مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره هفتم شماره دوم. ص ۱۸۹-۱۸۱، ۱۳۹۹

پیش‌بینی متاستاز سرطان پستان با استفاده از مدل‌های فازی و داده‌های بیماران ایرانی مبتلا به سرطان پستان

آرزو همایی فصیح[۱]، علیرضا آتشی[۲،۳]*، حجت احسنی طهرانی[۴]

• دریافت مقاله: ۹۸/۷/۲۶ • پذیرش مقاله: ۹۹/۱/۳۰

مقدمه: متاستاز سرطان پستان، گسترش سرطان پستان به سایر اندام‌های بدن، یکی از مهم‌ترین علل مرگ ناشی از سرطان پستان در زنان محسوب می‌شود. پیش‌بینی متاستاز در مراحل اولیه به انتخاب بهترین روش درمانی و بهبود کیفیت زندگی بیماران کمک می‌کند.

روش: در این مطالعه بنیادی از مجموعه داده‌های موجود بیماران ایرانی از مرکز تحقیقات سرطان پستان پژوهشکده سرطان معتمد تهران، استفاده شد. مطالعه حاضر سیستم استنتاج فازی ممدانی، تاکاگی سوگنو و سیستم استنتاج تطبیقی عصبی – فازی (انفیس) برای پیش‌بینی متاستاز سرطان پستان در مراحل اولیه را مورد استفاده قرار داد.

نتایج: بهترین خطای پیش‌بینی با استفاده از سیستم استنتاج تطبیقی عصبی – فازی (انفیس) مبتنی بر الگوریتم خوشه‌بندی میانگین مراکز فازی به دست آمد. نظرات متخصصان در مرکز تحقیقات سرطان پستان در پژوهشکده سرطان معتمد و خطای پیش‌بینی مدل ارزیابی شده نشان داد که این سیستم پیش‌بینی به خوبی شکل گرفته است.

نتیجه‌گیری: بهترین سیستم پیش‌بینی پیشنهادی می‌تواند به عنوان یک سیستم تصمیم‌یار بالینی برای کمک به پزشکان در فرآیند درمان مورد استفاده قرار گیرد.

کلید واژه‌ها: سرطان پستان، متاستاز، سیستم استنتاج فازی ممدانی، سیستم استنتاج فازی تاکاگی سوگنو، سیستم استنتاج تطبیقی عصبی – فازی (انفیس)

•**ارجاع:** همایی فصیح آرزو، آتشی علیرضا، احسنی طهرانی حجت، پورحسین محمدامین. پیش‌بینی متاستاز سرطان پستان با استفاده از مدل‌های فازی و داده‌های بیماران ایرانی مبتلا به سرطان پستان. مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۹؛ ۷(۲): ۹-۱۸۱

۱. کارشناسی ارشد ریاضی کاربردی – آنالیز عددی، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود، شاهرود، ایران
۲. دکترای انفورماتیک پزشکی، گروه پژوهشی انفورماتیک پزشکی، دپارتمان پژوهش‌های بالینی، مرکز تحقیقات سرطان پستان، پژوهشکده معتمد جهاد دانشگاهی، تهران، ایران
۳. دکترای انفورماتیک پزشکی، استادیار ،گروه سلامت الکترونیک، دانشکده مجازی، دانشگاه علوم پزشکی تهران، تهران، ایران
۴. دکترای ریاضی کاربردی، دانشیار، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود، شاهرود، ایران
* **نویسنده مسئول:** علیرضا آتشی
آدرس: تهران، بلوار کشاورز، خیابان نادری، بن‌بست دولتشاهی، پلاک ۲، دانشکده مجازی
• **شماره تماس:** ۰۲۱۴۲۰۳۶۰۰۰
• **Email:** aratashi@sina.tums.ac.ir

[Downloaded from jhbmi.ir on 2025-01-05]

مقدمه

سرطان پستان یکی از شایع‌ترین انواع سرطان‌های تهاجمی در زنان در سراسر جهان است که حدود ۱۰ درصد از زنان را در مراحل مختلف زندگی مبتلا ساخته و ابعاد مختلف زندگی آن‌ها را تحت تأثیر قرار می‌دهد [۴-۱]. این سرطان، شایع‌ترین بدخیمی در میان زنان ایرانی است. در سال‌های اخیر، میزان شیوع بیماری روند روبه رشدی داشته و بررسی داده‌ها نشان می‌دهد که میزان بقای بیماران تا پنج سال و نیز ده سال پس از تشخیص، به ترتیب ۸۸ و ۸۰ درصد بوده است [۲،۳]. تمام تومورها سرطانی نیستند و شامل خوش‌خیم و بدخیم می‌باشند. تومورهای خوش‌خیم علی‌رغم رشد غیرطبیعی به‌ندرت مرگ‌آور هستند. البته در برخی از زنان با سابقه نمونه‌برداری از توده‌های خوش‌خیم پستان، خطر سرطان پستان افزایش یافته است. از طرف دیگر، تومورهای بدخیم جدی‌تر بوده و سرطانی محسوب می‌شوند؛ اما تشخیص زودهنگام این نوع از سرطان‌ها و عوارض دیگر آن‌ها مانند عود مجدد و یا متاستاز، شانس درمان موفقیت‌آمیز را بالا می‌برد [۴]. متاستاز سرطان پستان علت اکثریت مرگ‌ومیرهای ناشی از سرطان پستان به حساب می‌آید. تشخیص متاستاز سرطان پستان در مراحل اولیه به تعیین بهترین شیوه برای مهار و پیشگیری از پیشرفت سرطان پستان و یا کنترل بیماری و بهبود کیفیت زندگی بیمار، کمک خواهد کرد. در حال حاضر، تشخیص متاستاز سرطان پستان بر اساس نشانه‌های بالینی از گسترش به دیگر اندام‌ها، بیوپسی اندام‌های آسیب دیده، ارزیابی‌های رادیولوژیک، روش‌های تصویربرداری و نشانگرهای سرمی تومور صورت می‌گیرد [۷-۵]. شناخت روش‌های پیش تشخیص سطح سرطان می‌تواند بسیار حائز اهمیت باشد. مطالعات انجام شده نشان می‌دهند که استفاده از ابزارهای هوش مصنوعی در پیش‌بینی سطح سرطان‌ها به خصوص سرطان پستان بسیار مفید و کارآمد است [۸]؛ اما در هر نوع بررسی به خصوص در زمینه‌های تخصصی مانند پزشکی اولویت در شناخت و تحلیل داده‌های زمینه کاری است.

منطق فازی نوعی منطق ریاضی است که از آن برای مدیریت ابزارهای مهندسی که می‌توانند به سازمان‌دهی مجموعه بزرگی از داده‌ها کمک کنند، استفاده می‌شود. به عنوان مثال، در مجموعه‌های شامل اطلاعات بیماران، بیمارانی را شناسایی می‌کند که یک الگوی معمول از بیماری یا پیشرفت بیماری را دنبال می‌کنند و در پردازش مجموعه داده‌ها مؤثرتر از سیستم‌های دستی و یا روش‌های تشخیصی فعلی عمل می‌کند. به دلیل ماهیت منطق فازی و نزدیکی آن به شیوه استنتاجی ذهن انسان بسیاری از علوم کلاسیک به تدریج منطق فازی را پذیرفته و از آن در تکمیل روش‌های خود بهره جسته‌اند. از آن جمله می‌توان به تحلیل رگرسیون در علم آمار و احتمال اشاره کرد که به عنوان یکی از روش‌های داده‌کاوی، به دلیل کاربردهای وسیع در تحلیل‌های مربوط به متغیرهای کیفی، از جمله روش‌هایی بوده است که منطق فازی را در درون خود پذیرفته و نتایج درخور توجهی به دست آورده است. در تحقیقات مبتنی بر تحلیل رگرسیون که با عناوین متنوع و در زیرشاخه‌های گوناگونی انجام شده نتایج قابل ملاحظه‌ای حاصل گردیده است، این نتایج به شکل نسبتاً پراکنده در زمینه تشخیص پزشکی هم حاصل شده‌اند [۹،۱۰]. در مدل‌سازی، تجزیه و تحلیل و ارزیابی بیماری‌ها با افزایش پیچیدگی در روند بیماری‌ها، پزشکان بیش‌ازپیش وابسته به سیستم‌های مشاورهای در تصمیم‌گیری‌های پزشکی می‌شوند که بسیاری از آن‌ها با استفاده از منطق فازی طراحی گردیده‌اند [۱۳-۱۱]. استفاده از منطق فازی در مسائل پزشکی خصوصاً سرطان پستان به شکل‌های مختلفی صورت گرفته است و اخیراً تشخیص‌های رایانه محور تبدیل به بخشی از کار بالینی معمول برای تشخیص سرطان پستان در بسیاری از مؤسسات و بیمارستان‌های کشورهای توسعه یافته شده است [۸،۱۱،۱۲،۱۴].

با توجه به اهمیت متاستاز در سرطان و به خصوص سرطان پستان و اولویت تشخیص زود هنگام آن برای بیماران و نیز با توجه به ماهیت داده‌های این نوع سرطان که غالباً از نوع فازی هستند پژوهشگران این مطالعه درصدد برآمدند تا مدل‌های استنتاج فازی و نیز مدل‌های سیستم استنتاج تطبیقی عصبی-فازی (انفیس) (Adaptive Neuro - Fuzzy Inference System) ANFIS را که خود بر اساس یک سیستم استنتاج فازی طراحی می‌شوند، برای داده‌های مربوط به بیماران درگیر با سرطان پستان جهت پیش‌بینی بروز متاستاز این سرطان به کار گیرند. در این راستا با استفاده از داده‌های موجود سرطان پستان یک پایگاه داده در ایران، مدل‌های پیش‌بینی مبتنی بر داده‌های بیماران درگیر با متاستاز برای پیش‌بینی متاستاز سرطان پستان در بیماران در معرض متاستاز (با استیج‌ها و یا مراحل پایین) ارائه شده‌اند. فرض بر این است که این مدل‌ها نتایج بهتری را برای ما رقم می‌زنند.

[Downloaded from jhbmi.ir on 2025-01-05]

روش

در این پژوهش بنیادی، داده‌ها از کلینیک بیماری‌های پستان، پژوهشکده معتمد جهاد دانشگاهی دریافت شده‌اند. این داده‌ها شامل داده‌های ۳۰۳۷ مبتلا به سرطان پستان سطح صفر تا چهار و دارای ۲۲ ویژگی برای هر بیمار بودند؛ که از سال ۱۳۷۵ تا ۱۳۹۴ جمع‌آوری شده بودند. با توجه به عوامل خطر در سرطان پستان و متاستاز آن و با نظر پزشکان این پژوهشکده ویژگی‌ها یا متغیرهایی که تأثیر آن‌ها در ابتلاء به سرطان پستان و نیز متاستاز سرطان پستان هنوز مشخص نشده‌اند و یا متغیرهای فاقد ارزش، از بین متغیرهای موجود حذف گردیدند. در تمامی مدل‌های پیشنهادی در این پژوهش، ویژگی‌های سن تشخیص بیماری، سن یائسگی، داشتن یا نداشتن بارداری، مصرف و یا عدم مصرف داروهای پیشگیری از بارداری، سابقه شخصی ابتلاء به سرطان پستان، سابقه خانوادگی ابتلاء به سرطان پستان و مرحله سرعت و قدرت رشد تومور (گرید) به عنوان متغیرهای ورودی و مرحله بیماری سرطان (استیج) به عنوان تنها متغیر خروجی در نظر گرفته شدند. (جدول ۱) در ارتباط با متغیرهای داشتن یا نداشتن بارداری، مصرف و یا عدم مصرف داروهای پیشگیری از بارداری، سابقه شخصی ابتلاء به سرطان پستان و سابقه خانوادگی ابتلاء به سرطان پستان با بررسی نتایج مختلف حاصل از مدل‌هایی با توابع عضویت متفاوت برای متغیرها و به دلیل نوع داده‌های دریافتی از بیماران که غالباً به شکل بله یا خیر بودند و نیز تعدد نوع داده‌ها برای برخی متغیرها مانند سابقه خانوادگی ابتلاء به سرطان پستان تصمیم بر آن شد که در تعیین توابع عضویت این متغیرها مقادیر صفر و یک را با تقریب‌های مناسب جایگزین داده‌های موجود برای هر متغیر در هر بیمار شود. از آنجاکه وقوع متاستاز در مراحل ۳ به بعد قطعی است؛ بنابراین این مراحل را به عنوان مراحلی که در آن‌ها بیمار دچار متاستاز شده است، در نظر گرفته و تنها داده‌های بیماران دارای مرحله سرطان) استیج (مشخص را مورد بررسی قرار داده شد؛ لذا با توجه به گمشدگی بسیار زیاد داده‌ها در این پایگاه داده، تنها اطلاعات مربوط به ۵۱۹ بیمار قابل استفاده بودند. در واقع همگی این داده‌ها دارای خروجی مشخص بودند؛ اما برخی متغیرهای آن‌ها مانند سن یائسگی و یا مصرف و یا عدم مصرف داروهای پیشگیری از بارداری مفقود بودند. این فقدان مربوط به عدم دریافت اطلاعات کافی از بیمار و یا رخ ندادن بعضی متغیرها در بیمار مانند سن یائسگی شد. به عنوان یکی از فازهای پیش‌پردازش و آماده‌سازی داده‌ها، با توجه به حالات خاص هر داده، برای عود یک متغیر به جای هر مقدار تهی، صفر در نظر گرفته شد و در سایر موارد میانه یا مد جایگزین شد. در حالت دیگر برای جایگزین کردن داده‌های جدید به‌جای داده‌های مفقود می‌توان از (EM) Algorithm Maximization Expectation استفاده کرد که در این مطالعه با توجه به داده‌ها از لحاظ پزشکی لازم نبود. پس از آن برای هر یک از متغیرهای ورودی و خروجی با توجه به واقعیت‌های علوم زیستی و نظر پزشکان متخصص و با استفاده از قوانین فازی، توابع عضویت تعریف شدند. تمامی مقادیر عددی و غیرفازی با استفاده از این توابع عضویت به مقادیر فازی تبدیل گشتند. این فازی‌سازی مقادیر مختلف در گروه‌های متفاوت توسط نرم افزار Matlab نسخه ۱۱ جهت ورود داده‌ها به الگوریتم می‌باشد.

در این پژوهش الگوریتم‌های دو مدل سیستم استنتاج فازی ممدانی و سیستم استنتاج فازی تاکاگی سوگنو(Takagi - Sugeno fuzzy inference system)، مدل تطبیقی عصبی – فازی (انفیس) با سیستم استنتاج فازی اولیه ایجاد شده به روش خوشه بندی میانگین مراکز فازی و مدل تطبیقی عصبی – فازی (انفیس) با سیستم استنتاج فازی اولیه ایجاد شده به روش خوشه‌بندی کاهشی، با نوشتن کدهای مربوطه با استفاده از نرم‌افزار Matlab ایجاد شدند. در مدل‌های ممدانی و تاکاگی سوگنو به مجموعه قواعدی جهت استنتاج نیاز بود که این مجموعه قواعد با بررسی پایگاه داده اولیه و نظر پزشکان خبره تعیین شدند. پس از حذف قواعد تکراری و فاقد ارزش بالینی، ۱۷۸ قاعده به عنوان قوانین معنی‌دار به لحاظ بالینی در نظر گرفته شدند.

در مدل تطبیقی عصبی – فازی (انفیس) نیاز به داده‌های آموزش و تست برای آموزش دادن و ایجاد بهترین ساختار و ارزیابی عملکرد آن بود؛ لذا مجموعه داده ورودی که به عنوان بردار ورودی در فرآیند آموزش انفیس در نظر گرفته شد، ماتریسی دارای ۵۱۹ سطر (اطلاعات مربوط به ۵۱۹ بیمار) و ۷ ستون (شامل متغیرهای سن تشخیص بیماری، سن یائسگی، داشتن یا نداشتن بارداری، مصرف و یا عدم مصرف داروهای پیشگیری از بارداری، سابقه شخصی ابتلا به سرطان پستان، سابقه خانوادگی ابتلاء به سرطان پستان و مرحله سرعت و قدرت رشد تومور) بود. مجموعه داده خروجی که به عنوان بردار خروجی در فرآیند آموزش انفیس در نظر گرفته شد، ماتریسی دارای ۵۱۹ سطر (اطلاعات مربوط به ۵۱۹ بیمار) و ۱ ستون (شامل مرحله بیماری سرطان یا استیج) بود. به همین ترتیب

[Downloaded from jhbmi.ir on 2025-01-05]

مجموعه داده‌های ورودی و خروجی که در فرآیند تست انفیس در نظر گرفته شدند شامل ماتریسی دارای ۴۳۹ سطر و ۷ ستون و نیز ماتریسی دارای ۴۳۹ سطر و ۱ ستون بودند. این نمونه ۴۳۹ تایی با استفاده از داده‌های کامل و تنها با جایگزین کردن سن یائسگی (به شیوه از پیش مطرح شده) در بیماران فاقد سن یائسگی به دست آمد. در این راستا ترکیبی از مجموعه داده‌های اولیه شامل ۳۰۳۷ داده و یک نمونه شامل ۴۰۶ داده جدید، دریافت شده از پژوهشکده معتمد جهاد دانشگاهی مورد استفاده قرار گرفتند. از آنجایی که تعدادی از داده‌های مورد استفاده در آموزش انفیس در تست انفیس نیز استفاده شدند این بررسی ارزیابی داخلی محسوب می‌شود. پس از ساخت مدل‌های فازی ممدانی و تاکاگی سوگنو و نیز انفیس‌ها با کدنویسی در نرم‌افزار Matlab، برای ارزیابی کارایی و بازدهی سیستم‌های ایجاد شده از شاخص‌های آماری استفاده شد.

نتایج

با توجه به داده‌های موجود بیماران، تمامی این بیماران زن و بزرگسال بودند و تحت درمان بوده یا پروسه درمانی خود را به پایان رسانده‌اند (مرگ یا سلامت). سن تشخیص بیماری در ۶۲/۴ درصد از این بیماران در بازه سنی ۳۰ تا ۵۰ سال قرار گرفته، ۳۳/۵۲ درصد از آن‌ها بیش از پنجاه سال داشته‌اند و تنها ۴/۰۴ درصد از بیماران در بازه سنی ۲۴ تا ۲۹ سال قرار داشته‌اند. در متغیر سن یائسگی ۱۰/۲۱ درصد، ۸۳/۶۲ درصد و ۶/۱۶ درصد از بیماران به ترتیب در سنین یائسگی زودرس، یائسگی عادی و یائسگی دیررس قرار داشته‌اند. بیماران دارای مراحل سرعت و قدرت رشد تومور (گرید) ۱، ۲ و ۳ بودند که در این میان بیماران گرید ۲ با ۷۳/۵۹ درصد بالاترین تعداد را شامل شدند. بیماران گرید ۳ با ۳۰/۰۵ درصد و بیماران گرید ۱ با ۱۰/۲۱ درصد مراتب بعدی را تشکیل دادند. مرحله بیماری سرطان (استیج) بیمارانی که داده‌های آن‌ها مورد استفاده قرار گرفتند، شامل ۳a، ۳b، ۳c و ۴ بودند که با توجه به حاصل شدن نتایج بهتر و دقیق‌تر برای مقادیر ۳، ۳/۳، ۳/۵ و ۴ آن‌ها جایگزین مقادیر اولیه استیج شد. ۵۱/۰۵ درصد از بیماران در استیج ۳، ۸/۲۸ درصد از بیماران در مرحله استیج ۳/۳، ۲۴/۴۷ درصد از بیماران در استیج ۳/۵ و ۱۶/۱۸ درصد از بیماران در استیج ۴ قرار داشتند.

جدول ۱: متغیرهای ورودی و خروجی

نام متغیر	بازه متغیر	نوع متغیر
سن یائسگی	۸-۹۰	ورودی
سن تشخیص بیماری	۲۴-۹۰	ورودی
داشتن یا نداشتن بارداری	۰-۱	ورودی
مصرف و یا عدم مصرف داروهای پیشگیری از بارداری	۰-۱	ورودی
سابقه شخصی ابتلا به سرطان پستان	۰-۴	ورودی
سابقه خانوادگی ابتلا به سرطان پستان	۰-۱	ورودی
مرحله سرعت و قدرت رشد تومور (گرید)	۰-۳	ورودی
مرحله بیماری سرطان (استیج)	۰-۴	خروجی

برای بررسی سیستم‌های استنتاج فازی ممدانی و تاکاگی سوگنو ساخته‌شده، با مقایسه مقادیر خروجی اصلی و مقادیر حاصله از مدل‌ها و استفاده از شاخص ریشه میانگین مربعات خط (RMSE) نتایج به صورت جدول ۲ به دست آمد.

همچنین نتایج حاصل از سیستم‌های انفیس مبتنی بر خوشه‌بندی میانگین مراکز فازی و نیز خوشه بندی کاهشی در جدول ۳ و نمودارهای ۱ و ۲ م قابل مشاهده است.

[Downloaded from jhbmi.ir on 2025-01-05]

جدول ۲: نتایج حاصل از سیستم‌های استنتاج فازی ممدانی و تاکاگی سوگنو

نام مدل	RMSE	STD
ممدانی	۰/۳۷۰۸۹۴۲۰۸	۰/۱۵۶۰۹۸۳۱۲
تاکاگی سوگنو	۰/۳۸۹۳۱۱۷۹۷	۰/۱۷۳۰۲۶۹۸۸

جدول ۳: نتایج حاصل از سیستم‌های انفیس مبتنی بر خوشه‌بندی میانگین مراکز فازی و نیز خوشه بندی کاهشی

شاخص آماری	انفیس مبتنی بر FCM	انفیس مبتنی بر خوشه‌بندی کاهشی
RSMEداده‌های آموزش	۰/۲۷۷۲۸	۰/۳۱۸۳۷
RSME داده‌های تست	۰/۳۴۸۷۴	۰/۳۶۵۱۵
STD داده‌های آموزش	۰/۲۷۵۵	۰/۳۱۸۶۸
STD داده‌های تست	۰/۳۴۸۹۶	۰/۳۶۵۲۹

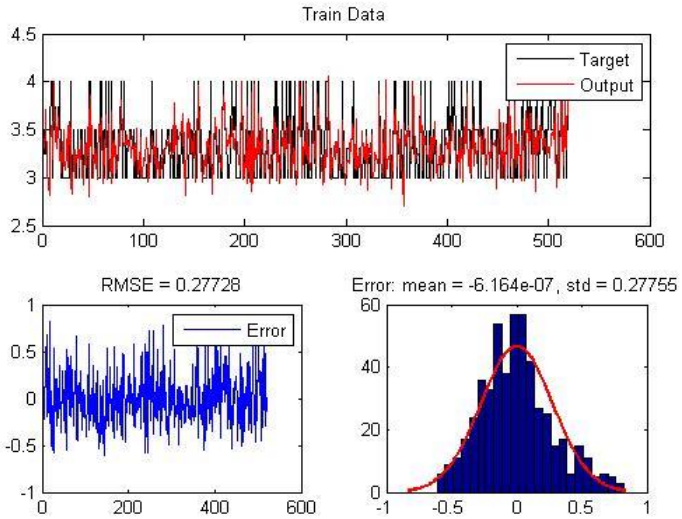
بحث و نتیجه‌گیری

در این مطالعه با استفاده از روش‌های ممدانی، تاکاگی سوگنو و انفیس با دو ساختار متفاوت برای سیستم استنتاج فازی اولیه و نیز داده‌های موجود در کلینیک بیماری‌های پستان، پژوهشکده معتمد جهاد دانشگاهی مدل‌هایی جهت پیش‌بینی متاستاز در بیماران با استیج‌های اولیه (حداکثر ۲) ایجاد شد. پیش‌بینی زودهنگام متاستاز سرطان پستان از جهات متعددی اهمیت دارد. برای نمونه، در مورد بیمارانی که عوامل خطر مشابهی با عوامل خطر بیماران درگیر با متاستاز دارند می‌توان با بررسی استیج احتمالی آن‌ها و در نظر گرفتن محدوده استیج‌های دارای متاستاز، تمهیدات لازم برای انتخاب بهترین شیوه درمان بیمار و جلوگیری از متاستاز در نظر گرفت. با بررسی جدول نتایج و فرآیند کلی ایجاد مدل‌ها این نکته مشخص شد که طراحی انفیس با سیستم استنتاج فازی اولیه ایجاد شده به روش میانگین مراکز فازی، نتایج دقیق‌تر و بهتری نسبت به سایر مدل‌های پیشنهادی ارائه می‌دهد. در انفیس قواعد به صورت خودکار و بدون نیاز به دانش فرد خبره تعیین می‌شوند؛ لذا در صورت عدم وجود دانش مورد نظر نیز با سنجش تمام نمونه‌ها بهترین ساختار را ارائه می‌دهد. انفیس ساخته شده با استفاده از الگوریتم کاهشی در مرتبه دوم قرار دارد. مدل‌های ممدانی و تاکاگی سوگنو نیز به ترتیب در مراتب بعدی قرار دارند. دلیل ضعف این دو مدل می‌تواند این باشد که همواره دانش افراد خبره در حیطه پزشکی و داده‌های مربوط به بیماران حقیقی از جامعیت کافی برخوردار نیستند، پس نتایج دقت قابل‌قبول و نه چندان دلخواهی دارند.

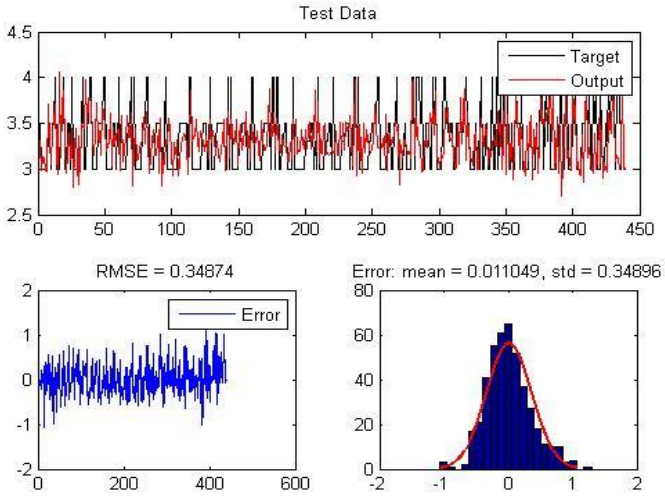
همان‌طور که بیان شد از آنجا که ایجاد متاستاز می‌تواند آغازی بر وخامت شدید حال بیمار باشد که به احتمال زیاد مرگ را به همراه دارد؛ لذا شناسایی و اطلاع از احتمال وقوع این عارضه می‌تواند به تصمیم‌گیری‌های در حوزه سرطان و در نظر گرفتن تدابیر لازم برای پیشگیری و مراقبت‌های لازم کمک نماید. مدل‌های پیش‌بینی از این حیث می‌توانند بسیار مفید باشند؛ اما باید در نظر داشت که در حوزه ارزیابی مدل‌های پیش‌بینی پزشکی باید به هر دو پارامتر ویژگی و حساسیت مدل توجه نمود، زیرا در نظر گرفتن یکی از آن‌ها به تنهایی می‌تواند گمراه کننده باشد. علاوه بر این باید به مقدار پارامتر منفی کاذب نیز توجه خاصی داشت. درصد این پارامتر در مدل‌های پیش‌بینی حوزه پزشکی بسیار حائز اهمیت است چون فرد بیمار به اشتباه، سالم در نظر گرفته می‌شود که می‌تواند عواقب بسیار خطرناکی داشته باشد. در مدل پیش‌بینی انفیس مبتنی بر میانگین مراکز فازی ارائه شده در این تحقیق این مقدار ۶/۲ درصد بود که مقدار نسبتاً کمی است و از این حیث می‌توان این مدل را قابل قبول دانست. روش‌های پیشنهاد شده در این مطالعه، به کمک سیستم‌های فازی مبتنی بر مجموعه‌ای از قوانین اگر – آنگاه فازی، دانش مورد نیاز را واکنشی نموده و عمل دسته‌بندی را بدین ترتیب انجام می‌دهند. قوانین فازی از آن جهت مورد توجه می‌باشند که امکان تفسیر آن‌ها توسط انسان خبره یا متخصص پزشکی آموزش دیده وجود دارد. در واقع دانش مد

[Downloaded from jhbmi.ir on 2025-01-05]

نظر ما می‌تواند به صورت یک پایگاه قوانین فازی در نظر گرفته شود که در طول فرایند داده کاوی و به کمک یک الگوریتم بهینه‌سازی با توجه به معیارهایی نظیر دقت و قابلیت تفسیر، بهینه می‌شود. انفیس مبتنی بر میانگین مراکز فازی ارائه شده دارای دو ویژگی اصلی سیستم‌های داده‌کاوی یعنی قابلیت اعتماد بالا و قابلیت تفسیر مناسب می‌باشد که از این نظر با الگوریتم‌های مشابه داده‌کاوی که معمولاً در پزشکی استفاده می‌شوند، قابل رقابت است [۱۵].



شکل ۱: نمایش نتایج انفیس مبتنی بر FCM برای داده‌های آموزش



شکل ۲: نمایش نتایج انفیس مبتنی بر FCM برای داده‌های تست

[Downloaded from jhbmi.ir on 2025-01-05]

مطالعات انجام شده در مورد پیش‌بینی‌های سرطان با استفاده از موتور فازی متفاوت هستند. در پیش‌بینی متاستاز، مطالعه Tez [۱۶] استفاده از استنتاج فازی را برای سرطان تیروئید و مطالعه Girometti و همکاران [۱۷] استفاده از استنتاج فازی را برای سرطان پانکراس موفقیت‌آمیز ارزیابی کرده‌اند؛ اما در مطالعه Nagata و همکاران [۱۸] برای بیماران مبتلا به سرطان زبان، سیستم استنتاج فازی علی‌رغم ویژگی قابل قبول بالای ۹۰ درصد، حساسیت نسبتاً پایینی معادل ۵۰ درصد را رقم زده است. در مطالعه‌ای مشابه مطالعه حاضر، Dehnavi و همکاران [۱۹] با استفاده از مارکرهای سلولی- ملکولی با موفقیت سیستمی مشابه سیستم پیشنهادی فعلی را آزمایش نموده‌اند که به دلیل عدم تشابه متغیرها و به تبع آن شاخص نهایی، امکان مقایسه بین این دو مطالعه وجود ندارد. در حوزه‌های پیش‌بینی دیگر سرطان پستان همانند تشخیص زودرس بیماری، کم و بیش در مطالعاتی خارجی یا داخلی از سیستم‌های فازی یا انفیس بهره گرفته شده است که بیشتر آن‌ها ارزش تست مدل برای حوزه‌های مهندسی را دارا هستند و جزء روش پیشنهادی ارزش بالینی خاصی ندارند [۲۲-۲۰، ۱۵]؛ تنها در یک مطالعه با رویکرد بالینی Atashi و همکاران [۲۳] مدل انفیس را برای تشخیص زود هنگام سرطان به ترتیب با حساسیت و ویژگی ۸۹ و ۷۹ درصد معرفی کرده‌اند که حداقل در مقایسه نتایج، مطالعه حاضر دارای شاخص‌های بهتری است.

از نقاط قوت مطالعه حاضر می‌توان به استفاده از مجموعه داده‌های واقعی بیماران اشاره کرد. داده‌های واقعی بیماران نسبت به داده‌های موجود مانند وینسکانسین مناسب‌تر، قابل اعتمادتر و نزدیک‌تر به واقعیت هستند. از طرفی نتایج حاصله با اطمینان بیشتری برای استفاده در تصمیم گیری‌های بالینی مجموعه‌های مورد بحث کاربردی می‌شوند. نقطه قوت دیگر این مطالعه وجود مجموعه جامعی از داده‌ها با تعداد بالای نمونه است که توان مطالعه را بالا می برد. همچنین در این مطالعه از کمک‌ها و نظرات تخصصی متخصصان سرطان برای کاربردی کردن هر چه بیشتر نتایج استفاده شده است.

از محدودیت‌های مطالعه، وجود مقادیر معتنابهی داده تهی و گمشده می‌باشد که جانشانی آن‌ها با روش‌های مرسوم از جمله روش انجام شده در مطالعه حاضر ممکن است در نتایج تأثیرگذار باشد. از طرفی، مجموعه‌های مورد استفاده با توجه به نظر پزشکان ایرانی تعریف شده‌اند که ممکن است برای خارج از ایران یا سایر مجموعه‌های بیماران مناسب و قابل تعمیم نباشند. با توجه به این‌که همواره با کمبود و محدودیت پایگاه داده‌های کامل مربوط به سرطان پستان روبه‌رو هستیم و دیگر این که همواره باید تعداد زوج داده‌ها بیشتر از تعداد پارامترها در شبکه تطبیقی باشند، استفاده از شبکه عصبی به دلیل کمتر بودن پارامترها و نیز الگوریتم ژنتیک احتمالاً می‌توانند نتایج را بهبود بخشند که استفاده از آن‌ها در مطالعات آتی پیشنهاد می‌شود.

به‌طورکلی از مدل‌های ساخته شده در این پژوهش و خصوصاً انفیس مبتنی بر روش خوشه‌بندی میانگین مراکز فازی می‌توان برای پیش‌بینی وقوع متاستاز در افراد مبتلا به سرطان پستان در مراحل اولیه (استیج‌های ۰ تا ۲) استفاده نمود. به این شکل که می‌توان با در نظر گرفتن داده‌های مربوط به بیماران در مراحل اولیه و وارد کردن آن‌ها در مدل‌های طراحی شده، استیج احتمالی در صورت درگیری با متاستاز را برای آن بیمار مشخص نمود. در این تشخیص مدل‌های پیشنهادی بر این اساس که بیماران درگیر با متاستاز و با شرایط و عوامل خطر مشابه در چه مرحله‌ای قرار داشته‌اند، استیج احتمالی را برای بیماران مراحل اولیه که اطلاعات و داده‌های آن‌ها در مدل وارد شده است، پیش بینی می‌کنند. بهترین مدل پیش‌بینی پیشنهادی، می‌تواند به عنوان یک سیستم پشتیبانی تصمیم گیری بالینی برای کمک به پزشکان در تصمیم‌گیری در مورد درمان بیماران مورد استفاده قرار گیرد.

تعارض منافع

در این مطالعه هیچ‌گونه تضاد منافعی وجود نداشت.

References

1. Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, et al. The global breast cancer burden: variations in epidemiology and survival. Clin Breast Cancer 2005;6(5):391-401. doi: 10.3816/cbc.2005.n.043

2. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55(2):74-108. doi: 10.3322/canjclin.55.2.74

3. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61(2):69-90. doi: 10.3322/caac.20107

4. National Breast Cancer Foundation. Early Detection [cited 2019 Jun 2]. National Breast Cancer Foundation,

[Downloaded from jhbmi.ir on 2025-01-05]

Inc. 2012. Available from: https://www.nationalbreastcancer.org/early-detection-of-breast-cancer/
5. Scully OJ, Bay BH, Yip G, Yu Y. Breast cancer metastasis. Cancer Genomics Proteomics 2012;9(5):311-20.
6. Lacroix M. Significance, detection and markers of disseminated breast cancer cells. Endocr Relat Cancer 2006;13(4):1033-67. doi: 10.1677/ERC-06-0001
7. Sun YF, Yang XR, Zhou J, Qiu SJ, Fan J, Xu Y. Circulating tumor cells: advances in detection methods, biological issues, and clinical relevance. J Cancer Res Clin Oncol 2011;137(8):1151-73. doi: 10.1007/s00432-011-0988-y
8. Castanho MJ, Hernandes F, De Ré AM, Rautenberg S, Billis A. Fuzzy expert system for predicting pathological stage of prostate cancer. Expert Systems with Applications 2013; 40(2): 466-70. https://doi.org/10.1016/j.eswa.2012.07.046
9. Bakker DP, Busscher HJ, van Zanten J, de Vries J, Klijnstra JW, van der Mei HC. Multiple linear regression analysis of bacterial deposition to polyurethane coatings after conditioning film formation in the marine environment. Microbiology (Reading) 2004;150(Pt 6):1779-84. doi: 10.1099/mic.0.26983-0
10. Ying H, Lin F, MacArthur RD, Cohn JA, Barth-Jones DC, Ye H, et al. A self-learning fuzzy discrete event system for HIV/AIDS treatment regimen selection. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 2007; 37(4): 966-79. doi: 10.1109/TSMCB.2007.895360
11. Barro S, Marín R. Fuzzy Logic in Medicine. eBook. Springer Science & Business Media; 2013.
12. LaBrunda M, LaBrunda A. Fuzzy logic in medicine. In: Khosrow-Pour M, editor. Breakthrough Discoveries in Information Technology Research: Advancing Trends: Advancing Trends. IGI Global; 2009. p. 218-24. doi: 10.4018/978-1-60566-966-3
13. Shleeg AA, Ellabib IM. Comparison of Mamdani and Sugeno fuzzy interference systems for the breast cancer risk. International Journal of Computer and Information Engineering 2013; 7(10): 387-91.
14. Saleh AA, Barakat SE, Awad AA. A fuzzy decision support system for management of breast cancer. International Journal of Advanced Computer Science and Applications 2011; 2(3): 34-40. doi: 10.14569/IJACSA.2011.020307
15. Einipour A, Abbasi Dezfouli M, Yektaie MH. Breast cancer detection using fuzzy classification system based on Ant-community algorithm. 2nd Iran Data Mining Conference; 2008 Nov 11-12; Tehran: Amirkabir University of Technology, Gita Data Mining Research Institute; 2008. [In Persian]
16. Tez M. Fuzzy Logic Can Improve Decision Making in Thyroid Cancer. Ann Surg 2010;252(4):706-7.
doi: 10.1097/SLA.0b013e3181f47aed
17. Girometti R, Fabris F, Sgarro A, Zanella G, Pullini S, Cereser L, et al. Differentiating Small (≤ 1 cm) Focal Liver Lesions as Metastases or Cysts by means of Computed Tomography: A Case-Study to Illustrate a Fuzzy Logic-Based Method to Assess the Impact of Diagnostic Confidence on Radiological Diagnosis. Comput Math Methods Med 2014;2014:587976. doi: 10.1155/2014/587976
18. Nagata T, Schmelzeisen R, Mattern D, Schwarzer G, Ohishi M. Application of fuzzy inference to European patients to predict cervical lymph node metastasis in carcinoma of the tongue. International Journal of Oral and Maxillofacial Surgery 2005; 34(2): 138-42. https://doi.org/10.1016/j.ijom.2004.03.012
19. Dehnavi AM, Sehhati MR, Rabbani H. Hybrid method for prediction of metastasis in breast cancer patients using gene expression signals. J Med Signals Sens 2013; 3(2): 79–86.
20. Hosseini R, Khezri R, Mazinani M. A fuzzy rule-based expert system for the prognosis of the risk of development of the breast cancer. International Journal of Engineering 2014; 27(10): 1557-64.
21. Alipoor M, Haddadnia J. An Accurate Intelligent Breast Cancer Diagnosis System. Iranian Journal of Breast Diseases 2009; 2(2): 33-40. [In Persian]
22. Mahdavi H. Detection of Breast Cancer Progress Using Adaptive Nero Fuzzy Inference System and Data Mining Techniques. Journal of Computer and Robotics 2013; 6(2): 23-8.
23. Atashi A, Nazeri N, Abbasi E, Dorri S, Alijani Z M. Breast cancer risk assessment using adaptive neuro-fuzzy inference system (ANFIS) and subtractive clustering algorithm. Multidiscip Cancer Investig 2017; 1(2): 20-6. doi: 10.21859/mci-01029

[Downloaded from jhbmi.ir on 2025-01-05]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2020; 7(2): 181-189

Prediction of Breast Cancer Metastasis Using Fuzzy Models based on Data from Iranian Breast Cancer Patients

HomaieFasih Arezou[1], Atashi Alireza[2,3*], Ahsani Tehrani Hojjat[4]

• **Received:** 18 Oct 2019 • **Accepted:** 18 Apr 2020

Introduction: The metastasis of breast cancer, the spread of cancer to different body parts, is considered as one of the most important factors responsible for the majority of deaths caused by breast cancer in women. Diagnosing the breast cancer metastasis at the earliest stages helps to choose the best treatment and improve the quality of life for patients.

Method: In the present fundamental research, the dataset of Iranian patients available at Breast Cancer Research Center of Motamed Cancer Institute in Tehran was utilized. This study used Mamdani fuzzy inference system, Takagi-Sugeno fuzzy inference system and adaptive neuro-fuzzy inference system (ANFIS) to predict breast cancer metastasis at early stages.

Results: The best prediction error was obtained using adaptive neuro-fuzzy inference system based on fuzzy c-means approach. The opinion of the experts at Breast Cancer Research Center of Motamed Cancer Institute and the prediction error of the assessed model indicated that this prediction system is well-formed.

Conclusion: The optimal proposed prediction system can be used as a clinical decision support system to assist medical practitioners in the healthcare practice.

Keywords: Breast Cancer, Metastasis, Mamdani Fuzzy Inference System, Takagi-Sugeno Fuzzy Inference System, Adaptive Neuro-Fuzzy Inference System

• **Citation:** HomaieFasih A, Atashi A, Ahsani Tehrani H .Prediction of Breast Cancer Metastasis Using Fuzzy Models based on Data from Iranian Breast Cancer Patients. Journal of Health and Biomedical Informatics 2020; 7(2): 181-9. [In Persian]

1. M.Sc. in Applied Mathematics-Numerical Analysis, School of Mathematical Sciences, Shahroud University of Technology, Shahroud, Iran
2. Ph.D. in Medical Informatics, Medical Informatics Research Group, Clinical Research Dept., Breast Cancer Research Center, Motamed Cancer Institute, ACECR, Tehran, Iran
3. Ph.D. in Medical Informatics, Assistant Professor, e-Health Dept., Virtual School, Tehran University of Medical Sciences, Tehran, Iran
4. Ph.D. in Applied Mathematics, Associate Professor, School of Mathematical Sciences, Shahroud University of Technology, Shahroud, Iran

***Corresponding Author:** Alireza Atashi
Address: No. 2, Dolatshahi Alley, Naderi St., Keshavarz Blvd., Tehran, Iran
• **Tel:** 02142036000 • **Email:** aratashi@sina.tums.ac.ir

[Downloaded from jhbmi.ir on 2025-01-05]