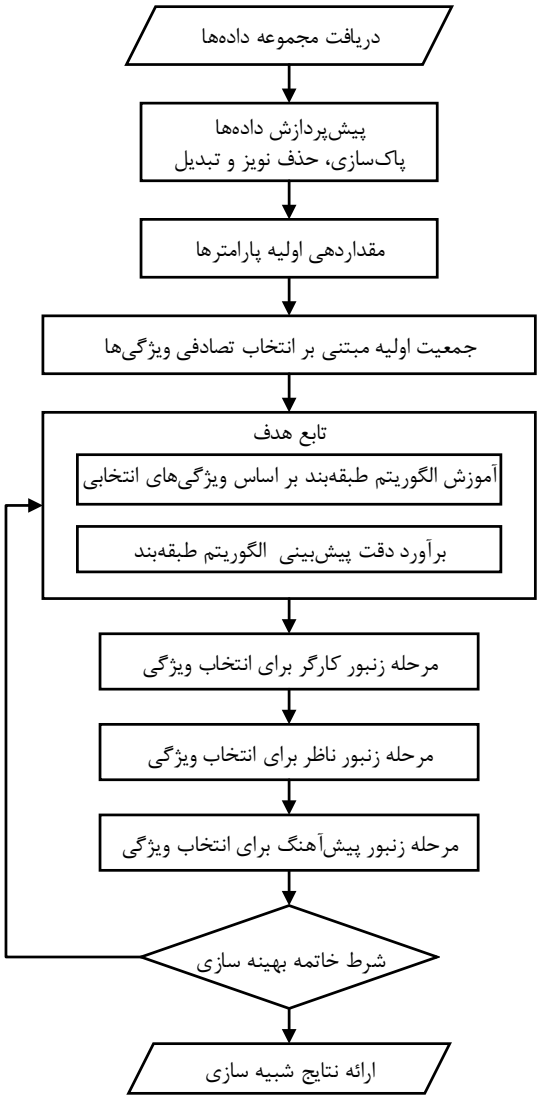
اولین زنبورعسل دیده‌بان طبق الگوریتم پایه ABC فرایند جستجو را بازنشانی کرده و به طور تصادفی فضای جدیدی از راه حل را کشف می‌کند. متعاقبا، یک عملگر جهش مکانی را برای جایگذاری زنبور عسل دیده‌بان دوم می‌گیرد. در زنبور عسل دیده‌بان دوم، ما به دنبال یک محل در سراسر فضای جستجو با توجه به راه حلی با

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]

بالاترین برازندگی که تاکنون تولید شده است هستیم. نرخ احتمال جهش ۰/۰۱ تنظیم شده است (شکل ۲). ساختار روش پیشنهادی در شکل ۳ نشان داده شده است.



شکل۲: عملگر جهش



شکل ۳: ساختار روش پیشنهادی

پیاده‌سازی روش پیشنهادی: عمل کلاسه‌بندی با مقادیر ویژگی‌های انتخاب شده، که بر اساس هریک از زنبورها محاسبه می‌گردد، به روش ماشین بردار پشتیبان انجام می‌شود. هر زنبور که کلاسه‌بندی را با خطای پیش‌بینی کم‌تری انجام دهد، شایسته‌تر خواهد بود. تخمین خطا بر مبنای نمونه‌گیری به روش 10-Fold Cross Validation (۱۹) انجام می‌شود. از معیار accuracy برای انتخاب و حذف یک ویژگی استفاده شده است. تابع برازندگی نشان‌دهنده شایستگی یا توانایی هر کروموزوم است، مراحل الگوریتم تنظیم تابع برازندگی روش پیشنهادی به شرح زیر می‌باشد.

۱. مجموعه داده‌ها فراخوانی می‌شود.
۲. مقادیر ویژگی‌ها نرمال می‌شود.
۳. مجموعه ای از مقادیر (صفر یا یک) به هریک از ویژگی‌ها اختصاص داده می‌شود.
۴. بر اساس ویژگی‌هایی که وزن یک دارند، طبقه‌بندی انجام می‌شود.
۵. میزان خطای طبقه‌بند محاسبه می‌شود.

یافته‌ها

نتایج شبیه‌سازی نهایی با انتخاب ویژگی‌های مربوط به عادات غذایی، عوامل فرهنگی، علایم بالینی و نتایج آزمایشگاهی صورت گرفته است. جدول ۳ ویژگی‌های انتخاب شده و حذف شده از عادات غذایی را نشان می‌دهد. همان‌گونه که مشاهده می‌شود، مصرف دخانیات در ابتلا به سرطان پستان تاثیرگذار است. جدول ۴ نشان‌دهنده ویژگی‌های انتخاب شده و حذف شده از عوامل فرهنگی می‌باشد. با توجه به مقادیر جدول فوق، کار مداوم در شیفت شب ارتباط معنادار با سرطان پستان داشته است. جدول ۵ مهم‌ترین اطلاعات در زمینه تشخیص سرطان پستان را دارد. در این جدول، ویژگی‌های مربوط به فرم تکمیل‌شده توسط تکنسین و ماموگرافیست نشان داده شده است. مقایسه تصاویر ماموگرافی قبلی با فعلی، سفتی قابل لمس در پستان، توده و ترشح از مهم‌ترین ویژگی‌های این جدول می‌باشد. پس از انتخاب نهایی ویژگی‌ها توسط الگوریتم بهینه‌ساز، مدل پیشنهادی با سه روش فازی (۲۰)، شبکه عصبی MLP (۲۱) و شبکه عصبی RBF (۲۲) مقایسه شده است. ارتباط بین کلاس‌های واقعی و کلاس‌های پیش‌بینی شده با استفاده از ماتریس Confusion قابل محاسبه است. در شکل ۴، پارامترهای مورد نیاز ماتریس Confusion ذکرشده است. برای مقایسه مدل پیشنهادی با سایر روش‌ها از

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
DOI: 10.30699/acadpub.ijbd..11.3.71]

معیارهای Accuracy، Sensitivity، Specificity، Precision و F-Measure طبق روابط جدول ۱ استفاده می‌شود (۲۳):

جدول ۱: معیارهای مقایسه مدل پیشنهادی با سایر روش‌ها

Accuracy = (TP + TN) / All	(۶)
Sensitivity = TP / (TP + FN)	(۷)
Specificity = TN / (FP + TN)	(۸)
Precision = TP / (TP + FP)	(۹)
Recall = TP / (TP + FN)	(۱۰)
F Measure= 2 * Precision *Recall --- Precision + Recall	(۱۱)

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.

شکل ۴، مقایسه درصد خطا برای پیش‌بینی سرطان پستان بین الگوریتم‌های GA[۴] (۲۴)، ACO[۵] (۲۵)، ABC[۶] (۲۶)، PSO[۷] (۲۷) و GBC در ۵۰۰ مرحله نشان داده شده است. همان‌طور که مشاهده می‌شود، دقت پیش‌بینی سرطان پستان با استفاده از الگوریتم پیشنهادی افزایش می‌یابد. شکل ۵ نشان‌دهنده نمودار نتایج تشخیص روش‌های مختلف با معیار Accuracy است. همان‌گونه که مشاهده می‌شود مدل پیشنهادی دقت بیش‌تری نسبت به سایر روش‌ها دارد. هم‌چنین مقایسه نتایج پیش‌بینی بیماری با معیارهای Sensitivity و Specificity به ترتیب در جدول‌های ۶ و ۷ نمایش داده شده‌است. جدول ۸ و ۹ نتایج روش‌ها به‌ترتیب با معیار Precision و F-Measure مورد مقایسه قرار گرفته است. نتایج مقایسه نشان‌دهنده برتری عملکرد مدل پیشنهادی می‌باشد.

بحث

مدل پیشنهادی با معرفی روشی جهت پیش‌بینی سرطان پستان با استفاده از بهبود الگوریتم زنبور عسل ژنتیکی، به انتخاب ویژگی‌ها از بین عادات غذایی، عوامل فرهنگی، علایم بالینی و نتایج آزمایشگاهی می‌پردازد. در این مطالعه بر اساس یک پایگاه داده بومی، مجموعه‌ای از الگوهای کمتر شناخته شده و موثر در بروز سرطان پستان مورد پردازش قرار گرفته است. پژوهش‌های مربوط به پیش‌بینی سرطان پستان 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 داده های از دست رفته به‌کار گرفته شده است.

ویژگی‌های انتخاب شده در پیش‌بینی سرطان پستان در جدول‌های ۳، ۴ و ۵ قابل مشاهده است. ویژگی‌های منتخب در ارتباط با تشخیص سرطان پستان از بین عادات غذایی و عوامل فرهنگی شامل: مجرد بودن، عدم فعالیت ورزشی، مصرف سیگار و قلیان، کار در شیف شب است. همچنین مقایسه نتیجه ماموگرافی قبلی و فعلی، مدت مصرف قرص ضد بارداری، هیستروکتومی، جایگزینی هورمون، میزان تورفتگی نوک پستان، میزان درد، نوع ترشح و توده موجود در تصاویر ماموگرافی موثر در تشخیص سرطان پستان شناخته شد. هیچ‌گونه رابطه معناداری بین بروز سرطان پستان و ویژگی‌هایی مانند استفاده از ماکروفر در آشپزخانه، نوع برنج وروغن مصرفی پیدا نشده است.

آتشی و همکاران تعداد ۱۰۰ رابطه انجمنی با ضرایب اطمینان بالاتر از ۹۰٪ کشف کردند. تعداد ۱۰ رابطه از این روابط معنی‌دار گزارش شد (۲۸). کاوراسیا و همکاران از درخت تصمیم جهت پیش‌بینی سرطان پستان با دقت ۷۴٪ استفاده کردند (۲۹).

طلوعی و همکاران پیش بینی عود مجدد سرطان پستان به کمک سه تکنیک داده کاوی را ارایه دادند (۳۰).

بررسی‌های صورت گرفته نشان می‌دهد که دقت در سه الگوریتم داده کاوی، یعنی درخت تصمیم گیری، ANN و SVM به ترتیب برابر ۰/۹۳۶، ۰/۹۴۷، ۰/۹۵۷ بوده است.

[4] Genetic algorithm
[5] Ant colony optimization
[6] Artificial bee colony
[7] Particle swarm optimization

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]

جدول ۳: ویژگی‌های انتخاب شده و حذف شده از عادات غذایی جهت پیش‌بینی سرطان

ویژگی‌های انتخاب شده پس از فرایند بهینه‌سازی	حذف ویژگی‌ها در فرایند بهینه‌سازی
۱- آیا سیگار مصرف می‌کنید؟	۱- آیا در آشپزخانه از ماکروفر استفاده می‌کنید؟
۲- چه مدتی سیگار مصرف می‌کنید؟	۲- بیشتر از کدام برنج مصرف می‌کنید؟
۳- یا قلیان مصرف می‌کنید؟	۳- کدام نوع روغن را مصرف می‌کنید؟
۴- چه مدتی قلیان مصرف می‌کنید؟	۴- چند لیوان شیر در هفته می‌خورید؟
۵- آیا در معرض دود سیگار هستید؟	۵- آیا همراه با غدا ماست می‌خورید؟
۶- چه مدتی در معرض دود سیگار بوده‌اید؟	۶- آیا غذای آماده مانند ساندویچ می‌خورید؟
۷- آیا در معرض دود قلیان هستید؟	۷- سبزی را هر چند وقت می‌خورید؟
۸- چه مدتی در معرض دود قلیان بوده‌اید؟	۸- میوه را هر چند وقت می‌خورید؟
	۹- شیرینی را هر چند وقت می‌خورید؟
	۱۰- گوشت قرمز را هر چند وقت می‌خورید؟
	۱۱- گوشت مرغ را هر چند وقت می‌خورید؟
	۱۲- گوشت ماهی را هر چند وقت می‌خورید؟

جدول ۴: ویژگی‌های انتخاب شده و حذف شده از عوامل فرهنگی جهت پیش‌بینی سرطان

ویژگی‌های انتخاب شده پس از فرایند بهینه‌سازی	حذف ویژگی‌ها در فرایند بهینه‌سازی
۱- اگر شاغل هستید آیا شیفت شب دارید؟	۱- در کدام منطقه زندگی می کنید؟
۲- وضعیت تاهل شما چیست؟	۲- آیا از لوازم آرایش استفاده می‌کنید؟
۳- آیا از خانواده شما مبتلا به سرطان پستان وجود دارد؟	۳- معمولا جنس پارچه سوتین شما چیست؟
۴- آیا ورزش می‌کنید؟	۴- ضد عرق برای زیر بغل شما چیست؟
۵- چگونه ورزش می‌کنید؟	۵- درآمد ماهانه خانواده شما کدام است؟
	۶- شب‌ها در چه ساعتی می‌خوابید؟
	۷- صبح‌ها در چه ساعتی از خواب بیدار می‌شوید؟
	۸- رابطه عاطفی با همسرتان چگونه است؟
	۹- رابطه عاطفی شما با والدین چگونه است؟
	۱۰- کدام واقعه ناگوار را داشته‌اید؟
	۱۱- کدام مشکل را داشته‌اید؟

جدول ۵: ویژگی‌های انتخاب شده و حذف شده از فرم تکمیل‌شده توسط تکنسین و ماموگرافیست جهت پیش‌بینی سرطان

ویژگی‌های انتخاب شده پس از فرایند بهینه‌سازی	حذف ویژگی‌ها در فرایند بهینه‌سازی
۱- تعداد ماموگرافی قبلی؟	۱- نوع قرص ضد بارداری مصرفی؟
۲- مدت مصرف قرص ضد بارداری؟	۲- مدت شیردهی برای نوزادن چند ماه است؟
۳- هیستروکتومی؟	۳- علت قطع قاعدگی طبیعی شما چه بوده است؟
۴- جایگزینی هورمون HRT	۴- آیا به موارد خاصی حساسیت یا آلرژی دارد؟
۵- تو رفتگی نوک پستان N/R	۵- برای مدت طولانی داروی خاصی را مصرف می‌کند؟
۶- درد؟	
۷- ترشح؟	
۸- توده؟	
۹- سفتی قابل لمس؟	
۱۵- قبلا جراحی پستان داشته است؟	
۱۶- آیا از خانواده شما مبتلا به هر نوع سرطان وجود دارد؟	
۱۷- آیا قبلا سابقه بیماری و ضایعه پستان داشته‌است؟ (مانند آبسه، عفونت پستان)	
۱۸- آیا سابقه تروما به پستان داشته‌است؟	
۱۹- مدت عادت ماهیانه شما در سال اخیر کدام است؟	
۲۰- چند ماه از آخرین پریود شما می‌گذرد؟	

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
DOI: 10.30699/acadpub.ijbd..11.3.71]

کیانی و همکاران در مدل توسعه یافته به ویژگی و حساسیت به ترتیب ۶۵٪ و ۹۶٪ رسیدند. در این مطالعه گذشته‌نگر، از داده‌های ۸۰۹ بیمار مبتلا به سرطان پستان و ۸۹ ویژگی از هر بیمار، استفاده شد(۳۱).

لند و ورهگن استفاده از ماشین بردار پشتیبان بر روی مجموعه داده‌های سرطان پستان برای پیش‌بینی استفاده کردند (۳۲). سلاما و همکاران روش شبکه فازی- عصبی را برای پیش‌بینی سرطان پستان به کار بردند. دقت ۸۷٪ در تشخیص سرطان پستان با استفاده از درخت C4.5 نتیجه کار سوریا و همکاران بود (۳۳). دقت پیش‌بینی این روش ۹۵/۰۶٪ بود (۳۴). در حالی که مدل پیشنهادی با دقت ۹۷/۲٪ دارای عملکرد بهتری نسبت به روش روش‌های فازی، شبکه عصبی MLP و شبکه عصبی RBF در تشخیص سرطان پستان است.

چاراسیا و همکاران با استفاده از روش SVM به دقت ۹۶/۴٪ دست یافتند (۳۴). سروستانی و همکان برای پیش‌بینی درجه بدخیمی سرطان پستان به مقایسه میانگین مربع خطا در شبکه‌های عصبی چند لایه، رقابتی و پایه شعاع پرداختند که بهترین دقت را شبکه عصبی پایه شعاعی داشت (۲۸). لاوانیا و همکارن با استفاده داده‌های پایگاه داده WBCD و درخت تصمیم با دسته‌بندی دو مرحله‌ای به دقت ۹۴/۸۴٪ رسیدند (۳۵). آرونا و همکاران با استفاده از دسته‌بندی کننده‌های با ناظر مانند روش نایو بیز و درخت تصمیم برروی مجموعه داده‌های پایگاه داده WBCD به پیش‌بینی سرطان پستان پرداختند. بر اساس معیارهای Accuracy، Sensitivity، Specificity، Precision، Recall و F-Measure در نتایج شبیه‌سازی، مدل پیشنهادی نتایج بهتری نسبت به این مقاله دارد (۳۶). نتایج آزمایشات نشان داد که ماشین بردار پشتیبان دقت ۹۶/۷٪ را دارد. کیان و همکارن با به‌کارگیری روش RBF و MLP برای پیش‌بینی سرطان پستان به ترتیب به دقت ۹۶/۱۸٪ و ۹۵/۷۴٪ رسیدند (۳۷).

در روش پیشنهادی با هدف طراحی و ارزیابی یک مدل پزشک‌یار در تشخیص سرطان پستان، به انتخاب ویژگی های موثر در بروز این بیماری پرداخته شد. همچنین با به‌کارگیری الگوریتم GBC، دقت ماشین بردار پشتیبان افزایش یافت. آزمایشات و شبیه‌سازی نشان داد سیستم پزشک‌یار معرفی شده در این پژوهش بر روی مجموعه داده بیماران بومی مبتلا به سرطان پستان بیمارستان مرتاض یزد به دقت ۹۷/۲٪ رسیده است که بالاتر از تحقیقات مشابه بر روی مجموعه دادهای متفاوت بوده است.

نتیجه‌گیری

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 سرطان پستان با استفاده از هوش مصنوعی و یادگیری ماشین، احتمال درمان موفق را بالا می‌برد.

روش پیشنهادی با بهبود الگوریتم زنبور عسل ژنتیکی به انتخاب ویژگی‌های موثر در پیش‌بینی سرطان پستان از بین عادات غذایی، عوامل فرهنگی، علایم بالینی و نتایج آزمایشگاهی می‌پردازد در این مقاله برای پیش‌بینی و تشخیص سرطان پستان، از الگوریتم GBC برای بهینه‌سازی نتایج ماشین بردار پشتیبان استفاده شد و یک مدل جدید ارایه گردید. نتایج شبیه‌سازی نشان می‌دهد که مدل پیشنهادی با دقت پیش‌بینی ۰/۹۷۲ از روش‌های فازی، شبکه عصبی MLP و شبکه عصبی RBF دقت بیش‌تری دارد. کار در شیفت شب، وضعیت تاهل، سابقه خانوادگی و ورزش در پیش‌بینی موثر بوده‌اند. استعمال دخانیات مانند سیگار و قلیان در بروز ابتلا به سرطان پستان موثر بوده است. تعداد ماموگرافی قبلی، مدت مصرف قرص ضد بارداری، هیستروکتومی، جایگزینی هورمون HRT، تو رفتگی نوک پستان N/R، درد، ترشح، توده، سفتی قابل لمس در پستان، سابقه خانوادگی در تشخیص سرطان پستان تاثیر داشته‌اند.

تقدیر و تشکر

نویسندگان بر خود لازم می‌دانند از مدیریت محترم بیمارستان فوق تخصصی مرتاض یزد به خاطر تامین مجموعه داده‌های مورد نیاز جهت انجام این پژوهش تشکر نمایند.

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]

تعارض منافع

نویسندگان اعلام می‌دارند که هیچ تعارض منافعی در پژوهش حاضر وجود ندارد.

References

1. Nakai K, Xia W,Liao H, et al. The role of PRMT1 in EGFR methylation and signaling in MDA-MB-468 triple-negative breast cancer cells. Breast Cancer. 2018; 25(1): 74-80.
2. Rheinbay E, Parasuraman P,Grimsby J, et al. Recurrent and functional regulatory mutations in breast cancer. Nature 2017; 547(7661): 55-60.
3. Inoue M, Nakagomi H, Nakada H, et al. Specific sites of metastases in invasive lobular carcinoma: a retrospective cohort study of metastatic breast cancer. Breast Cancer 2017; 24(5): 667-72.
4. Quist J,Mirza H, Chon U cheang M, et al. Association of a four-gene decision tree signature with response to platinum-based chemotherapy in patients with triple negative breast cancer. Clinical oncology 2017; 1006.
5. Quail D F,C.olson o,Bhardwaj P, et al. Obesity alters the lung myeloid cell landscape to enhance breast cancer metastasis through IL5 and GM-CSF. Nature cell biology 2017; 19(8): 974.
6. Denkert C, Liedtke C, Tutt A, et al. Molecular alterations in triple-negative breast cancer- the road to new treatment strategies. The Lancet 2017; 389(10087): 2430-42.
7. Tomioka N, Azuma M, Ikarashi M, et al. The therapeutic candidate for immune checkpoint inhibitors elucidated by the status of tumor-infiltrating lymphocytes (TILs) and programmed death ligand 1 (PD-L1) expression in triple negative breast cancer (TNBC). Breast Cancer. 2018; 25(1): 34-42.
8. Heitz A E, Baumgartner R, Baumgartner K, et al. Healthy lifestyle impact on breast cancer-specific and all- cause mortality. Breast cancer research and treatment. 2018; 167(1): 171-81.
9. Howlader N, cronin K, Kurian A, et al. Differences in Breast Cancer Survival by Molecular Subtypes in the United States. Cancer Epidemiology and Prevention Biomarkers 2018; 27(6):619-26.
10. Ster B, Dobnikar A. Neural networks in medical diagnosis: Comparison with other methods. Proc Int Conf Eng Appl neural networks 2016; 427-30.
11. Sheikhpour R, Sarram M, Zare Mirakabad M, et al. Breast Cancer Detection Using Two-Step Reduction of Features Extracted from Fine Needle Aspirate and Data Mining Algorithms. Iranian Quarterly Journal of Breast Disease 2015; 7(4).
12. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. Artificial Intelligence in Medicine 2005; 34(2): 113-27.
13. Mohd Zamani M, Musirin I, Suliman S, et al. Multi-area economic dispatch performance using swarm intelligence technique considering voltage stability. International Journal on Advanced Science. Engineering and Information Technology 2017; 7(1): 1-7.
14. Lu H, Rudy S, Liu H. Neurorule: A connectionist approach to data mining. arXiv preprint arXiv. 2017; 1701 (01358).
15. Saed S. Support Vector Machine. Support Vector Machines. Np, nd Web 26 2017.
16. Gu B, S. Sheng V. A robust regularization path algorithm for $\nu $-support vector classification. IEEE Transactions on neural networks and learning systems 2017; 28(5): 1241-8.
17. Alshamlan MH, Ghada HB, A. Alohali Y. Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification. Computational biology and chemistry 2015; 56: 49-60.
18. Hancer E, Xue B, Zhang M, et al. Pareto front feature selection based on artificial bee colony optimization. Information Sciences 2018; 422: 462-79.

[Downloaded from ijbd.ir on 2026-11-13]

[DOR: 20.1001.1.17359406.1397.11.3.2.2]

[DOI: 10.30699/acadpub.ijbd..11.3.71]

19. Rohani A, Taki , Abdollahpour M. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). Renewable Energy 2018; 115: 411-22.

20. Taneja Sh, Suri B, Gupta S, et al. A Fuzzy Logic Based Approach for Data Classification. Data Engineering and Intelligent Computing. Springer, Singapore 2018; 605-16.

21. Hramov AE, Forolov NS, Maksimenko VA, et al. Artificial neural network detects human uncertainty. Chaos: An Interdisciplinary Journal of Nonlinear Science 2018; 28(3): 033607.

22. Agrawal Sh, Agrawal j, Kaur Sh, et al. A comparative study of fuzzy PSO and fuzzy SVD- based RBF neural network for multi-label classification. Neural Computing and Applications 2018; 29(1): 245-56.

23. Prahlad A, Parag G, Rajiv K, et al. Data mining systems and methods for heterogeneous data sources. U.S. Patent No. 2018; 9: 738-910.

24. Kadri R, F. Boctor F. An efficient genetic algorithm to solve the resource-constrained project scheduling problem with transfer times: The single mode case. European Journal of Operational Research 2018; 265(2): 454-62.

25. Mellal M, Williams E J. A survey on ant colony optimization, particle swarm optimization, and cuckoo algorithms. Handbook of research on emergent applications of optimization algorithms. IGI Globa. 2018; 37-51.

26. Karaboga De, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. Journal of global optimization.2007; 39(3): 459-71.

27. Kenned J. Particle swarm optimization. Encyclopedia of machine learning. Springer US 2011; 760-6.

28. Atashi A, Kiani B. Discover hidden patterns in the actual data set of patients with breast cancer using the data mining technique. Journal of Informatics. Health and Biomedicine 1394; 8 (1): 60-5.

29. 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.

30. Toloei Ashlaghi A, Pour Abrahamid A, Ebrahimi M, at al. Prediction of recurrence of breast cancer using three data mining techniques. Journal of Breast Diseases of Iran 2012; 5 (4): 34-23.

31. Kiani B, Atashi A. Create a data mining prognostic model to predict the recurrence of breast cancer. Journal of Informatics. Health and Biomedical Sciences 1393; 1(1): 26-31.

32. Land WH, Verheggen EA. Multiclass Primal Support Vector Machines For Breast Density Classification. Int J Comput Biol Drug Des 2009; 2(1):21-57.

33. Salama GI, Abdelhalim M, Zeid M. Breast cancer diagnosis on three different datasets using multi-classifiers. Breast Cancer (WDBC) 2012; 190-6.

34. 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.

35. 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.

36. Aruna S, Rajagopalan DS, Nandakishore LV. Knowledge based analysis of various statistical tools in detecting breast cancer. Computer Science & Information Technology 2011; 2: 37-45.

37. Kiyan T, Yildirim T. Breast cancer diagnosis using statistical neural networks. IU-Journal of Electrical & Electronics Engineering 2011; 4(2): 1149-53.

[Downloaded from ijbd.ir on 2026-11-13]
[DOR: 20.1001.1.17359406.1397.11.3.2.2]
[DOI: 10.30699/acadpub.ijbd..11.3.71]