

تعیین پارامترهای منطق فازی با بکارگیری الگوریتم ژنتیک برای تشخیص بیماری کبد

مریم رضایی فرخزاد^۱، فرهاد سلیمانیان قره‌چپق^{۲*}

• پذیرش مقاله: ۹۷/۷/۷

• دریافت مقاله: ۹۶/۹/۱۸

مقدمه: بیماری کبد یکی از شایع‌ترین مشکلات مزمن کبدی و سیروز می‌باشد. مشکلات کبدی شامل طیف وسیعی از بیماری‌ها و نارسایی‌هایی هستند که به بافت کبد یا عملکرد آن آسیب می‌رسانند. تشخیص زودهنگام و درمان این بیماری می‌تواند مرگ‌ومیر و درجه بیماری را کاهش دهد.

روش: این مطالعه از نوع توصیفی-تحلیلی است. پایگاه داده آن شامل ۵۸۳ رکورد مستقل شامل ۱۱ ویژگی موجود در پایگاه داده یادگیری ماشین UCI می‌باشد. در این مقاله با استفاده از منطق فازی که پارامترهای آن با الگوریتم ژنتیک تعیین می‌شوند روشی برای تشخیص بیماری کبد پیشنهاد شد. برای این منظور، ابتدا ویژگی‌های مجموعه داده با استفاده از ویژگی آنالیز رتبه‌بندی شد و سپس بهینه‌سازی داده‌های مجموعه داده با استفاده از الگوریتم ژنتیک انجام گرفت. در نهایت بیماری کبد با استفاده از سیستم استنتاج فازی genfis2 و genfis3 تشخیص داده شد.

نتایج: نتایج نشان داد که دقت تشخیص بیماری کبد با استفاده از سیستم استنتاج فازی genfis2 با هشت ویژگی برابر با ۹۱/۶۶ درصد است و با سیستم استنتاج فازی genfis3 با شش ویژگی برابر با ۸۹/۸۷ درصد است و همچنین خطای genfis2 و genfis3 به ترتیب برابر با ۰/۰۳۴ و ۰/۰۴۷ کمتر است.

نتیجه‌گیری: بیماری کبد یکی از شایع‌ترین بیماری‌ها در بین افراد جامعه است. تشخیص به موقع بیماری ضمن کاهش هزینه‌ها، می‌تواند شانس درمان موفقیت‌آمیز بیمار را افزایش دهد. با توجه به نتایج به دست آمده، مشاهده شد که مدل پیشنهادی با دقت نسبتاً بالایی، افراد مبتلا به بیماری کبد را تشخیص داد.

کلید واژه‌ها: تشخیص بیماری کبد، منطق فازی، سیستم استنتاج فازی، الگوریتم ژنتیک

• **ارجاع:** رضایی فرخزاد مریم، سلیمانیان قره‌چپق فرهاد. تعیین پارامترهای منطق فازی با بکارگیری الگوریتم ژنتیک برای تشخیص بیماری کبد. مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۷؛ ۵(۳): ۳۹۷-۳۸۴.

۱. کارشناسی ارشد مهندسی کامپیوتر، گروه مهندسی کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران

۲. دکترای مهندسی کامپیوتر، استادیار، گروه مهندسی کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران

* **نویسنده مسئول:** ارومیه، کیلومتر دو جاده فرودگاه، دانشگاه آزاد اسلامی ارومیه، دانشکده فنی و مهندسی

• **Email:** bonab.farhad@gmail.com

• **شماره تماس:** ۰۹۱۴۱۷۶۴۴۲۷

مقدمه

کبد یکی از اعضای مهم بدن است که سم زدایی از داروها، دفع مواد زائد ناشی از تخریب و نوسازی گلبول های قرمز خون به صورت صفرا، تولید عوامل انعقادی خون، ذخیره قند به صورت گلیکوژن و نیز تنظیم سوخت و ساز قند و چربی از مهم ترین نقش های آن در بدن می باشد؛ البته نقش کبد در جذب چربی و دفاع در مقابل میکروب ها و سموم جذب شده از راه مواد غذایی را نیز نباید نادیده گرفت [۱]. در سال های اخیر آمار مرگ و میر ناشی از اختلالات کبدی به شدت افزایش یافته است. تشخیص به موقع این بیماری می تواند در پیشگیری از عوارض آن و نیز کنترل و درمان بیماری مؤثر واقع شود [۲۴]. Brawes یکی از مهم ترین مشکلات در تشخیص بیماری را تشخیص نادرست و دیر هنگام پزشکان می داند؛ ممکن است از سوی پزشک در حوزه تشخیص و درمان خطایی رخ دهد که منجر به صدمات جبران ناپذیری گردد [۲، ۳].

یکی از برنامه های مهم انفورماتیک پزشکی، استفاده از سیستم های خبره برای تشخیص بیماری براساس مجموعه ای از علائم می باشد [۳، ۴]. این سیستم ها برپایه هوش مصنوعی هستند که با بررسی تست های آزمایشگاهی، به پزشکان در تشخیص بیماری و درمان مؤثرتر کمک می کنند و هزینه، زمان، تخصص انسان و تشخیص نادرست را به حداقل می رسانند [۵]. در صورت ایجاد مدلی به منظور پیش بینی و تشخیص زودهنگام بیماری با دقت، ویژگی و حساسیت بالا می توان به پزشکان در تشخیص به موقع و انجام اقدامات درمانی مناسب در مراحل اولیه کمک کرد و بدین صورت مانع از پیشرفت بیماری شد [۶].

هدف از این مطالعه، تشخیص بیماری کبدی با استفاده از ترکیب منطق فازی و الگوریتم ژنتیک بود و پارامترهای منطق فازی با استفاده از الگوریتم ژنتیک تعیین می شوند. هدف این مطالعه از انتخاب منطق فازی داشتن مزایایی همچون مفهوم بسیار ساده، انعطاف پذیری بالا، توان تحمل داده های غیردقیق به شکل مطلوب، قابلیت مدل سازی توابع غیرخطی پیچیده، عمل کردن بر مبنای تجربه متخصصان، قابل ادغام با تکنیک های کنترلی مرسوم، بر مبنای زبان طبیعی بودن، می باشد. الگوریتم ژنتیک نیز که جستجوی قوی در فضاهای پیچیده فراهم می کند، برای دست یافتن به عملکرد بهتر منطق فازی استفاده شد.

با مطالعه پژوهش ها مشاهده شد که محققان بسیاری به تشخیص بیماری با استفاده از روش های مختلف پرداخته اند و

هر کدام توانسته اند به دقت های قابل قبولی دست یابند. Rathika و Rajamani به مروری بر سرطان کبد با استفاده از تکنیک داده کاوی در سیستم استنتاج عصبی-فازی تطبیقی پرداخته اند. از تصاویر دو بعدی سی تی اسکن به عنوان ورودی استفاده کرده و در گام پردازش داده ها، به حذف نویز در تصاویر سی تی اسکن، فرآیند تقسیم بندی، عملیات مورفولوژیکی و تکنیک های استخراج ویژگی پرداخته اند. همچنین از سیستم استنتاج عصبی-فازی تطبیقی در تشخیص زود هنگام سرطان کبد در انسان بحث کرده اند و امیدوار هستند که روش ارائه شده و یا ترکیبی از سیستم استنتاج عصبی-فازی تطبیقی با دیگر تکنیک های داده کاوی به حوزه پزشکی در تشخیص زود هنگام سرطان کبد کمک نماید [۷].

Hashmi و Saleem Khanb در مقاله خود تحت عنوان «بررسی تست خون برای بیماری کبدی با استفاده از منطق فازی»، رایج ترین روش های آزمایشگاهی برای سنجش سلامت عمومی و سلامت کبد را بررسی و کارکرد کبد، شمارش کامل خون و عکس بیان نموده اند. ورودی های منطق فازی را گلبول سفیدخون، هماتوکریت، هموگلوبین و تعداد پلاکت و خروجی را رفع عفونت، کم خونی و ترومبوسیتوز (Thrombocytosis) در نظر گرفته اند. مدل فازی پیشنهادی را به چهار مرحله فازی سازی، موتور استنتاج و انتخاب قوانین و غیرفازی سازی تقسیم نموده اند [۸]. Obayya و Areed با استفاده از سیستم استنتاج عصبی-فازی تطبیقی به طبقه بندی تومور کبدی به دو دسته خوش خیم و بدخیم با آنالیز تصاویر سی تی اسکن کبد پرداخته اند. به منظور بررسی اثر هر نوع از ویژگی های تومور بر فرآیند طبقه بندی و تصمیم گیری درست بر طبقه بندی تومور خوش خیم و بدخیم، این دو مجموعه از ویژگی ها را به صورت جداگانه آموزش داده اند. نتایج نشان می دهد که روش پیشنهادی دارای توان بالقوه ای در شناسایی نوع تومور دارد [۹]. Fan و همکاران مدل ترکیبی استدلال موردی و درخت تصمیم گیری فازی را برای تشخیص بیماری کبدی پیشنهاد کرده و به دقت ۸۱/۶ درصد دست یافته اند [۱۰].

Satarkar و Ali، سیروز را یکی شایع ترین بیماری های کبدی عنوان نموده [۱۱] و سیستم خبره فازی را برای تشخیص این نوع بیماری معرفی نموده اند. سیستم پیشنهادی دارای سه ورودی (نرخ (Aspartame) Aminotransferase/Alanime Aminotransferase International, AST/ALT، تعداد پلاکت و)

تلاش شد که دقت تشخیص بیماری کبد نسبت به مطالعات مشابه افزایش یابد.

روش

مجموعه داده استفاده شده شامل ۱۶۷ بیمار غیرکبدی (سالم) و ۴۱۶ بیماری کبدی که در شمال شرق آندراپرادش هند از ۴۴۱ مرد و ۱۴۲ زن جمع‌آوری شده است [۱۷]. بیماران بالای ۸۹ تا ۹۰ سال فرض شده‌اند. در کل مجموعه داده فوق دارای ۱۱ ویژگی و ۵۸۳ نمونه است. ویژگی‌های این مجموعه داده عبارت‌اند از: سن بیمار، جنسیت بیمار، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز یا آلکالین فسفاتاز، آمینوترانسفراز آلانین، آسپارات ترانس آمیناز، مجموع پروتئین‌ها، آلبومین، نسبت آلبومین و نسبت گلوبولین است. کل ویژگی‌ها عددی هستند و هیچ داده گم‌شده‌ای موجود نمی‌باشد. این پایگاه داده ۱۰ ویژگی خام دارد و تمامی آزمایش‌ها بر روی این ۱۰ ویژگی انجام شده است؛ بنابراین این پایگاه داده شامل ۱۰ علامت بیماری و یک ویژگی تشخیص است که فیلد هدف به وجود بیماری کبدی بر اساس علائم موجود در بیمار اشاره دارد که یک مقدار عددی (۲ عدم ابتلا به بیماری کبدی) و (۱ ابتلا به بیماری کبدی) است.

داده‌کاوی یا کشف دانش از میان پایگاه‌های اطلاعاتی علمی انجام می‌گیرد که برای تصمیم‌گیری‌های هوشمندانه باتوجه به پیشرفت روزافزون تکنولوژی اطلاعات بسیار کاربرد دارد. تکنیک‌های داده‌کاوی برای پیدا کردن الگوهای جالب در تشخیص پزشکی و درمان مورد استفاده است [۱۶]. انواع مختلفی از تکنیک‌های داده‌کاوی در استخراج اطلاعات که می‌تواند برای پیش‌بینی مورد استفاده قرار گیرد وجود دارد که آن جمله می‌توان به منطق فازی اشاره کرد. تئوری مجموعه‌های فازی، یک روش نظام‌مند محاسباتی است که این تئوری بر مبنای محاسبات عددی بر روی مقادیر تولیدشده توسط تابع عضویت برای هریک از متغیرهای زبان‌شناختی عمل می‌نماید. علاوه بر آن، انتخاب قواعد if-then فازی، جزء اصلی سیستم استنتاج فازی را تشکیل می‌دهد. به کمک این قواعد می‌توان به شکل مؤثری تخصص یک انسان را در یک زمینه خاص مدل نماید. براین اساس در این مقاله به ارائه روشی جهت تشخیص بهتر بیماری کبد پرداخته می‌شود. فلوجارت روش پیشنهادی در شکل ۱ نشان داده شد.

برای رتبه‌بندی ویژگی‌ها براساس ویژگی آنتروپی، ابتدا از تابع Rank Features در نرم‌افزار Matlab نسخه ۲۰۱۷ استفاده

INR (Normalized Ratio) و یک خروجی (خطر ابتلا به سیروز که با اعداد ۰ تا ۶ مشخص شده) می‌باشد. در سیستم پیشنهادی، از روش استنتاج مددانی استفاده شده است. Nasira و Kalaiselvi، در پژوهش خود رویکرد جدیدی برای تشخیص سرطان کبد و دیابت با استفاده از سیستم استنتاج عصبی-فازی تطبیقی و k نزدیک‌ترین همسایه بهبود یافته، ارائه نموده‌اند. ایشان از مجموعه داده دیابت Pima هند به‌عنوان ورودی و طبقه‌بندی براساس اطلاعات ویژگی‌ها را به‌عنوان خروجی در نظر گرفته‌اند. نتایج نشان می‌دهد که دقت طبقه‌بندی در روش پیشنهادی بهتر از روش‌های دیگر است [۱۲].

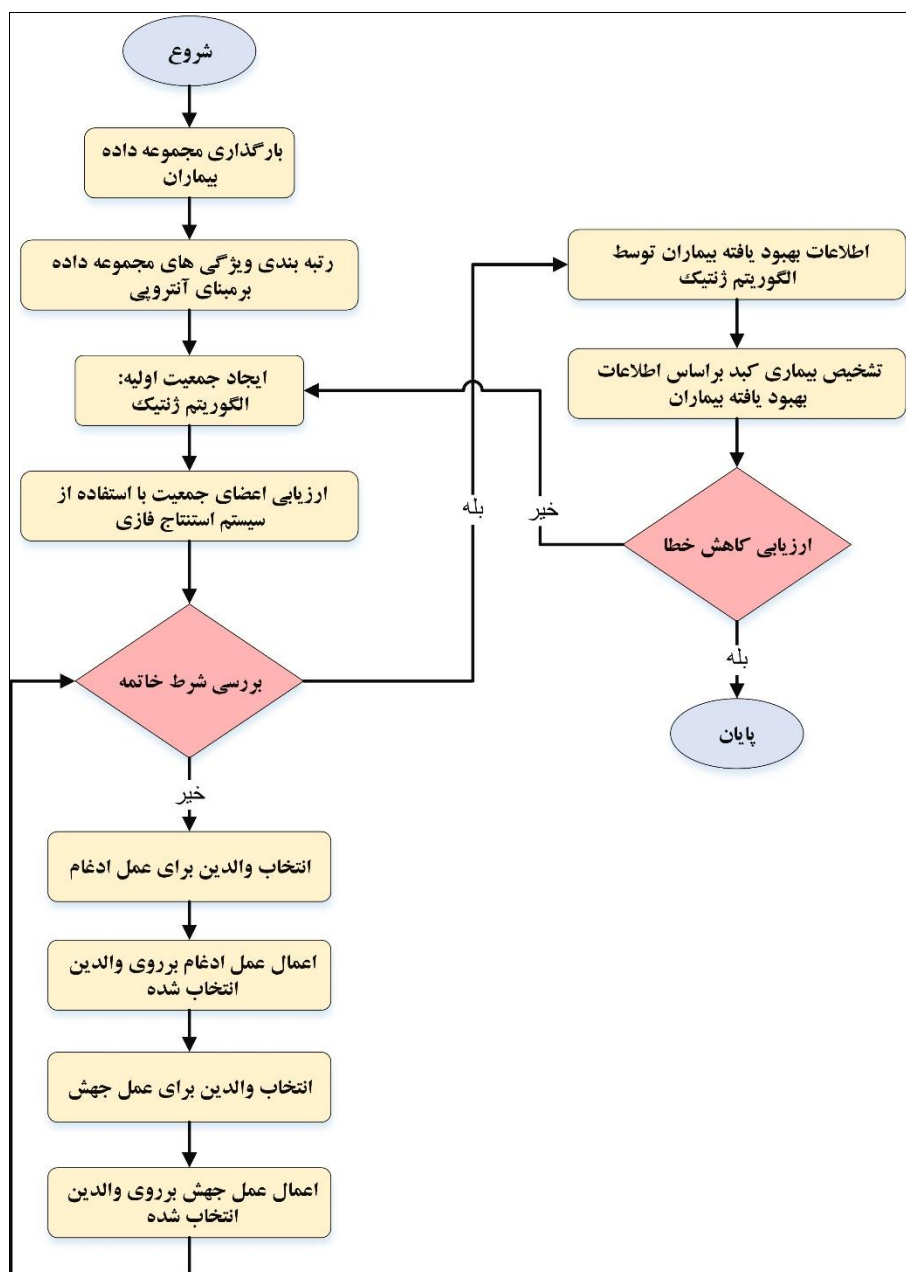
نشاط و یعقوبی، در پژوهشی تحت عنوان «طراحی یک سیستم فازی تشخیص درجه شدت هیپاتیت B و مقایسه آن با سیستم فازی شبکه عصبی تطبیقی»، از مجموعه داده‌های جمع‌آوری شده در بخش بیوپسی (نمونه‌برداری) کبد بیمارستان امام‌رضا (ع) مشهد ایران که شامل ۳۰۰ رکورد و هر رکورد دارای ۷ فیلد می‌باشد، استفاده کرده‌اند. قوانین استنتاج قوی فازی، با کمک تحقیق و تجربه فرد متخصص مشخص شده است. دقت سیستم پیشنهادی در تشخیص شدت هیپاتیت 0.96 ± 0.02 درصد به‌دست آمده است [۱۳].

Olaniyi و Khashman [۱۴] به تشخیص بیماری کبد با استفاده از شبکه‌های عصبی مصنوعی پس انتشار و تابع پایه شعاعی پرداخته‌اند. ایشان از مجموعه داده‌های BUPA (برگرفته از دانشگاه کالیفرنیا در ایروین) استفاده نموده و با ۶ ورودی و ۲ خروجی با استفاده از شبکه عصبی پس انتشار (با ۱۰ نرون نهان) و شبکه عصبی پایه شعاعی (با ۲۰ نرون نهان) به ترتیب به دقت ۶۳ درصد و ۷۰ درصد دست یافته‌اند حال آن که بهرامی‌راد و همکاران [۱۵]، برای طبقه‌بندی بیماری کبدی، به‌ترتیب از دو مجموعه داده AP (Andhra Pradesh) و BUPA استفاده نموده و به دقت ۷۳/۳۹ و ۶۷/۳۹ دست یافته‌اند.

اگرچه تحقیقات مربوط به بیماری کبد بیشتر بالینی و بیولوژیکی است؛ اما استفاده از تحقیقات آماری به یک امر رایج تبدیل شده است. توجه بیشتر به ویژگی‌های بالینی و بیولوژیکی بیماری کبد و تشخیص دقیق‌تر این بیماری حتی به اندازه یک درصد باعث خواهد شد که شانس بقاء بیمار افزایش یافته و روند دوره درمان با سرعت بیشتری طی شود؛ لذا در این مطالعه که با هدف تشخیص بیماری کبد صورت گرفته است، با استفاده از ترکیب الگوریتم ژنتیک و سیستم استنتاج فازی

به عنوان فاصله Kullback-Leibler یا واگرایی نیز شناخته می شود. خروجی تابع Features Rank برابر Index است که ویژگی های مجموعه داده را رتبه بندی نموده است و براساس اولویت به صورت نزولی نشان می دهد.

شد. با استفاده از این تابع، ویژگی های موجود در مجموعه داده گروه بندی شد؛ که در آن Inputs یک ماتریس 11×583 ، Outputs یک ماتریس 1×583 از خروجی مجموعه داده می باشد که هر دو ترا نهاده شده اند. آنتروپی معیار تحلیل مؤلفه اساسی استفاده شده در این مقاله است. آنتروپی نسبی همچنین



شکل ۱: فلوچارت روش پیشنهادی

بالا بردن عملکرد آموزش می باشد. انتخاب مناسب روش نرمال سازی به نوبه خود کارایی سیستم استنتاج فازی را بالا می برد. برای نرمال سازی از روش نرمال سازی "min_max" استفاده شد که این روش دوباره مقیاس بندی

فرآیند نرمال سازی برای داده های اولیه، تأثیر بسزایی بر روی آماده سازی دارد تا آموزش بهتری انجام گیرد. بدون نرمال سازی آموزش سیستم استنتاج فازی بسیار کند صورت می گیرد [۱۸، ۱۹]. در واقع، هدف از نرمال سازی (پیش پردازش) داده ها،

انجام می‌دهد. به‌طوری که ویژگی‌ها یا خروجی‌ها را از یک محدوده مقادیر به محدوده مقادیر جدید انتقال می‌دهد. اغلب محدوده مقادیر بین ۰ و ۱ یا بین ۱- تا ۱ می‌باشد؛ که در این

مطالعه محدوده بین ۱- و ۱ در نظر گرفته شد. فرمول نرمال‌سازی "min_max" به‌صورت رابطه ۱ می‌باشد.

$$x' = (newx_{max} - newx_{min}) \times \frac{(x_i - x_{min})}{(x_{max} - x_{min})} + newx_{min} \quad (1)$$

که در رابطه ۱، x_{max} حداکثر مقدار ستون فعلی، x_{min} حداقل مقدار ستون فعلی و x_i محتویات خانه نام از ستون فعلی است. در این مقاله، داده‌ها به‌صورت ستونی نرمال‌سازی شد که با ایجاد یک حلقه for در هر مرحله یکی از ستون‌ها (یکی از پارامترهای نمونه) نرمال‌سازی شد.

الگوریتم ژنتیک [۲۰] دارای چندین پارامتر است که قبل از اجرای آن، باید مقداردهی گردند. اولین پارامترها که باید مقداردهی شوند؛ حد پایین و بالای ژن‌ها می‌باشد که به‌ترتیب آن‌ها را برابر کمترین و بیشترین مقدار مجموعه داده (۱- و ۱)،

در نظر گرفته شد. حداکثر تعداد تکرار حلقه الگوریتم برابر ۲۰۰ و تعداد جمعیت برابر تعداد رکوردهای مجموعه داده می‌باشد. درصد ادغام و جهش برابر ۰/۵ در نظر گرفته شد. تعداد والد‌های شرکت‌کننده در عمل ادغام و جهش نیز براساس روابط ۲ و ۳ محاسبه شد. در الگوریتم ژنتیک باید ترکیب و جهش انجام شود. باید مقدار پارامترهای pc و pm که به ترتیب نشان دهنده احتمال ادغام و احتمال جهش هستند، تعیین شوند. در این مقاله مقدار آن‌ها برابر با $pc=pm=0/5$ لحاظ شد.

$$nCrossover = round(pc * nPop / 2) * 2; \quad (2)$$

$$nMutation = round(pm * nPop); \quad (3)$$

که در روابط ۲ و ۳ pc درصد ادغام، nPop تعداد جمعیت اصلی و pm درصد جهش می‌باشد. اولین مرحله در الگوریتم ژنتیک، ایجاد جمعیت اولیه از کروموزوم‌ها است. برای تولید جمعیت اولیه، بیشتر از روش تولید تصادفی کروموزوم‌ها استفاده شد. تنوع کروموزوم‌ها در روش تصادفی بالا است، چرا که کروموزوم‌ها متعلق به نواحی مختلف فضای جواب هستند. در نتیجه در تکرارهای اولیه الگوریتم، تکامل نسل‌ها سریع‌تر انجام شد با افزایش تکرار، تشابه کروموزوم‌ها نیز افزایش یافت تا این‌که در نهایت الگوریتم به یک یا چند حل، شاخص همگرا شود. در این مقاله، نیز از روش تصادفی برای انتخاب کروموزوم‌ها از مجموعه داده بیماران کبدی استفاده شد. بدین‌صورت که به‌صورت تصادفی مقادیر nPop بیمار به‌عنوان جمعیت اولیه انتخاب شد.

برای ارزیابی اعضای جمعیت از سیستم استنتاج فازی (دو تابع genfis2 و genfis3) استفاده شد. هر سیستم استنتاج فازی شامل توابع مختلف فازی برای طبقه‌بندی نمونه‌ها است [۲۱]. هر دو تابع یک مجموعه داده آموزشی را به‌عنوان ورودی دریافت کرده و برطبق آن یک سیستم استنتاج فازی اولیه می‌سازند. ایجاد یک مدل اولیه برای توابع فازی می‌تواند در

طبقه‌بندی و شناسایی هرچه بهتر داده‌ها مؤثر باشد، زیرا هرچه دقت مدل اولیه بیشتر باشد، دقت قوانین نهایی فازی استخراج شده از سیستم استنتاج فازی نیز بیشتر می‌شود؛ لذا در این مقاله، روند تعیین پارامترهای سیستم استنتاج فازی توسط الگوریتم ژنتیک انجام گرفت. در سیستم استنتاج فازی از الگوریتم ژنتیک برای تنظیم پارامترهای توابع عضویت بهره گرفته شد و از میانگین مربعات خطا به منظور تنظیم مقدار بهینه برای پارامترهای مربوط به توابع عضویت خروجی استفاده شد.

genfis2: این تابع یک ساختار سیستم استنتاج فازی از نوع سوگنو [۲۲] با استفاده از خوشه‌بندی کاهشی ایجاد می‌کند. برای استخراج قوانین، ابتدا از تابع subclus برای تعیین تعداد قوانین و توابع عضویت اولیه استفاده کرده و سپس از میانگین مربع خطا برای تعیین معادلات مربوط به هر قانون استفاده می‌کند. این تابع یک ساختار سیستم استنتاج فازی را تحلیل می‌کند که شامل مجموعه‌ای از قوانین فازی برای پوشش فضای ویژگی است. مقادیر این تابع در جدول ۱ قابل مشاهده است. به‌طور پیش‌فرض نوع تابع عضویت ورودی و خروجی به‌ترتیب ناقوس شکل و خطی است.

جدول ۱: مقادیر تابع genfis2

Name	genfis2
Type	sugeno
And Method	prod
OR Method	probor
Defuzzification Method	wtaver
Implication Method	prod
Aggregation Method	max

genfis3: این تابع یک ساختار سیستم استنتاج فازی از نوع سوگنو با استفاده از خوشه‌بندی فازی C-means برای استخراج یک مجموعه از قوانین که رفتار داده‌ها را مدل می‌کنند، ایجاد می‌کند، مقادیر این تابع در جدول ۲ قابل

مشاهده است. به‌طور پیش‌فرض نوع تابع عضویت ورودی و خروجی در این تابع به ترتیب گاوسی و خطی است؛ اما اگر نوع سیستم استنتاج فازی، ممدانی [۲۳] در نظر گرفته شود، تابع عضویت خروجی گاوسی خواهد بود.

جدول ۲: مقادیر تابع genfis3

Name	genfis3
Type	Sugeno/ Mamdani
And Method	Prod
OR Method	probor
Defuzzification Method	wtaver
Implication Method	Prod
Aggregation Method	Sum

یک بخش دیگر الگوریتم ژنتیک، ایجاد کروموزوم‌های فرزندان از طریق کروموزوم‌های والدین است. عملگر ادغام عملگری است که در یک لحظه بر روی یک یا چندین کروموزوم والد اعمال می‌شود و یک یا چند فرزند را ایجاد می‌کند. کارایی این

عملگر در رسیدن به جواب بهینه در مسائل مختلف، متفاوت است. عملگر ادغام بر روی دو کروموزوم انتخاب‌شده به‌صورت تصادفی به‌صورت روابط ۴ و ۵ اعمال شد.

$$C_{pop1} = \alpha * pop1 + (1 - \alpha) * pop2 \quad (4)$$

$$C_{pop2} = \alpha * pop2 + (1 - \alpha) * pop1 \quad (5)$$

که در روابط ۴ و ۵ $pop1$ و $pop2$ دو والد که به‌صورت تصادفی انتخاب شده‌اند. مقدار α نیز به‌صورت معادله ۶ به‌دست

آمد.

$$\alpha = \text{unifrnd}(-d, 1+d, \text{size}(pop1)) \quad (6)$$

رابطه ۶ در بازه بین $-d$ تا $d+1$ به تعداد size والد اعداد تصادفی تولید می‌کند. هدف از جهش، بیان یک خاصیت ژنتیکی است که باعث افزایش تنوع در جواب‌های جمعیت می‌شود. جهش به‌عنوان یک عملگر الگوریتم ژنتیک، در کنار عملگر ادغام قرار می‌گیرد. به‌عبارت دیگر، این عملگر در

کروموزوم‌های متفاوت تغییرات تصادفی برنامه‌ریزی نشده ایجاد می‌کند و ژن‌هایی را که در جمعیت اولیه وجود نداشته‌اند را وارد جمعیت می‌کند. برای محاسبه عملگر جهش از رابطه ۷ استفاده شد.

$$\beta = (\text{VarMax} - \text{VarMin}) / 10 \quad (7)$$

$$n\text{Pap}(j) = \text{pop}(j) + \beta \times b \quad (8)$$

نتایج

برای تشخیص این که چه تعداد از ویژگی‌های بیماری کبد در تشخیص بهتر بیماری کبد مؤثر هستند، از روش سعی و خطا استفاده شد؛ که نتایج آن در جدول ۴ و ۵ قابل مشاهده است. همچنین تابع برازندگی برای طبقه‌بندی نمونه‌ها طبق معادله ۹ تعریف شد.

که در رابطه ۷، VarMax و VarMin به ترتیب برابر کمترین و بیشترین مقدار مجموعه داده یعنی ۱- و ۱ تولیدشده در بازه بین صفر و یک است. طبق رابطه ۸، جمعیت جدید ایجاد می‌شود که پارامتر b یک عدد تصادفی در بازه ۰ و ۱ است. جمعیت‌های اولیه، ادغام و جهش‌یافته باهم ادغام شدند و پس از مرتب‌سازی آن‌ها براساس کمترین تفاضل خطا، nPop جمعیت اولیه به عنوان جمعیت جدید انتخاب شد.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (9)$$

مثبت بوده و الگوریتم طبقه‌بندی دسته آن‌ها را به اشتباه منفی تشخیص داده است.

جدول ۴، نشان می‌دهد که اگر تعداد ویژگی‌ها برابر ۸ ویژگی باشد میانگین مربعات خطا کمتر است؛ لذا درصد صحت بیشتر است. کاهش بعد یکی از مهم‌ترین تکنیک‌های پیش پردازش داده است. در کاهش بعد داده مزایای زیادی وجود دارد. برای مثال با کاهش بعد، ویژگی‌های نامرتبط حذف می‌شوند، نویز کاهش می‌یابد و در نتیجه مدل بهتری ساخته می‌شود.

پارامتر TN بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم طبقه‌بندی نیز دسته آن‌ها را به درستی منفی تشخیص داده است.

TP: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها مثبت بوده و الگوریتم طبقه‌بندی نیز دسته آن‌ها را به درستی مثبت تشخیص داده است.

FP: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم طبقه‌بندی دسته آن‌ها را به اشتباه مثبت تشخیص داده است.

FN: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها

جدول ۴: نتایج پارامترها براساس genfis2

تعداد ویژگی	ویژگی‌ها	میانگین مربع خطا
۳	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین و مجموع بیلی روبین	۰/۰۲۸۲
۴	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین و بیلی روبین مستقیم	۰/۰۳۱۰
۵	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم و آلکالین فسفاتاز یا آلکالین فسفاتاز	۰/۰۳۳۷
۶	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز و آلومین	۰/۰۳۱۸
۷	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلومین و سن بیمار	۰/۰۳۱۱
۸	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلومین، سن بیمار و جنسیت بیمار	۰/۰۲۸۱
۹	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلومین، سن بیمار، جنسیت بیمار و مجموع پروتئین‌ها	۰/۰۸۲۹
۱۰	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلومین، سن بیمار، جنسیت بیمار، مجموع پروتئین‌ها و نسبت آلومین و نسبت گلوبولین	۰/۰۶۷۶

همچنین اگر تعداد ویژگی‌ها برابر با ۴ ویژگی باشد، میانگین مربعات خطا کمتر است.

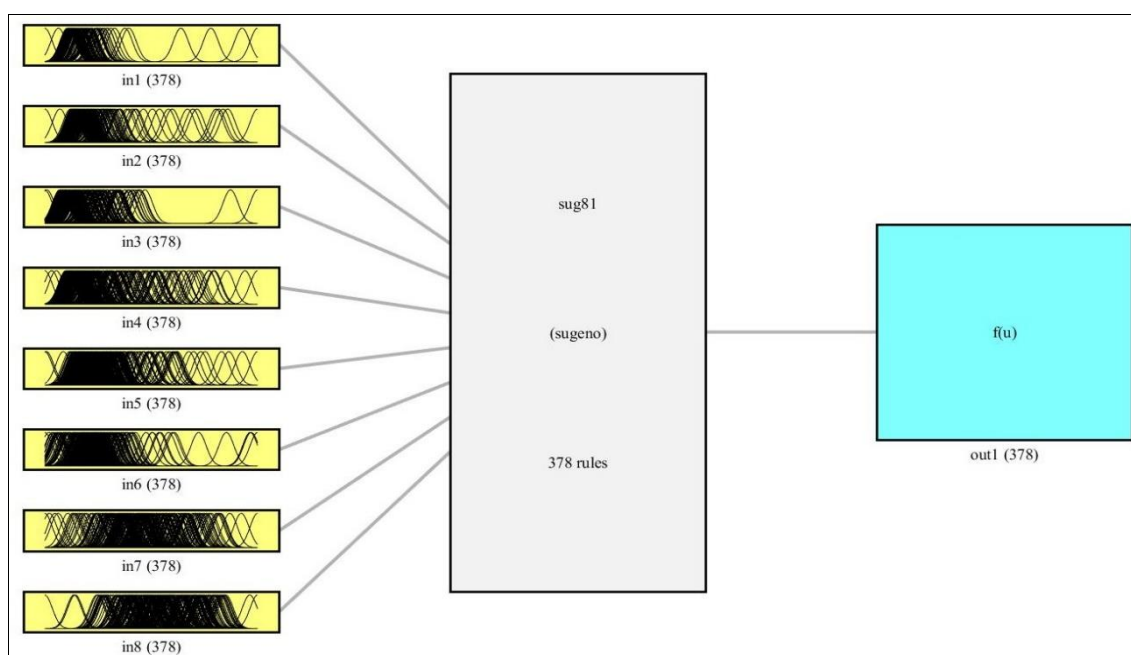
جدول ۵ نشان داد که اگر تعداد ویژگی‌ها برابر ۶ ویژگی باشد میانگین مربعات خطا کمتر است؛ لذا درصد صحت بیشتر است.

جدول ۵: نتایج پارامترها براساس genfis3

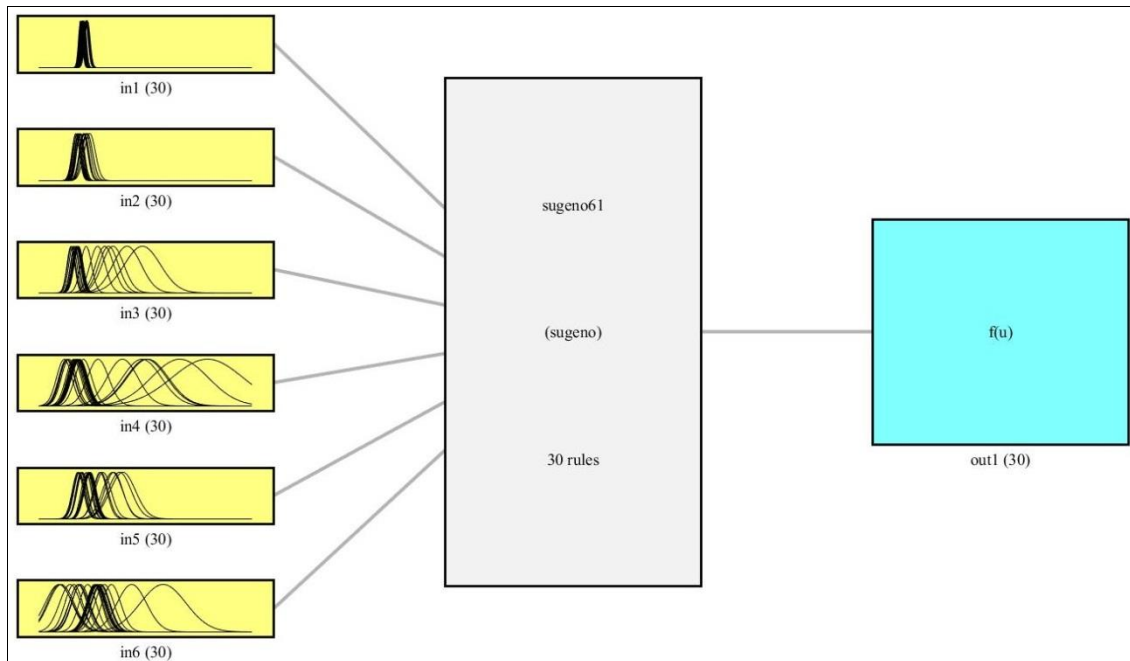
تعداد ویژگی	ویژگی‌ها	میانگین مربع خطا
۳	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین و مجموع بیلی روبین	۰/۰۷۷۱
۴	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین و بیلی روبین مستقیم	۰/۰۳۸۹
۵	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم و آلکالین فسفاتاز	۰/۰۵۱۰
۶	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز و آلبومین	۰/۰۳۸۶
۷	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین و سن بیمار	۰/۰۵۷۱
۸	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین، سن بیمار و جنسیت بیمار	۰/۰۶۳۴
۹	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین، سن بیمار، جنسیت بیمار و مجموع پروتئین‌ها	۰/۱۱۶۴
۱۰	آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین، سن بیمار، جنسیت بیمار، مجموع پروتئین‌ها و نسبت آلبومین و نسبت گلوبولین	۰/۱۱۹۲

پس کمترین میانگین مربع خطا برای genfis2 برابر ۰/۰۲۸۱ برای هشت ویژگی آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین، سن بیمار و جنسیت بیمار و برای genfis3 برابر ۰/۰۳۸۶ برای شش ویژگی آسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز و آلبومین است. مشاهده شد که تفاوت در ویژگی‌های سن و جنسیت بیمار است. برای بررسی بیشتر این هشت و شش ویژگی لازم است به نتایج تشخیص بیماری کبد براساس این تعداد ویژگی پرداخته شود تا مشخص شود که آیا این تعداد شش ویژگی در تشخیص بهتر بیماری کبد کافی

است یا دو ویژگی سن و جنسیت بیمار نیز باید مدنظر قرار داده شود تا تشخیص بهتری صورت گیرد. شمای کلی سیستم استنتاج فازی برای genfis2 و genfis3 به ترتیب در شکل ۲ و ۳ نشان داده شده‌اند. سیستم‌های استنتاج فازی genfis2 از نوع سوگنو و شامل ۸ ورودی و ۱ خروجی و تابع عضویت ورودی‌ها از نوع گائوسی و خروجی از نوع خطی و تعداد توابع عضویت برای تک تک ورودی‌ها برابر ۳۷۸ و تعداد قوانین برابر ۳۷۸ است و سیستم استنتاج فازی genfis3 از نوع سوگنو می‌باشد که شامل ۶ ورودی و ۱ خروجی و تابع عضویت ورودی‌ها از نوع گائوسی و خروجی از نوع خطی و تعداد توابع عضویت برای تک تک ورودی‌ها برابر ۳۰ و تعداد قوانین برابر ۳۰ است.



شکل ۲: سیستم استنتاج فازی براساس genfis2



شکل ۳: سیستم استنتاج فازی براساس genfis3

نشان داده شد. این جدول نشان داد که درصد صحت در genfis2 برابر با ۹۱/۶۶ و در genfis3 برابر با ۸۹/۸۷ درصد می‌باشد.

برای بررسی این هشت و شش ویژگی، به صورت تصادفی، ۴۶۶ نمونه (۸۰ درصد) از داده‌ها را به عنوان داده‌های آموزش و ۱۱۷ نمونه (۲۰ درصد) از داده‌ها به عنوان داده‌های آزمایش انتخاب شدند. در جدول ۶ نتایج معیارهای دقت و میانگین مربع خطا

جدول ۶: نتایج داده‌های آزمایشی براساس هشت و شش ویژگی

نوع استنتاج منطق فازی	معیار	درصد
genfis2 با هشت ویژگی	دقت	۹۱/۶۶
	میانگین مربع خطا	۰/۰۳۴۳
genfis3 با شش ویژگی	دقت	۸۹/۸۷
	میانگین مربع خطا	۰/۰۴۷۹

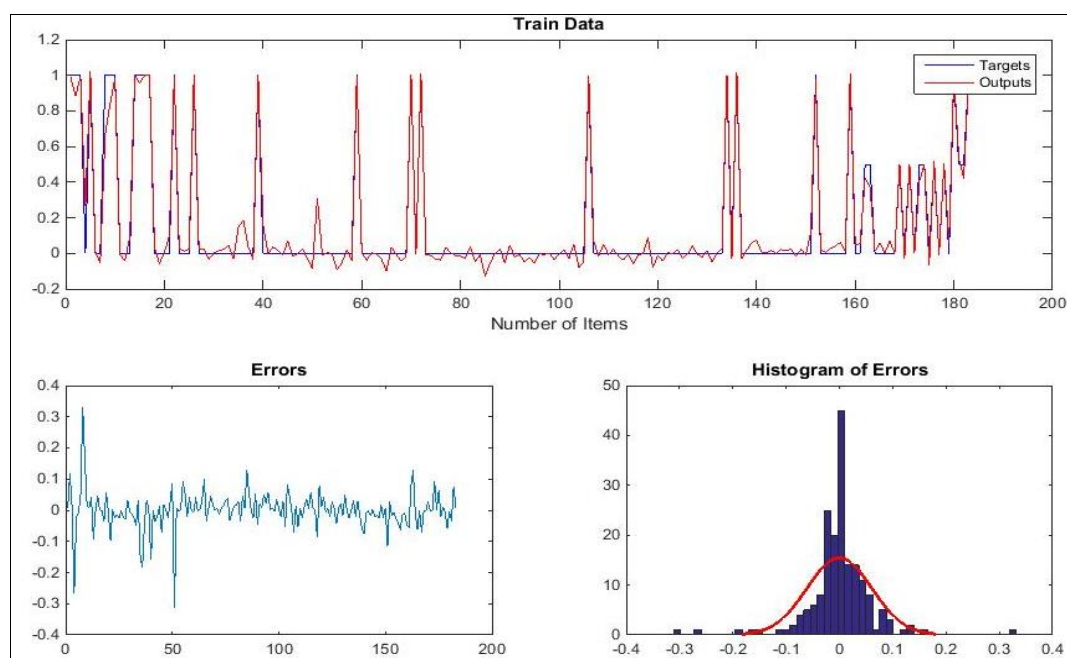
پشتیبانی تصمیم و سیستم‌هایی که قابلیت تفسیری از قوانین را دارند استفاده می‌شوند، خروجی الگوریتم استنتاج ممدانی غیرخطی و فازی است؛ اما خروجی سیستم استنتاج سوگنو خطی است، الگوریتم ممدانی در بسیاری از عملکردها که دقت خروجی و انعطاف‌پذیری مهم است به خوبی عمل نمی‌کند، ولی موتور استنتاج سوگنو انعطاف‌پذیرتر و از دقت عملکرد بالاتری برخوردار می‌باشد و از لحاظ محاسباتی نسبت به ممدانی مؤثرتر است، زیرا فرآیند غیرفازی کردن به زمان محاسباتی کمتری نیاز دارد.

در شکل ۴ تا ۷ نمودار خروجی و هدف سیستم استنتاج فازی پیشنهادی براساس، تفاضل و هیستوگرام خطای مطلق هر آیت

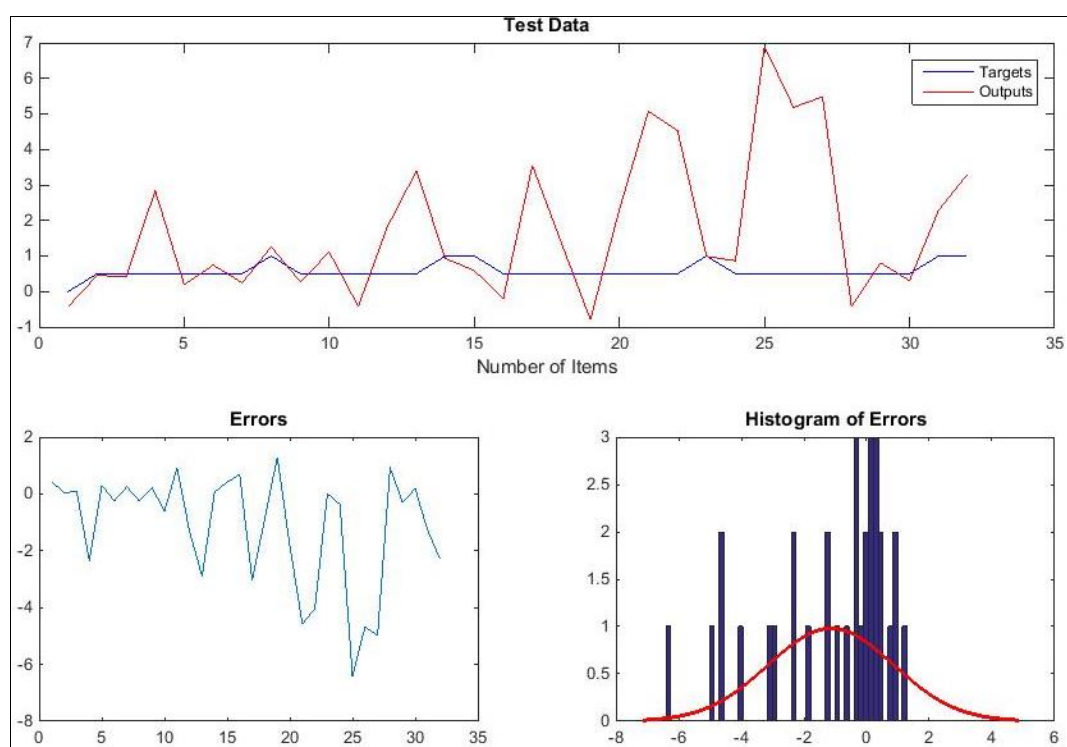
طبق نتایج به دست آمده دقت سیستم استنتاج فازی برای هشت ویژگی بهتر از شش ویژگی و همچنین خطای آن نیز کمتر است. کاهش خطا در genfis2 (بر مبنای تابع سوگنو) محسوس‌تر از genfis3 (بر مبنای تابع ممدانی) بوده است. پس برای تشخیص بیماری کبد از ویژگی‌های آپارات ترانس آمیناز، آمینوترانسفراز آلامین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلبومین، سن بیمار و جنسیت بیمار استفاده شد. الگوریتم استنتاج سوگنو بیشتر در سیستم‌های کنترلی و سیستم‌هایی که احتیاج به محاسبات ریاضی دارند استفاده شده؛ اما در الگوریتم استنتاج ممدانی نتایج منطقی با یک ساختار نسبتاً ساده بیان شد و بیشتر در سیستم‌های

genfis2 بهتر از genfis3 است.

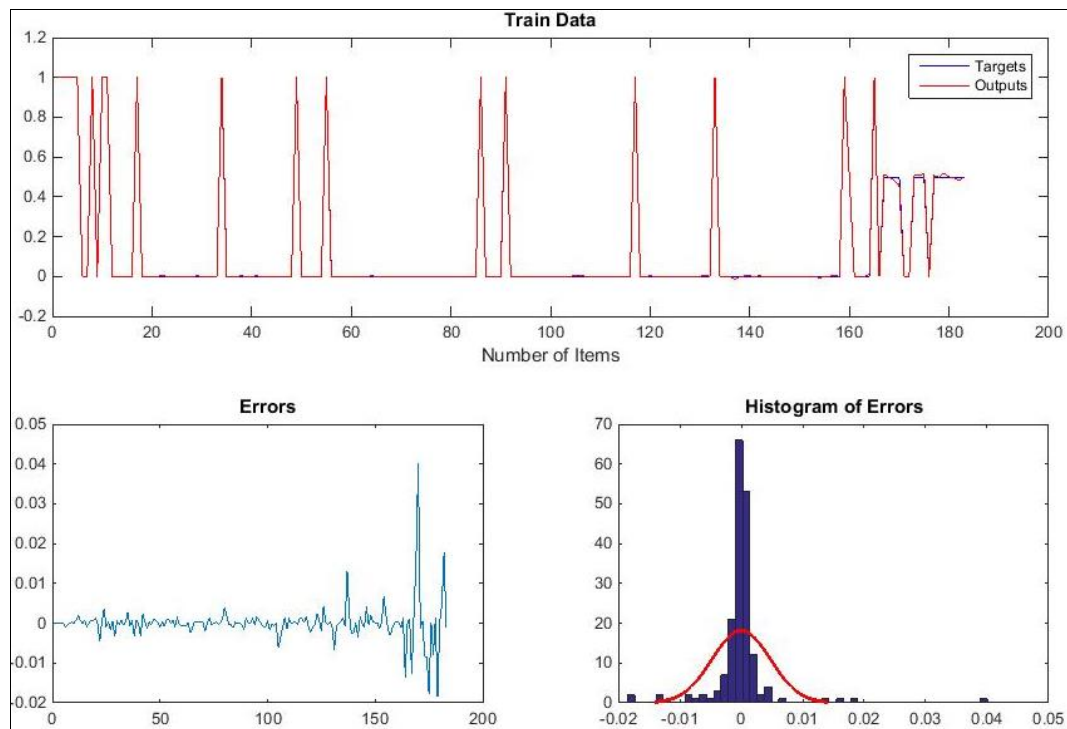
برای داده‌های آموزش و آزمایش نشان داده شد. همان‌طور که قابل مشاهده است، نتایج داده‌های آموزش و آزمایش



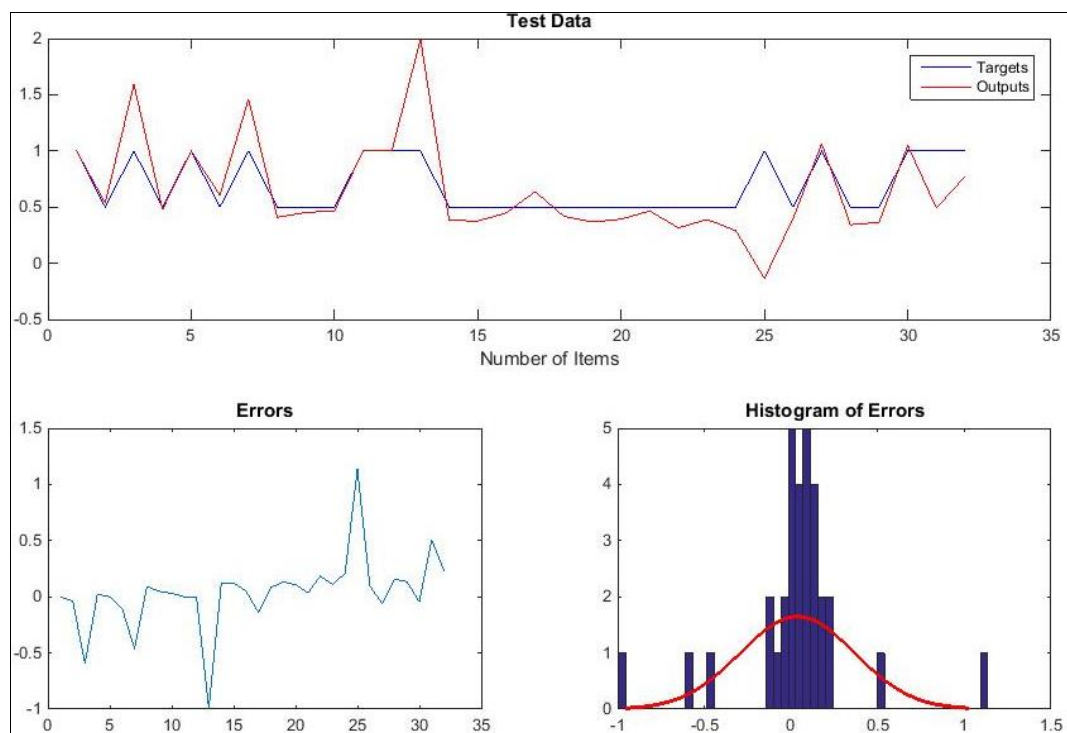
شکل ۴: نمودار خروجی و هدف سیستم استنتاج فازی پیشنهادی براساس genfis2، تفاضل و هیستوگرام خطای مطلق هر آیتم برای داده‌های آموزش



شکل ۵: نمودار خروجی و هدف سیستم استنتاج فازی پیشنهادی براساس genfis2، تفاضل و هیستوگرام خطای مطلق هر آیتم برای داده‌های آزمایش



شکل ۶: نمودار خروجی و هدف سیستم استنتاج فازی پیشنهادی براساس *genfis3*، تفاضل و هیستوگرام خطای مطلق هر آیتم برای داده‌های آموزش



شکل ۷: نمودار خروجی و هدف سیستم استنتاج فازی پیشنهادی براساس *genfis3*، تفاضل و هیستوگرام خطای مطلق هر آیتم برای داده‌های آزمایش

ماشین مانند ماشین بردار پشتیبان، نیوی بیز، k نزدیکترین همسایه برای طبقه‌بندی استفاده نمود.

میانگین مربع خطا براساس genfis2 برای ویژگی‌هایی مانند اسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز، آلومین، سن بیمار و جنسیت بیمار برابر ۰/۰۲۸۱ می‌باشد. میانگین مربع خطا براساس genfis3 برای ویژگی‌هایی مانند اسپاراتات ترانس آمیناز، آمینوترانسفراز آلانین، مجموع بیلی روبین، بیلی روبین مستقیم، آلکالین فسفاتاز و آلومین برابر ۰/۰۳۸۶ است.

متأسفانه اکثر پژوهشگران برای تشخیص بیماری کبد به جای مجموعه داده‌های استاندارد از عکس‌های CT استفاده نموده‌اند. دقت روش پیشنهادی نسبت به بهرامی‌راد و همکاران که برای طبقه‌بندی بیماری کبدی، به‌ترتیب از دو مجموعه داده AP و BUPA استفاده کرده‌اند، بیشتر است. درصد صحت آن‌ها برای رگرسیون لجستیک خطی (۶۹/۵۷ درصد)، شبکه عصبی مصنوعی (۷۳/۹۱ درصد)، ماشین بردار پشتیبان (۶۹/۲۳ درصد) و درخت رگرسیون (۶۶/۳۵ درصد) به دست آمده است [۱۵]. Khashma و Olaniyi [۱۴] برای تشخیص بیماری کبد با ۶ ورودی و ۲ خروجی با استفاده از شبکه عصبی پس‌انتشار (با ۱۰ نرون نهان) و شبکه عصبی پایه شعاعی (با ۲۰ نرون نهان) به‌ترتیب به دقت ۶۳٪ و ۷۰٪ دست یافته‌اند، که به‌مراتب دقت روش پیشنهادی (با ۸۹/۸۷ و ۹۱/۶۶ درصد) بیشتر است.

باتوجه به این که تمامی مطالعات نواقص و کاستی‌هایی دارند این مقاله نیز از این قاعده مستثنی نیست از جمله در مطالعه حاضر می‌توان وابستگی دقت تشخیص مدل پیشنهادی را به مقدار عملگرهای ادغام و جهش و همچنین تعیین توابع فازی ذکر کرد، همچنین باتوجه به اینکه نوشتن قوانین فازی به طور دقیق ممکن نیست؛ لذا می‌توان آن را به عنوان یک محدودیت و عیب به شمار آورد.

تعارض منافع

بدین وسیله نویسندگان تصریح می‌نمایند که هیچ‌گونه تعارض منافی در خصوص پژوهش حاضر وجود ندارد.

همان طور که از شکل ۴ تا ۷ قابل مشاهده است، تقریباً نمودار خروجی سیستم استنتاج فازی منطق بر هدف است که این انطباق برای خروجی سیستم استنتاج فازی genfis2 محسوس‌تر از genfis3 می‌باشد. نمودار نوسانات تفاضل خطای مطلق خروجی بسیار کم و همچنین نمودار هیستوگرام خطای مطلق دارای پراکندگی کمی می‌باشند.

بحث و نتیجه‌گیری

در مطالعه حاضر، هدف ارائه روش نوین برای تعیین پارامترهای منطق فازی با استفاده از الگوریتم ژنتیک برای تشخیص بیماری کبد بود. برای این منظور از مجموعه داده بیماران کبدی سایت UCI استفاده گردید. پس از رتبه‌بندی ویژگی‌های مجموعه داده به کمک آنتروپی و بهبود مقادیر ویژگی‌ها با کمک الگوریتم ژنتیک؛ تشخیص بیماری با استفاده از منطق فازی ارزیابی شد. در نتیجه سیستم استنتاج فازی genfis2 با ۸ ورودی و یک خروجی و سیستم استنتاج فازی genfis3 با ۶ ورودی و یک خروجی پیشنهاد گردید. نتایج نشان داد که دقت تشخیص بیماری کبد با کمک سیستم استنتاج فازی genfis2 با هشت ویژگی (۹۱/۶۶) بیشتر از genfis3 با شش ویژگی (۸۹/۸۷) و همچنین خطای genfis2 (۰/۰۳۴) کمتر از genfis3 (۰/۰۴۷) است. دقت روش پیشنهادی نسبت به تکنیک‌های داده کاوی [۱۵] که برای طبقه‌بندی بیماری کبدی، به‌ترتیب از دو مجموعه داده AP و BUPA استفاده و به دقت ۷۳/۳۹ و ۶۷/۳۹ رسیده‌اند، نیز به‌مراتب بیشتر است؛ و این روش را می‌توان به عنوان یکی از روش‌های مطمئن برای تشخیص بیماری کبد چرب استفاده کرد. برای بهبود عملکرد سیستم جاری و همچنین گسترش مدل پیشنهادی کارهای مختلفی را می‌توان انجام داد. برای کارهای آینده می‌توان از سیستم استنتاج فازی بر مبنای ترکیب دیگر الگوریتم‌های فراابتکاری مانند کلونی زنبور مصنوعی، الگوریتم کرم شب‌تاب و الگوریتم بهینه‌سازی گرگ خاکستری استفاده نمود. استفاده از الگوریتم‌های فراابتکاری برای انتخاب بهتر ویژگی‌ها است و همچنین می‌توان از روش‌های یادگیری

References

- Shuja A, Malespin M, Scolapio J. Nutritional Considerations in Liver Disease. Gastroenterol Clin North Am 2018;47(1):243-52.

- Shukla A, Tiwari R, Kaur P, Janghel RR. Diagnosis of Thyroid Disorders using Artificial Neural Networks. IEEE International Advance Computing Conference 2009 Mar 6-7; Patiala, India: IEEE; 2009. p. 1016-20.

3. Isa I, Saad Z, Omar S, Osman MK, Ahmad KA, Mat Sakim HA. Suitable MLP Network Activation Functions for Breast Cancer and Thyroid Disease Detection. Second International Conference on Computational Intelligence, Modelling and Simulation; 2010 Sep 28-30; Tuban, Indonesia: IEEE; 2010. p. 39-44.
4. Sarasvathi V, Santhakumaran A. Towards artificial neural network model to diagnose thyroid problems. Global Journal of Computer Science and Technology 2011; 11(5): 53-5.
5. Chuang CL. Case-based reasoning support for liver disease diagnosis. Artif Intell Med 2011;53(1):15-23.
6. Afzali F, Heidari Z, Montazeri M, Ahmadian L, Zahedi M J. Futures studies in health: choosing the best intelligent data mining model to predict and diagnose liver cancer in early stage. Journal of Health and Biomedical Informatics 2015; 2(3):133-40. Persian
7. Rajamani R, Rathika M. Analysis of liver cancer using Adaptive Neuro Fuzzy Inference System (ANFIS). International Journal of Innovative Research in Computer and Communication Engineering 2015; 3(7): 6757-63.
8. Hashmi A, Saleem Khanb M. Diagnosis blood test for liver disease using fuzzy logic. International Journal of Sciences: Basic and Applied Research 2015; 20(1): 151-83.
9. Obayya MI, Areed NF. Liver Cancer Identification using Adaptive Neuro-Fuzzy Inference System. International Journal of Computer Applications 2016; 140(8): 1-7.
10. Fan CY, Chang PC, Lin JJ, Hsieh JC. A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. Applied Soft Computin 2011;11(1):632-44.
11. Satarkar SL, Ali MS. Fuzzy expert system for the diagnosis of common liver disease. International Engineering Journal for Research & Development 2013; 1(1): 1-7.
12. Kalaiselvi C, Nasira GM. A novel approach for the diagnosis of diabetes and liver cancer using ANFIS and improved KNN. Research Journal of Applied Sciences, Engineering and Technology 2014; 8(2): 243-50.
13. Neshat M, Yaghobi M. Designing a fuzzy expert system of diagnosing the Hepatitis B Intensity rate and comparing it with adaptive neural network fuzzy system. Proceedings of the World Congress on Engineering and Computer Science, 2009 Oct 20-22; San Francisco, USA: WCECS; 2009.
14. Olaniyi EO, Khashman A. Liver Disease Diagnosis Based on Neural Networks. Proceedings of the 16th International Conference on Neural Networks (NN'15); 2013 Nov 7-9; At Rome, Italy: Advances in Computational Intelligence; 2013. p. 48-53.
15. Bahramirad S, Mustapha A, Eshraghi M. Classification of Liver Disease Diagnosis: A Comparative Study. Informatics and Applications. Second International Conference on Informatics & Applications (ICIA); 2013 Sep 23-25; Lodz, Poland: IEEE; 2013. p. 42- 6.
16. Mirsharif M, Rouhani S. Data mining approach based on neural network and decision tree methods for the early diagnosis of risk of gestational diabetes mellitus. Journal of Health and Biomedical Informatics 2017; 4 (1):59-68. Persian
17. UCI. Machine Learning Repository. [cited 2018 Aug 20] Available from: <https://archive.ics.uci.edu/ml/datasets>
18. Rajaobelina L, Brun I, Toufaily E. A relational classification of online banking customers. International Journal of Bank Marketing 2013; 31(3): 187-205.
19. Nguyen T, Singh M. Impact of internet banking on customer satisfaction and loyalty: a conceptual model. The Fourth International Conference on Electronic Business; 2004 Dec 5-9; Beijing, China: ICEB; 2004. p. 298-303.
20. Mitchell M. An Introduction to Genetic Algorithms (Complex Adaptive Systems). Cambridge, MA, USA: MIT Press; 1996.
21. Dhimish M, Holmes V, Mehrdadi B, Dales M. Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection. Renewable Energy 2018;117:257-74.
22. David RC, Precup RE, Petriu EM, Rădac MB, Preitl S. Gravitational search algorithm-based design of fuzzy control systems with a reduced parametric sensitivity. Information Sciences 2013;247:154-73.
23. Renault A, Dagli C. Genetic Algorithm Optimization of SoS Meta-Architecture Attributes for Fuzzy Rule Based Assessments. Procedia Computer Science 2016;95:95-102.
24. Xu D, Fu H, Jiang W. Research on Liver Disease Diagnosis Based on RS_LMBP Neural Network. 12th International Conference on Computational Intelligence and Security (CIS); 2016 Dec 16-19; Wuxi, China: IEEE; 2016. p. 646-9.

Determining Fuzzy Logic Parameters by using Genetic Algorithm for the Diagnosis of Liver Disease

Rezaei Farokhzad Maryam¹, Soleimanian Gharehchopogh Farhad^{2*}

• Received: 9 Dec, 2017

• Accepted: 29 Sep, 2018

Introduction: Liver disease is one of the most common chronic liver problems and cirrhosis. Liver problems include a wide range of diseases and disorders that damage the liver tissue or its function. Early diagnosis and treatment of this disease can reduce the severity of the disease and mortality rate.

Method: In this descriptive-analytic study, database was consisted of 583 independent records, including 11 features in the UCI machine learning database and through using fuzzy logic that its parameters are determined by Genetic Algorithm (GA), a method for the diagnosis of liver disease is proposed. For this purpose, first, the features of the dataset were ranked using the entropy feature and then, the dataset data were optimized using GA. Ultimately, liver disease was diagnosed using the genfis2 and genfis3 Fuzzy Inference System (FIS).

Results: The results show that the accuracy of detection of liver disease using the genfis2 FIS with 8 features is 91.66% and using the genfis3 FIS with 6 features, it is equal to 89.87%. Moreover, the rates of error for genfis2 and genfis3 were 0.034 and 0.047 respectively.

Conclusion: Liver disease is one of the most common diseases in population. Early diagnosis of disease while reducing costs can increase the chance of treatment success. According to the obtained results, the proposed model can identify people with liver disease with a fairly high degree of accuracy.

Keywords: Liver Disease Detection, Fuzzy Logic, Fuzzy Inference System, Genetic Algorithm

• **Citation:** Rezaei Farokhzad M, Soleimanian Gharehchopogh F. Determining Fuzzy Logic Parameters by using Genetic Algorithm for the Diagnosis of Liver Disease. *Journal of Health and Biomedical Informatics* 2018; 5(3): 384-397.

1. M.Sc. in Computer Engineering, Computer Engineering Dept., Urmia Branch, Islamic Azad University, Urmia, Iran
2. Ph.D., in Computer Engineering, Assistant Professor, Computer Engineering Dept., Urmia Branch, Islamic Azad University, Urmia, Iran

*Correspondence: Urmia, 2km Airport, Islamic Azad University, Urmia, Faculty of Engineering

• Tel: 09141764427

• Email: bonab.farhad@gmail.com