نشریه علمی‌ترویجی محاسبات نرم
سال ششم/ شمارهٔ دوم/ پاییز و زمستان ۱۳۹۶/ صفحهٔ ۳۲ـ۴۷

دریافت مقاله: ۱۳۹۶/۸/۱۳
پذیرش مقاله: ۱۳۹۸/۳/۱۰

ارائهٔ مدل هوشمند هایبریدی فازی-تکامل ژنتیکی تفاضلی در یک سیستم خبرهٔ فازی برای پیش‌بینی خطر ابتلا به بیماری قلبی

رعنا آخوندی[۱]، راحیل حسینی[۲*]

[۱] کارشناسی ارشد، گروه هوش مصنوعی، واحد شهر قدس، دانشگاه آزاد اسلامی، تهران، ایران
Rana.akhoondi@qodsiau.ac.ir
[۲] استادیار، گروه مهندسی کامپیوتر، واحد شهر قدس، دانشگاه آزاد اسلامی، تهران، ایران
Rahil.hosseini@qodsiau.ac.ir

چکیده: با توجه به اهمیت روزافزون سیستم‌های هوشمند تصمیم یار به‌عنوان ابزاری در اختیار پزشکان، هدف از این مقاله، ارائهٔ یک سیستم هوشمند برای پیش‌بینی بیماری قلبی به‌دلیل پیشگیری از خطر ابتلا به بیماری قلبی است. برای این منظور، یک سیستم خبرهٔ فازی پیشنهاد شده است. به‌منظور تنظیم پارامترهای توابع عضویت سیستم فازی، بهبود عملکرد و افزایش دقت آن، الگوریتم ژنتیک و الگوریتم تکامل تفاضلی روی سیستم خبرهٔ فازی اعمال شده‌اند. مدل‌های ترکیبی فازی-ژنتیک و فازی-تکامل تفاضلی با استفاده از روش‌های ارزیابی تحلیل منحنی ROC و اعتبارسنجی ۱۰- بخشی روی مجموعه داده‌ای شامل ۳۸۰ نمونه که از بیمارستان پارسیان تهیه شده، مورد ارزیابی قرار گرفته است. بر اساس نتایج به‌دست‌آمده، سیستم خبرهٔ فازی دارای صحت عملکرد ۸۵/۵۲٪ می‌باشد که پس از اعمال مدل ترکیبی تکاملی فازی-ژنتیکی به ۹۷/۹۳٪، و مدل ترکیبی فازی-تفاضل تکاملی به ۹۷/۶۷٪ افزایش یافته است. نتایج بیانگر آن است که در مدل پیشنهادی فازی- تکامل ژنتیکی و تفاضلی صحت عملکرد سیستم خبرهٔ فازی به‌طور چشمگیری بهبود یافته است. قابلیت تفسیر نتایج در سیستم خبرهٔ فازی و صحت عملکرد بالای مدل‌های تکاملی، به‌منظور پیش‌بینی زودهنگام بیماری‌های قلبی و ارائهٔ مراقبت‌های لازم بسیار امیدبخش است.

واژه‌های کلیدی: سیستم خبرهٔ فازی، الگوریتم ژنتیک، الگوریتم تکامل تفاضلی، تحلیل منحنی ROC، پیش‌بینی بیماری قلبی.

۱. مقدمه

بیشتر پزشکان و افراد خبرهٔ پزشکی، پیش‌بینی بیماری‌ها و پیشگیری از پیشرفت آن‌ها را امری ضروری‌تر و مهم‌تر از تشخیص و درمان آن‌ها تلقی می‌کنند. پیش‌بینی بیماری‌ها و جلوگیری از پیشرفت آن‌ها بسیار ساده‌تر و کم‌هزینه‌تر از تشخیص آن‌هاست. بنابراین محققان و پژوهشگران به‌کارگیری روش‌های مفید هوش محاسباتی و الگوریتم‌های قدرتمند فراابتکاری را در راستای تحقق این کاربرد مهم، مناسب دانسته و می‌کوشند تا با بهینه‌سازی این روش‌ها و ترکیب آن‌ها به‌منظور بهره‌گیری از مزایای هرکدام، دقت این سیستم‌های حیاتی را در عرصهٔ پزشکی بهبود دهند.

روش کلی در تحقیقات پیش‌بینی بیماری قلبی این است که ابتدا با استفاده از روش‌های مختلف انتخاب ویژگی، ویژگی‌های مهم و مرتبط با بیماری را از مجموعه داده‌های پزشکی استخراج کرده و سپس با اعمال تکنیک‌های مختلف طبقه‌بندی خطر ابتلا به بیماری قلبی را در افراد در معرض خطر بیماری قلبی و یا حتی افراد سالم در چند سال آینده پیش‌بینی می‌کنند.

انواع روش‌های مختلف هوش محاسباتی می‌تواند برای طراحی سیستم‌های کاربردی مورد استفاده قرار گیرد اما از آنجایی که اکثر سیستم‌های تشخیص پزشکی دارای منابع عدم قطعیت (پایگاه داده‌ها) هستند، باید از روش‌هایی برای طراحی آن‌ها استفاده کرد که دارای قابلیت کنترل عدم قطعیت این سیستم‌ها باشند. قدرتمندترین سیستم کنترل و مدیریت عدم قطعیت، سیستم استنتاج فازی است که قادر است کار پردازش با کلمات را به‌خوبی انجام دهد. بنابراین سیستم خبرهٔ فازی طراحی‌شده در [۱] به‌عنوان مدل پایه برای پیش‌بینی بیماری قلبی در نظر گرفته می‌شود. هدف از این مقاله، اعمال الگوریتم‌های ژنتیک و تکامل تفاضلی روی این سیستم خبره و ارائهٔ مدل‌های هایبریدی Fuzzy-GA و Fuzzy-DE به‌منظور تنظیم پارامترهای توابع عضویت و بهینه‌سازی نتایج آن سیستم است.

۲. مروری بر کارهای گذشته

یک الگوریتم ترکیبی ANN-GWO[1] به‌منظور پیش‌بینی بیماری قلبی در [۲] ارائه شده است. این مقاله به بررسی قدرت ترکیب دو روش هوش محاسباتی بهینه‌سازی گرگ خاکستری (GWO) و شبکه‌های عصبی مصنوعی (ANN) برای پیش‌بینی بیماری قلبی می‌پردازد. پارامترهای پزشکی که شامل ۱۳ پارامتر (جنسیت، فشار خون، کلسترول و...) هستند به‌عنوان ورودی و ۴ کلاس (طبیعی، اولین سکته، دومین سکته و پایان زندگی) به‌عنوان خروجی‌های سیستم در نظر گرفته می‌شوند. عملکرد الگوریتم ANN-GWO با عملکرد شبکهٔ عصبی پس انتشار استاندارد[2] (BPNN) مقایسه شده و با استفاده از معیار RMSE[3] ارزیابی می‌شود. میزان RMSE برای BPNN و ANN-GWO به‌ترتیب ۰/۰۰۴۳ و ۰/۰۰۳۶ است. طبق نتایج به‌دست‌آمده در روش پیشنهادی سرعت همگرایی و دقت پیش‌بینی افزایش می‌یابد.

عملکرد تعدادی از الگوریتم‌های یادگیری ماشین باناظر[4] برای پیش‌بینی بیماری قلبی با کاهش تعداد ویژگی‌ها در [۳] مورد بررسی قرار گرفته است. در این مقاله، از روش انتخاب ویژگی PCA[5] برای کاهش ویژگی‌های مورد استفادهٔ مدل‌های پیش‌بینی مانند ماشین بردار پشتیبان[6]، بیز ساده[7] و درخت تصمیم[8] به‌منظور افزایش دقت آن‌ها استفاده شده است و سپس عملکرد آن‌ها قبل و بعد از اعمال روش PCA (یک روش آماری) مورد ارزیابی و مقایسه قرار گرفته است. مجموعه دادهٔ مورد استفاده در این مطالعه، از پایگاه دادهٔ Cleveland که شامل ۳۰۳ رکورد و ۱۱ ویژگی می‌باشد، فراهم شده است. مجموعهٔ داده شامل ویژگی‌هایی مانند سن، جنسیت، درد قفسهٔ سینه، فشار خون و... و همچنین یک ویژگی تشخیص بیماری قلبی است که عدم حضور و احتمال حضور بیماری را به‌ترتیب با ۰ و ۱ مشخص می‌کند. پس از اعمال روش PCA بر روی

1. Artificial Neural Networks- Gray Wolf Optimization
2. Back-Propagation Neural Network
3. Root Mean Square Error
4. Supervised learning algorithms
5. Principal Component Analysis
6. Support Vector Machine
7. Naive Bayes
8. Decision Tree

مجموعهٔ داده، ویژگی‌ها به ۶ ویژگی کاهش می‌یابد. پس از کاهش اندازهٔ مجموعهٔ داده و اعمال آن روی مدل‌های پیش‌بینی، ماشین بردار پشتیبان دارای عملکرد بهتری نسبت به بیز ساده و درخت تصمیم است. دقت پیش‌بینی ماشین بردار پشتیبان قبل و بعد از کاهش ویژگی‌ها به‌ترتیب برابر با ۴۶٪ و ۶۷٪ است.

یک شبکهٔ عصبی کوانتومی[1] در [۴] برای پیش‌بینی بیماری قلبی طراحی شده است. در این مقاله، یک سیستم پیش‌بینی بیماری قلبی جدید بر اساس یادگیری ماشین ارائه شده است. این مقاله، از داده‌های ۶۸۹ بیمار که دارای علائم بیماری قلبی هستند، استفاده کرده است. همچنین برای ارزیابی مدل ارائه‌شده، مجموعه داده‌ای شامل ۵۲۰۹ بیمار قلبی که از مطالعهٔ فرامینگهام[2] تهیه شده است، مورد استفاده قرار می‌گیرد. این مجموعهٔ داده دارای تشخیص‌های واقعی توسط پزشکان نیز می‌باشد. در سیستم ارائه‌شده از شبکهٔ عصبی کوانتومی استفاده می‌شود. دقت سیستم ارائه‌شده برای پیش‌بینی خطر بیماری قلبی ۹۸/۵۷٪ است. ویژگی‌های سن، جنسیت، کلسترول، HDL، فشارخون، دیابت، شاخص تودهٔ بدن و مصرف سیگار برای پیش‌بینی بیماری قلبی در سیستم ارائه‌شده استفاده می‌شوند. در نهایت، سیستم ارائه‌شده، داده‌های بیماران را برای پیش‌بینی بیماری قلبی در چهار گروه طبیعی، کم‌خطر، خطر متوسط و پرخطر دسته‌بندی می‌کند.

یک مدل پیش‌بینی بیماری قلبی در [۵] با استفاده از تکنیک‌های یادگیری ماشین ارائه شده است. روش پیشنهادی در این پژوهش برای طراحی مدل پیش‌بینی به‌کارگیری تکنیک‌های یادگیری ماشین است. روش‌های درخت طبقه‌بندی (ID3[3])، بیز ساده، جنگل تصادفی[4] و ماشین بردار پشتیبان از جمله روش‌هایی هستند که در طراحی این مدل پیش‌بینی به کار گرفته می‌شوند. مجموعه داده‌ای که برای آموزش مدل پیش‌بینی مورد استفاده قرار می‌گیرد، از UCI تهیه شده است. بر اساس ارزیابی‌های انجام‌شده، بهترین و مناسب‌ترین طبقه‌بندی، درخت طبقه‌بندی (ID3) است؛ زیرا از صحت و دقت بالایی برخوردار است و همچنین کارایی بالایی در پیش‌بینی دارد. مهم‌ترین مشکل این روش این است که از تکنیک‌های یادگیری باناظر استفاده می‌کند و این تکنیک‌ها نیازمند داده‌های برچسب‌دار هستند که برچسب‌گذاری داده‌ها فرایندی پرهزینه است.

یک مدل ترکیبی فازی- ژنتیک برای پیش‌بینی بیماری قلبی در [۶] ارائه شده است. در این مقاله، یک سیستم با استفاده از روش ترکیبی الگوریتم ژنتیک و منطق فازی طراحی می‌شود. مجموعه داده‌ای که برای ارزیابی این سیستم استفاده می‌شود از UCI تهیه شده است. با اعمال الگوریتم ژنتیک روی این مجموعهٔ داده تعداد ویژگی‌ها از ۱۳ به ۷ کاهش یافته و همین امر سبب افزایش دقت ۱/۵۴درصدی آن نسبت به روش‌های موجود شده است. برخی از ویژگی‌های موجود در این مجموعهٔ داده عبارت‌اند از: سن، جنسیت، نوع درد قفسهٔ سینه، قند خون در حالت استراحت، قند خون در حالت فعالیت، کلسترول، نوار قلب در حالت استراحت، بیشترین ضربان قلب، آنژین صدری ناشی از ورزش و تعداد رگ‌های گرفته‌شده.

سیستمی دقیق مبتنی بر الگوریتم کرم شب‌تاب[5] برای پیش‌بینی بیماری قلبی در [۷] ارائه شده است. در این مقاله، یک سیستم تشخیص بیماری قلبی با استفاده از مجموعه‌های ناهموار بر اساس کاهش ویژگی‌ها و سیستم فازی نوع ۲ ارائه می‌شود. در این مقاله از دو روش ازدحام ذرات و الگوریتم کرم شب‌تاب برای کاهش ویژگی‌ها استفاده شده است. نتایج آزمایش‌ها نشان می‌دهد که سیستم ارائه‌شده عملکرد بهتری در مقایسه با سایر روش‌های یادگیری ماشین از قبیل بیز ساده، ماشین بردار پشتیبان و شبکهٔ عصبی مصنوعی دارد. برای ارزیابی این روش از مجموعه دادهٔ UCI استفاده می‌شود.

یک شبکهٔ عصبی مصنوعی برای پیش‌بینی بیماری قلبی در [۸] ارائه شده است. در این پژوهش، از شبکهٔ عصبی چندلایه با الگوریتم یادگیری پس انتشار و شبکهٔ عصبی چندلایه با الگوریتم ژنتیک استفاده شده است. داده‌هایی که برای ارزیابی این مدل استفاده می‌شود، از طریق تکنیک‌های مختلف نمونه‌برداری جمع‌آوری شده است. مجموعه تست شامل ۴۳

1. quantum neural network
2. Framingham
3. Iterative Dichotomized 3
4. Random Forest
5. Firefly

ویژگی ورودی و یک خروجی از ۵ بیمار است. نتایج ارزیابی سیستم و دقتی برابر با ۹۹/۳۱٪ نشان می‌دهد.

استفاده از یک سیستم فازی- ژنتیک و مدل دینامیکی سیگنال‌های ECG بر اساس طبقه‌بندی سیگنال‌های ECG برای پیش‌بینی بیماری قلبی در [۹] ارائه شده است. در این مقاله، یک طبقه‌بند فازی بر اساس مدل ممدانی ایجاد می‌شود و پس از پیاده‌سازی این طبقه‌بند به‌منظور تعیین کارایی آن ۳۰۰ سیگنال ECG انتخاب شده و روی آن اعمال می‌شود. این سیگنال‌ها شامل ۴ گروه ECG بوده که هریک شامل ۷۵ عضو است. پس از اعمال ۳۰۰ سیگنال ECG نتایج به‌دست‌آمده نشان می‌دهد که این طبقه‌بند قادر است سیگنال‌های ECG را با دقتی برابر با ۹۹/۳۴٪ طبقه‌بندی کند. طبقه‌بند دیگری که برای کشف آریتمی قلبی از بین سیگنال‌های ECG طراحی می‌شود، طبقه‌بند فازی- ژنتیک است. پس از بررسی عملکرد طبقه‌بند فازی- ژنتیک نتایج نشان می‌دهد که این طبقه‌بند سیگنال‌های ECG را با دقت ۹۸/۶۷٪ طبقه‌بندی می‌کند.

یک سیستم پیش‌بینی بیماری قلبی در [۱۰] با استفاده از شبکهٔ عصبی و الگوریتم ژنتیک ارائه شده است. این سیستم، تعداد واحدهای لایه‌های پنهان شبکهٔ عصبی چندلایه با الگوریتم یادگیری پس‌انتشار را به‌منظور یادگیری شبکه با ساختار مناسب، محاسبه می‌کند و سپس از الگوریتم بهینه‌سازی ژنتیک برای مقداردهی اولیهٔ وزن‌های شبکهٔ عصبی استفاده می‌کند. این سیستم برای پیش‌بینی بیماری قلبی از ۱۲ پارامتر مانند جنسیت، سن، فشار خون و... استفاده می‌کند. همچنین یک ویژگی تشخیص بیماری قلبی به‌عنوان خروجی در این سیستم استفاده می‌شود که عدم حضور و حضور بیماری قلبی را به‌ترتیب با ۰ و ۱ مشخص می‌کند. برای ارزیابی این سیستم، از مجموعه داده‌ای که از انجمن قلب آمریکا تهیه شده و شامل عوامل خطر بیماری قلبی مربوط به ۵۰ نفر می‌باشد، استفاده شده است. با توجه به نتایج به‌دست‌آمده، روش عصبی- ژنتیک، بیماری قلبی را با دقت ۹۸٪ پیش‌بینی می‌کند.

یک سیستم هوشمند برای پیش‌بینی بیماری قلبی در [۱۱] با استفاده از روش‌های داده‌کاوی ارائه شده است. هدف اصلی این مقاله توسعهٔ یک سیستم هوشمند پیش‌بینی بیماری قلبی با استفاده از سه تکنیک داده‌کاوی درخت تصمیم، بیز ساده و شبکهٔ عصبی است. مجموعه داده‌ای که برای آموزش و ارزیابی سیستم ارائه‌شده استفاده می‌شود، مجموعه دادهٔ Cleveland شامل ۹۰۹ رکورد و ۱۵ ویژگی است. این مجموعهٔ داده شامل ۱۳ ویژگی ورودی مانند (سن، جنسیت، درد قفسهٔ سینه، قند خون، نوار قلب، کلسترول و...)، یک ویژگی پیش‌بینی برای تشخیص بیماری قلبی و یک ویژگی کلیدی برای شناسایی بیمار است. به‌منظور اعتبارسنجی کارایی مدل‌ها، از دو روش نمودار پیشرفت[1] و ماتریس طبقه‌بندی استفاده می‌شود. دقت بیز ساده، درخت تصمیم و شبکهٔ عصبی به‌ترتیب ۹۵٪، ۹۴/۹۳٪ و ۹۳/۵۴٪ هستند.

یک شبکهٔ عصبی پس‌انتشار برای پیش‌بینی بیماری قلبی در [۱۲] ارائه شده است. سیستم ارائه‌شده برای پیش‌بینی بیماری قلبی از ۱۳ ویژگی استفاده می‌کند. مجموعه داده‌ای که برای ارزیابی این سیستم استفاده می‌شود، از پایگاه دادهٔ Cleveland تهیه شده که شامل ۱۶۶ رکورد است. مجموعهٔ داده همچنین دارای ۱۳ ویژگی ورودی (سن، جنسیت، درد قفسهٔ سینه، قند خون، نوار قلب، کلسترول و...) و ۴ کلاس خروجی (۳ کلاس برای انواع وضعیت‌های بیماری قلبی و ۱ کلاس برای عدم حضور بیماری قلبی) است. بیشترین دقت شبکهٔ عصبی ارائه‌شده در مراحل آموزش و تست به‌ترتیب ۸۶٪ و ۹۲٪ است.

یک سیستم خبرهٔ فازی برای غربالگری بیماری قلبی با استفاده از پارامترهای کلینیکی در [۱۳] ارائه شده است. هدف از این مقاله، توسعهٔ یک سیستم خبرهٔ غربالگری فازی است که برای کشف و تشخیص بیماری قلبی در مراحل اولیه استفاده می‌شود. فاکتورهای خطر شناسایی‌شده عبارت‌اند از: سن، مصرف دخانیات، چاقی، فشار خون، دیابت، کلسترول و درد قفسهٔ سینه. برای ارزیابی این سیستم از مجموعه داده‌ای که شامل ۵۰۰ رکورد است و از بیمارستان AMRI تهیه شده است، استفاده می‌شود.

یک سیستم پشتیبانی از تصمیم‌گیری بالینی (CDSS[2]) مبتنی

1. Lift Chart
2. Clinical Decision Support System

بر قوانین فازی وزن‌دار برای پیش‌بینی سطح خطر بیماری قلبی در [۱۴] ارائه شده است. سیستم ارائه‌شده برای پیش‌بینی خطر بیماران قلبی از دو مرحله تشکیل شده است: ۱. روش خودکار برای تولید قوانین فازی وزن‌دار، ۲. توسعهٔ سیستم پشتیبانی از تصمیم‌گیری بالینی مبتنی بر قوانین فازی. برای ارزیابی سیستم از مجموعه داده‌ای که از UCI تهیه شده است، استفاده می‌شود. علائم حیاتی و ویژگی‌های اصلی در پیش‌بینی سطح خطر بیماران قلبی (سن، جنسیت، کلسترول، مصرف دخانیات و...) به‌عنوان ورودی‌ها و پیش‌بینی خطر بیماری قلبی به‌عنوان خروجی سیستم فازی ارائه‌شده استفاده می‌شود. جدول (۱) مروری بر کارهای گذشته را نشان می‌دهد.

جدول (۱): مروری بر کارهای گذشته

نام مقاله	روش	ورودی‌ها	خروجی‌ها	مزایا	عملکرد (معیار اندازه‌گیری)	مجموعه دادهٔ ارزیابی
سیستم تشخیص و پیش‌بینی بیماری قلبی [۱۳]	سیستم خبرهٔ فازی با استفاده از مدل استنتاج ممدانی	پارامترهای پزشکی	پیش‌بینی خطر بیماری قلبی در افراد و دسته‌بندی آن‌ها به سه گروه خطر	قابلیت درک آسان قوانین و نیاز به پارامترهای کلینیکی قابل دسترس و کم‌هزینه	٪۸۴/۲۰ (accuracy)	مجموعه داده‌ای شامل ۵۰۰ رکورد که از بیمارستان AMRI تهیه شده است.
سیستم پیش‌بینی سطح خطر بیماری قلبی [۱۴]	سیستم استنتاج فازی ممدانی و به‌کارگیری قوانین فازی وزن‌دار	پارامترهای پزشکی	پیش‌بینی خطر بیماری قلبی در دو دسته	بهبود عملکرد سیستم نسبت به شبکهٔ عصبی، تولید خودکار قوانین فازی و وزن آن‌ها	٪۵۷/۸۵ (accuracy)	UCI
سیستم پیش‌بینی بیماری قلبی [۴]	استفاده از شبکهٔ عصبی کوانتومی و مقایسهٔ آن با FRS	پارامترهای پزشکی	پیش‌بینی بیماری قلبی در چهار گروه طبیعی، کم‌خطر، خطر متوسط و پرخطر	افزایش قابل توجه دقت نسبت به FRS	٪۹۸۵۷/۵۷ (accuracy)	مجموعه داده‌هایی شامل ۶۸۹ و ۵۲۰۹ نمونه
سیستم پیش‌بینی بیماری قلبی [۱۲]	شبکهٔ عصبی پس انتشار	پارامترهای پزشکی	دسته‌بندی بیماران به ۴ دسته	عملکرد بهتر نسبت به روش‌های موجود	٪۹۲ (accuracy)	UCI (Cleveland)
الگوریتم‌های یادگیری ماشین برای پیش‌بینی بیماری قلبی [۳]	ارزیابی عملکرد مدل بل و بعد از اعمال روش انتخاب ویژگی PCA	پارامترهای پزشکی	پیش‌بینی حضور یا عدم حضور بیماری قلبی	بهبود عملکرد مدل‌های پیش‌بینی با کاهش ویژگی‌ها از ۱۱ به ۶ ویژگی	روش ماشین بردار پشتیبان دارای بیشترین دقت پیش‌بینی برابر با ٪۶۷ (accuracy)	UCI (Cleveland)
مدل پیش‌بینی بیماری قلبی [۵]	به‌کارگیری تکنیک‌های ID3، SVM، جنگل تصادفی و بیز ساده	پارامترهای پزشکی	پیش‌بینی خطر بیماری قلبی در دو دسته	استفاده از یک فرایند غربال‌گری ساده	روش ID3 دارای بهترین عملکرد برابر با ٪۹۲/۵ (accuracy)	UCI
سیستم هوشمند پیش‌بینی بیماری قلبی [۱۱]	تکنیک‌های داده‌کاوی مانند درخت تصمیم، بیز ساده و شبکهٔ عصبی	پارامترهای پزشکی	پیش‌بینی بیماری قلبی در دو دسته	قابلیت پاسخ‌گویی به جست‌وجوهای پیچیده، سهولت در تفسیر و دقت بالا	روش بیز ساده دارای بیشترین دقت پیش‌بینی برابر با ٪۹۵ (accuracy)	UCI (Cleveland)

ادامه جدول (۱): مروری بر کارهای گذشته

الگوریتم پیش‌بینی بیماری قلبی [۲]	ترکیب الگوریتم بهینه‌سازی گرگ خاکستری و شبکهٔ عصبی (ANN-GWO)	پارامترهای پزشکی	پیش‌بینی وضعیت بیمار در ۴ کلاس	بهره‌مندی از قابلیت‌های هر دو روش، افزایش سرعت همگرایی و بهبود دقت	شبکهٔ عصبی با عملکرد ۰/۰۰۴۳ ANN- (RMSE) GWO با عملکرد (RMSE) ۰/۰۰۳۶	UCI (Cleveland)
سیستم پیش‌بینی بیماری قلبی [۶]	ترکیب سیستم استنتاج فازی و الگوریتم ژنتیک	پارامترهای پزشکی	طبقه‌بندی افراد به دو دستهٔ بیمار و سالم	بهبود دقت طبقه‌بندی و پیاده‌سازی آسان	٪۸۶ (accuracy)	UCI
سیستم پیش‌بینی بیماری قلبی [۷]	ترکیب مجموعه‌های ناهموار و سیستم فازی نوع ۲ با الگوریتم کرم شب‌تاب	پارامترهای پزشکی	دسته‌بندی داده‌ها و بیان خطر بیماری قلبی	مدیریت داده‌های نویزدار و عدم قطعیت و بهبود عملکرد	٪۸۸/۳ (accuracy)	UCI
سیستم پیش‌بینی بیماری قلبی [۸]	شبکهٔ عصبی با الگوریتم یادگیری پس انتشار و الگوریتم ژنتیک	پارامترهای پزشکی	پیش‌بینی حضور یا عدم حضور بیماری قلبی	ابزاری مفید برای کمک به پزشکان	٪۹۹/۳۱ (accuracy)	مجموعه تست شامل ۴۳ ویژگی ورودی و یک خروجی از ۵ بیمار
سیستم پیش‌بینی بیماری قلبی [۹]	استفاده از طبقه‌بند فازی- ژنتیک برای پیش‌بینی بیماری قلبی از روی سیگنال ECG	پارامترهای سیگنال ECG	دسته‌بندی سیگنال‌های ECG به ۴ گروه متناظر ECG	بهبود دقت دسته‌بندی با اعمال الگوریتم ژنتیک روی طبقه‌بند فازی	٪۹۸/۶۷ (accuracy)	مجموعه داده‌ای شامل ۳۰۰ سیگنال ECG
سیستم پیش‌بینی بیماری قلبی [۱۰]	استفاده از شبکهٔ عصبی و الگوریتم ژنتیک	پارامترهای پزشکی	پیش‌بینی حضور یا عدم حضور بیماری قلبی	بهبود دقت پیش‌بینی با به‌کارگیری الگوریتم ژنتیک	٪۹۸ (accuracy)	مجموعه داده‌ای شامل ۵۰ نمونه که از انجمن قلب آمریکا تهیه شده است.

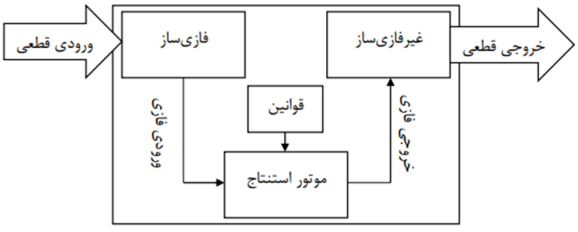
۳. روش‌های پیشنهادی

۱.۳. سیستم خبرهٔ فازی

یک سیستم خبرهٔ فازی[1] از سه بخش اصلی تشکیل شده است [۱]:

- موتور استنتاج: این بخش با پایگاه قوانین در ارتباط است و بر اساس آن دربارهٔ داده‌های ورودی تصمیم‌گیری فازی انجام می‌دهد.
- پایگاه داده: این بخش شامل تمام جزئیاتی است که برای طراحی یک سیستم فازی به آن‌ها نیاز می‌باشد.
- پایگاه دانش (پایگاه قوانین فازی): این بخش از سیستم خبره همان قوانین اگر- آنگاه فازی است که از طریق مشورت و مصاحبه با افراد خبره استخراج می‌شود.

معماری یک سیستم خبرهٔ فازی در شکل (۱) ارائه شده است.



شکل (۱): معماری یک سیستم خبرهٔ فازی

سیستم خبرهٔ فازی ابتدا ورودی‌ها را به شکل اعداد قطعی یا عبارات زبانی دریافت می‌کند و سپس بر اساس مجموعه‌های فازی آن‌ها را فازی می‌سازد و سپس موتور استنتاج برای هر

1. Fuzzy Expert System

ورودی بر اساس قوانین فازی تصمیم‌گیری فازی انجام می‌دهد و خروجی را به‌صورت فازی تولید می‌کند و به بخش غیرفازی‌ساز[1] می‌فرستد. در این بخش بر اساس روش غیرفازی‌ساز که برای سیستم طراحی شده است، خروجی فازی به عدد قطعی تبدیل، و در نهایت خروجی واقعی تولید می‌شود [۱].

۱.۱.۳. طراحی سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی

طبق تئوری طراحی سیستم فازی تمام پارامترها و جزئیاتی که برای طراحی یک سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی وجود دارد، به‌صورت کامل ارائه شده است [۱].

- موتور استنتاج: با پایگاه قوانین در ارتباط است و بر اساس آن دربارهٔ داده‌های ورودی تصمیم‌گیری فازی انجام می‌دهد.
- پایگاه داده: این بخش شامل تمام جزئیاتی است که برای طراحی یک سیستم فازی به آن‌ها نیاز می‌باشد.
- برای طراحی این سیستم از مدل استنتاج ممدانی استفاده شده است.
- برای این سیستم ۵ متغیر زبانی با مشورت پزشک متخصص قلب و عروق در نظر گرفته شده است که ۴ متغیر برای ورودی و یک متغیر مربوط به خروجی می‌باشد. سپس برای هرکدام از متغیرها عبارات زبانی در نظر گرفته شده که در جدول (۲) نشان داده شده است.

جدول (۲): عبارات زبانی متغیرهای ورودی و خروجی

متغیرهای زبانی	عبارات زبانی
فشار خون	پایین، طبیعی و بالا
کلسترول	طبیعی، بالا و خیلی بالا
قند خون	طبیعی و بالا
ضربان قلب	آهسته، طبیعی و سریع
بیماری قلبی (خروجی)	کم‌خطر و پرخطر

برای طراحی توابع عضویت از تابع عضویت گاوسی استفاده شده است؛ زیرا داده‌ها دارای توزیع نرمال هستند. پارامترهای توابع عضویت عبارات زبانی به بازهٔ هرکدام از عبارات زبانی بستگی دارد و بر اساس مقادیر به‌دست‌آمده به‌صورت تجربی تعیین شده است. تابع گاوسی در رابطهٔ (۱) نشان داده شده که در آن σ واریانس و c میانگین است.

$$f(x;\sigma.c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \quad (۱)$$

مقادیر پارامترهای سیستم خبرهٔ فازی با استفاده از مدل استنتاج ممدانی در جدول (۳) ارائه شده است.

پایگاه قوانین: این بخش از سیستم خبره همان قوانین اگر–آنگاه فازی می‌باشد. فرایند استخراج قوانین در این پژوهش به این صورت است که ابتدا تمام ترکیبات بین عبارات زبانی متغیرها تولید شده و به‌عنوان قوانین سیستم خبرهٔ فازی در نظر گرفته می‌شوند. روش مورد استفاده برای رفع مشکل پیچیدگی سیستم و استخراج قوانین مؤثر از پایگاه قوانین، روش جدول مراجعه[2] است. در این روش، آن قوانینی که توسط اکثر نمونه‌های مجموعهٔ داده فعال‌تر می‌شوند، انتخاب می‌گردند. قوانین انتخاب‌شده برای سیستم خبرهٔ پیش‌بینی بیماری قلبی در جدول (۴) ارائه شده است.

جدول (۳): پارامترهای سیستم خبرهٔ فازی

پارامتر	مقدار پارامتر
نوع مدل استنتاج فازی	ممدانی
روش AND	min
روش OR	max
روش Defuzzification	centroid
روش Implication	min
روش Aggregation	max
تعداد ورودی‌ها	۴
تعداد خروجی‌ها	۱
تعداد قوانین	۴

1. Defuzzifier

2. Look up table

جدول (۴): مجموعه قوانین سیستم خبرهٔ فازی

Rule #	Diagnostic fuzzy rules
R1	If blood pressure is normal and cholesterol is normal and blood sugar is normal and heart rate is normal then heart disease is low risk.
R2	If blood pressure is high then heart disease is high risk.
R3	If cholesterol is high then heart disease is high risk.
R4	If blood sugar is high then heart disease is high risk.

۲.۳. مدل ترکیبی فازی- ژنتیک (Fuzzy-GA)

به‌منظور تنظیم پارامترهای توابع عضویت سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی و همچنین بهبود نتایج آن مدل ترکیبی Fuzzy- GA طراحی و ارائه می‌شود.

مراحل طراحی این مدل شامل موارد زیر است [۱۵]:

۱. طراحی ساختار سیستم خبرهٔ فازی که شامل تعیین متغیرهای ورودی و خروجی، انتخاب نوع توابع عضویت آن‌ها و همچنین قوانین فازی است.

۲. شناسایی پارامترهای ضروری مانند پارامترهای توابع عضویت.

۳. ایجاد یک کروموزوم: برای تشکیل یک کروموزوم تمام پارامترهای شناسایی‌شده در مرحلهٔ قبل باید در یک آرایه مرتب شوند.

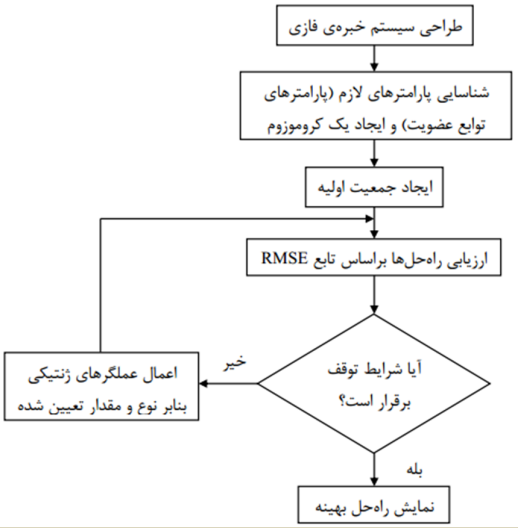
۴. طراحی یک تابع هدف: در این پژوهش به حداقل رساندن میانگین مربعات خطای سیستم (RMSE) به‌عنوان تابع هدف در نظر گرفته شده است. معیار RMSE خروجی‌های سیستم فازی را با خروجی‌های واقعی مجموعهٔ داده مقایسه می‌کند.

۵. انتخاب نوع و مقدار عملگرهای الگوریتم ژنتیک: در این مرحله، نوع و مقدار عملگرهای الگوریتم ژنتیک (انتخاب، ترکیب و جهش) تعیین می‌شود. در این پژوهش نرخ احتمال ترکیب و جهش به‌ترتیب برابر با ۰/۸ و ۰/۲ در نظر گرفته شده است.

۶. اعمال محدودیت‌های لازم به‌منظور ایجاد کروموزوم‌های معتبر.

۷. اجرای الگوریتم ژنتیک برای تکامل سیستم خبرهٔ فازی.

مراحل ایجاد مدل ترکیبی فازی- ژنتیک برای پیش‌بینی بیماری قلبی در قالب فلوچارتی در شکل (۲) نشان داده شده است.



شکل (۲): فلوچارت مدل ترکیبی Fuzzy- GA

۳.۳. مدل ترکیبی فازی- تکامل تفاضلی (Fuzzy-DE)

مدل ترکیبی پیشنهادی دیگر Fuzzy-DE است که به‌منظور بهینه‌سازی پارامترهای توابع عضویت سیستم خبرهٔ فازی طراحی شده و مراحل طراحی آن شامل موارد زیر است:

۱. طراحی ساختار سیستم خبرهٔ فازی.

۲. شناسایی پارامترهای لازم برای بهینه‌سازی سیستم. هدف در این پژوهش، تنظیم پارامترهای توابع عضویت سیستم خبرهٔ فازی است.

۳. ایجاد راه‌حلی اولیه بر اساس پارامترهای شناسایی‌شده.

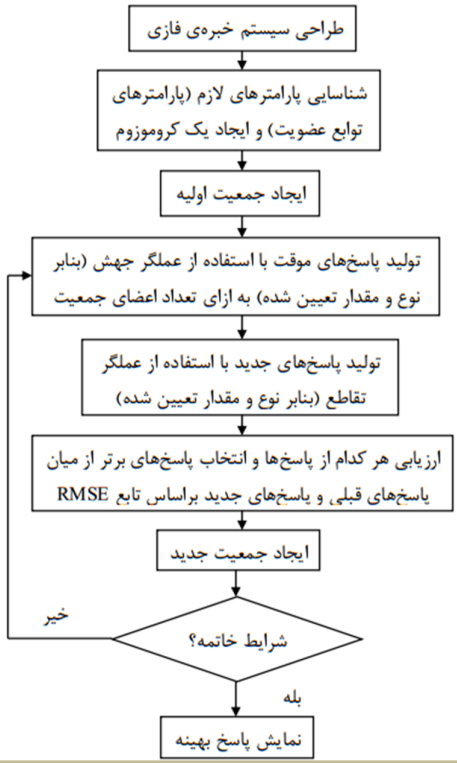
۴. طراحی یک تابع هدف مانند تابع RMSE برای سنجش میزان شایستگی راه‌حل‌ها.

۵. انتخاب مقادیر مناسب برای پارامترهای اساسی الگوریتم DE که شامل تعداد اعضای جمعیت، پارامتر مقیاس و احتمال ترکیب می‌باشند و مقادیر پارامترهای DE در این پژوهش به‌ترتیب برابر است با ۱۰۰، ۰/۸ و ۰/۹.

۶. تعیین محدوده‌ای مجاز برای راه‌حل‌ها.

۷. اجرای الگوریتم DE برای بهینه‌سازی سیستم خبرهٔ فازی.

مراحل اجرای مدل Fuzzy-DE در شکل (۳) نشان داده شده است.



شکل (۳): فلوچارت مدل ترکیبی Fuzzy-DE

۴.۳. مدل ترکیبی فازی- ژنتیک- تکامل تفاضلی (Fuzzy-GA-DE)

در این مقاله، به‌منظور بهره‌گیری از قابلیت‌های بهینه‌سازی هر دو الگوریتم تکاملی و هم‌افزایی کارایی هر روش، فرایند تکامل مدل در قالب اپراتورهای ژنتیکی در الگوریتم ژنتیک با جست‌وجوی هوشمندانهٔ الگوریتم تکامل تفاضلی بر مبنای فاصلهٔ اعضای جمعیت در فضای برداری به‌منظور تنظیم پارامترهای سیستم خبرهٔ فازی با یکدیگر ترکیب شده‌اند. مراحل طراحی مدل پیشنهادی Fuzzy-GA-DE شامل موارد زیر است:

۱. طراحی ساختار سیستم خبرهٔ فازی.

۲. شناسایی پارامترهای لازم برای بهینه‌سازی سیستم. هدف در این پژوهش، تنظیم پارامترهای توابع عضویت سیستم خبرهٔ فازی است.

۳. ایجاد راه‌حلی اولیه بر اساس پارامترهای شناسایی‌شده.

۴. طراحی یک تابع هدف مانند تابع RMSE برای سنجش میزان شایستگی راه‌حل‌ها.

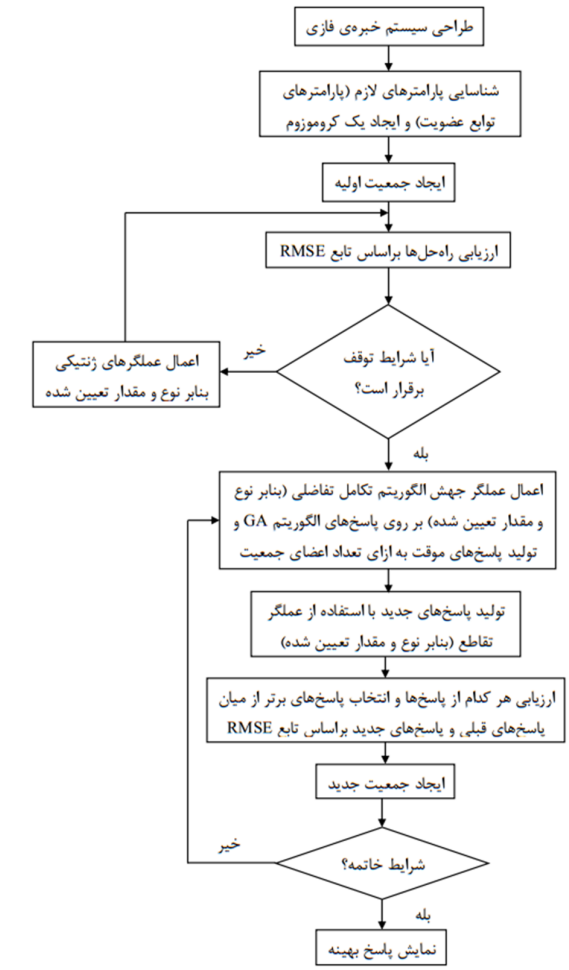
۵. انتخاب مقادیر مناسب برای پارامترهای اساسی الگوریتم DE که شامل تعداد اعضای جمعیت، پارامتر مقیاس و احتمال ترکیب هستند و مقادیر پارامترهای DE در این پژوهش به‌ترتیب برابر است با ۱۰۰، ۰/۸و ۰/۲.

۶. انتخاب مقادیر مناسب برای عملگرهای الگوریتم ژنتیک که احتمال عملگر ترکیب برابر با ۰/۸و احتمال عملگر جهش برابر با ۰/۲در نظر گرفته شده است.

۷. تعیین محدوده‌ای مجاز برای راه‌حل‌ها.

۸. اجرای الگوریتم GA-DE برای تنظیم پارامترهای سیستم خبرهٔ فازی.

مراحل اجرای مدل ترکیبی Fuzzy-GA-DE در شکل (۴) نشان داده شده است.



شکل (۴): فلوچارت مدل ترکیبی Fuzzy-GA-DE

۴. ارزیابی مدل‌های پیشنهادی

به‌منظور ارائهٔ یک سیستم کاربردی مفید و قابل اطمینان در دنیای واقعی و استفاده از آن به‌عنوان ابزار کمکی انسان لازم است دو گام مهم طراحی و ارزیابی سیستم اجرا شود. پس از اجرای گام مهم طراحی سیستم، ارزیابی آن با استفاده از داده‌های واقعی است که میزان ارزش سیستم را تعیین می‌کند و به سیستم طراحی‌شده جامهٔ عمل در تصمیم‌گیری‌های مهم را می‌پوشاند. صحت عملکرد سیستم‌های طراحی‌شده با استفاده از داده‌های واقعی با قابلیت اطمینان سیستم رابطه‌ای مستقیم دارد. به عبارت دیگر، هرچه کارایی سیستم‌های طراحی‌شده با اعمال داده‌های واقعی بیشتر باشد، آن سیستم مطمئن‌تر بوده و دارای پتانسیل بیشتری برای به‌کارگیری در دنیای واقعی و تصمیم‌گیری‌های حساس است. برای ارزیابی سیستم‌ها می‌توان از پایگاه داده‌هایی که از سازمان‌های مرتبط با حوزهٔ کاری سیستم‌ها فراهم می‌شوند، استفاده کرد.

۱.۴. پایگاه داده

برای ارزیابی کارایی سیستم طراحی‌شده برای پیش‌بینی خطر ابتلا به بیماری قلبی، از داده‌های بیماران واقعی استفاده شده است. این مجموعهٔ داده تهیه‌شده از بیمارستان پارسیان شامل داده‌های ۳۸۰ فرد است. ۳۲/۱٪ از داده‌های این مجموعهٔ داده مربوط به آقایان و درصد باقی‌ماندهٔ آن مربوط به خانم‌هاست. همچنین ۵۵/۲۶٪ از داده‌ها شامل افراد مبتلا به بیماری قلبی بوده و باقی‌ماندهٔ مجموعهٔ داده را افراد سالم تشکیل می‌دهند. علاوه بر این، مجموعهٔ داده دارای تشخیص‌های واقعی توسط افراد خبرهٔ پزشکی است که امکان مقایسهٔ خروجی‌های سیستم را با خروجی‌های واقعی فراهم می‌کند.

۲.۴. روش‌های ارزیابی بر اساس تحلیل منحنی ROC

معیار تحلیل منحنی ROC به‌دلیل نمایش موازنهٔ بین مزایا و هزینه‌های سیستم بر اساس تغییرات آستانهٔ تصمیم، معیاری محبوب و قابل اطمینان محسوب می‌شود. منحنی ROC تشخیص‌های درست سیستم را در مقابل تشخیص‌های نادرست آن ترسیم می‌کند. میزان کارایی و عملکرد سیستم‌های طبقه‌بندی به‌صورت نقطه‌ای روی منحنی ROC نمایش داده می‌شود. عواملی مانند تغییرات آستانه و توزیع داده‌ها در تغییر مکان این نقطه مؤثر است [۱۶ و ۱۷].

با استفاده از معیار ارزیابی تحلیل منحنی ROC، ماتریسی با عنوان ماتریس ابهام[1] ارائه می‌شود که شامل معیارهای مهم و ارزشمندی است که این معیارها عملکرد سیستم مورد ارزیابی را توصیف می‌کنند [۱۶ و ۱۷]. ساختار این ماتریس در جدول (۵) نشان داده شده است.

جدول (۵): ماتریس ابهام

		کلاس پیش‌بینی	
		مثبت	**منفی**
کلاس واقعی	مثبت	TP	FN
	منفی	FP	TN

همان طور که در جدول (۵) مشخص است، TP نشان‌دهندهٔ نمونه‌های مثبتی است که سیستم هم آن‌ها را به‌عنوان مثبت طبقه‌بندی کرده، FN نشان‌دهندهٔ نمونه‌های مثبتی است که سیستم آن‌ها را به‌عنوان منفی طبقه‌بندی کرده، FP نشان‌دهندهٔ نمونه‌های منفی است که سیستم آن‌ها را به‌عنوان مثبت طبقه‌بندی کرده و TN نشان‌دهندهٔ نمونه‌های منفی است که سیستم هم آن‌ها را به‌عنوان منفی طبقه‌بندی کرده است [۱۶و۱۷]. با استفاده از چهار پارامتری که از این ماتریس به دست می‌آیند، سه معیار ارزیابی accuracy، sensitivity و specificity استخراج شده که به‌ترتیب در رابطه‌های (۲)، (۳) و (۴) بیان شده است.

$$\text{accuracy}=\frac{TN+TP}{TN+TP+FN+FP} \qquad (۲)$$

$$\text{sensitivity}=\frac{TP}{TP+FP} \qquad (۳)$$

$$\text{specificity}=\frac{TN}{TN+FN} \qquad (۴)$$

۳.۴. روش اعتبارسنجی ده بخشی(10-Fold Cross Validation)

در این روش، ابتدا پایگاه داده به k بخش (در این پژوهش k=10) گسسته تقسیم می‌شود. سپس سیستم با k-1 بخش

1. Confusion Matrix

آموزش دیده و با ۱ بخش تست می‌شود. فرایند آموزش و تست به تعداد k مرتبه انجام می‌گیرد. در نهایت برای تعیین دقت و صحت سیستم، میانگین نتایج این k بخش در نظر گرفته می‌شود [۱۸ و ۱۹].

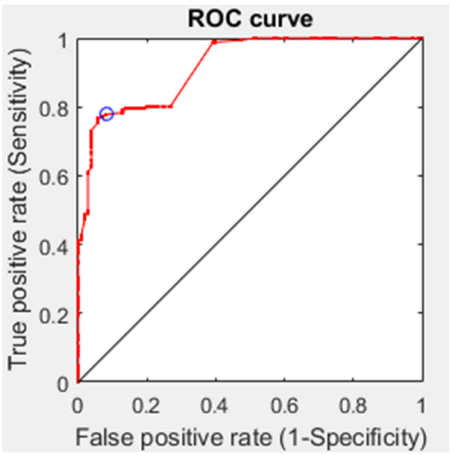
۴.۴. نتایج ارزیابی مدل‌های پیشنهادی

بر اساس ارزیابی‌های انجام‌شده ماتریس ابهام سیستم خبرهٔ فازی در جدول (۶) نشان داده شده است.

جدول (۶): ماتریس ابهام سیستم خبرهٔ فازی

		کلاس پیش‌بینی	
		مثبت	منفی
کلاس واقعی	مثبت	(TP) ۱۳۳	(FN) ۱۷
	منفی	(FP) ۳۸	(TN) ۱۹۲

طبق نتایج به‌دست‌آمده در جدول (۶)، سیستم خبرهٔ فازی در پیش‌بینی بیماری قلبی تا حدودی عملکرد خوبی داشته و نمونه‌های تقریباً زیادی را در دسته‌های درست طبقه‌بندی کرده است. همچنین منحنی ROC که بیانگر مزایا و معایب سیستم خبرهٔ فازی می‌باشد، در شکل (۵) نشان داده شده است.



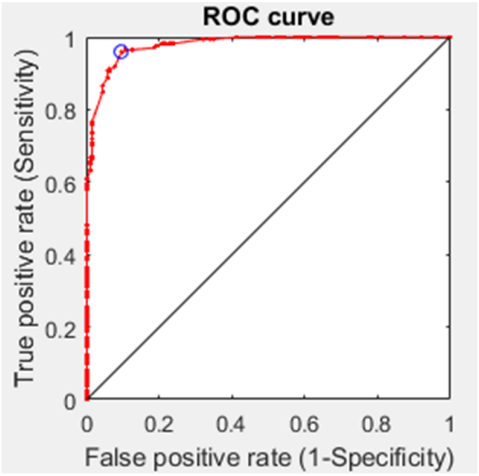
شکل (۵): منحنی ROC سیستم خبرهٔ فازی

از آنجایی که سطح زیر منحنی ROC (AUC) سیستم خبرهٔ فازی در شکل (۵) زیاد است، نشان‌دهندهٔ این است که سیستم عملکرد نسبتاً خوبی داشته اما با سیستم ایدئال هنوز فاصله دارد که برای رفع آن الگوریتم‌های GA و DE روی آن اعمال می‌شوند. پس از اعمال الگوریتم GA روی سیستم خبرهٔ فازی و ارزیابی آن، ماتریس ابهام مدل Fuzzy-GA در جدول (۷) ارائه شده است.

جدول (۷): ماتریس ابهام مدل ترکیبی Fuzzy-GA

		کلاس پیش‌بینی	
		مثبت	منفی
کلاس واقعی	مثبت	(TP) ۱۶۴	(FN) ۲۰
	منفی	(FP) ۷	(TN) ۱۸۹

بر اساس نتایج به‌دست‌آمده در جدول (۷)، کارایی مدل Fuzzy-GA نسبت به سیستم خبرهٔ فازی افزایش زیادی داشته زیرا دسته‌بندی‌های درست مدل Fuzzy-GA، (TP,TN)، نسبت به مدل فازی افزایش پیدا کرده است. منحنی ROC مدل هایبریدی Fuzzy-GA در شکل (۶) نشان داده شده است.



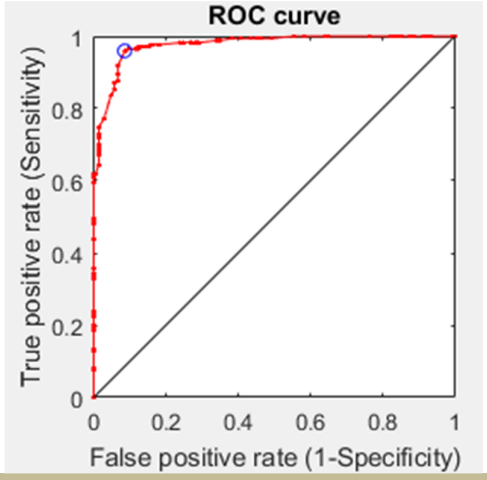
شکل (۶): منحنی ROC مدل ترکیبی Fuzzy-GA

با توجه به منحنی ROC مدل Fuzzy-GA در شکل (۶)، AUC این مدل نسبت به AUC سیستم خبرهٔ فازی افزایش یافته که بیانگر بهبود عملکرد و نتایج سیستم است و سیستم به یک سیستم ایده ال نزدیک‌تر شده است. پس از اعمال الگوریتم DE روی سیستم خبرهٔ فازی و ارزیابی آن، ماتریس ابهام مدل Fuzzy-DE در جدول (۸) ارائه شده است.

جدول (۸): ماتریس ابهام مدل ترکیبی Fuzzy-DE

		کلاس پیش‌بینی	
		مثبت	منفی
کلاس واقعی	مثبت	(TP) ۱۶۴	(FN) ۱۸
	منفی	(FP) ۷	(TN) ۱۹۱

با توجه به نتایج ارائه‌شده در جدول (۸)، مدل Fuzzy-DE، علاوه بر بهبود نتایج سیستم خبرۀ فازی، نتایج مدل Fuzzy-GA را نیز بهبود می‌دهد؛ زیرا دسته‌بندی‌ها در سیستم به تعداد بیشتری نسبت به دو مدل قبل افزایش یافته است. منحنی ROC مدل هایبریدی Fuzzy-DE در شکل (۷) ارائه شده است.



شکل (۷): منحنی ROC مدل ترکیبی Fuzzy-DE

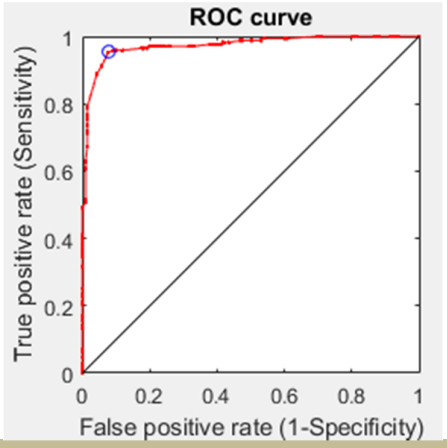
با توجه به شکل (۷)، AUC مدل Fuzzy-DE نسبت به AUC سیستم خبرۀ فازی افزایش یافته و این به معنای آن است که مدل Fuzzy-DE توانسته است به‌خوبی سیستم خبرۀ فازی و نتایج آن را بهبود بخشد.

پس از ترکیب دو الگوریتم ژنتیک و تکامل تفاضلی و اعمال آن‌ها روی سیستم خبرۀ فازی، مدل Fuzzy-GA-DE مورد ارزیابی قرار گرفته که نتایج و ماتریس ابهام آن در جدول (۹) نشان داده شده است.

جدول (۹): ماتریس ابهام مدل ترکیبی Fuzzy-GA-DE

		کلاس پیش‌بینی	
		مثبت	منفی
کلاس واقعی	مثبت	(TP) ۱۶۳	(FN) ۱۶
	منفی	(FP) ۸	(TN) ۱۹۳

با توجه به نتایج به‌دست‌آمده در جدول (۹)، کارایی مدل هایبریدی Fuzzy-GA-DE و همچنین دسته‌بندی‌های درست سیستم نسبت به هر سه مدل قبل افزایش یافته است. همان طور که انتظار می‌رفت، از ترکیب دو الگوریتم تکاملی به نتایج بهتری دست یافته می‌شود. نتایج موجود در جدول (۹) دستیابی به هدف مقاله و بهره‌گیری از قابلیت‌های هر دو الگوریتم تکاملی به‌منظور بهینه‌سازی و بهبود نتایج مدل‌های Fuzzy-GA و Fuzzy-DE را بیان می‌کند. منحنی ROC مدل هایبریدی Fuzzy-GA-DE در شکل (۸) نشان داده شده است.



شکل (۸): منحنی ROC مدل ترکیبی Fuzzy- GA- DE

با توجه به شکل (۸)، AUC مدل Fuzzy-GA-DE نسبت به AUC سیستم خبرۀ فازی افزایش یافته و بیانگر این مسئله است که مدل Fuzzy-GA-DE علاوه بر بهینه‌سازی سیستم خبرۀ فازی قادر به بهبود مدل‌های هایبریدی Fuzzy-GA و Fuzzy-DE می‌باشد.

عملکرد سیستم خبرۀ فازی و مدل‌های هایبریدی Fuzzy-GA، Fuzzy-DE و Fuzzy-GA-DE بر اساس معیارهای ارزیابی AUC، sensitivity و specificity و روش اعتبارسنجی ده بخشی در جدول (۱۰) نشان داده شده است.

با توجه به جدول (۱۰) هر دو الگوریتم‌های تکاملی اعمال‌شده روی سیستم خبرۀ فازی برای پیش‌بینی بیماری قلبی، نتایج این سیستم را از نظر معیارهای ارزیابی AUC، sensitivity، specificity و MSE بهبود بخشیده‌اند. همان طور که در جدول (۱۰) مشخص است، حداکثر AUC مربوط به مدل ترکیبی Fuzzy-GA برابر با ۹۷/۹۳٪ می‌باشد. مدل ترکیبی Fuzzy-GA-DE با سطح زیر منحنی ROC نزدیک به مدل‌های پیشنهادی و از نظر معیار ارزیابی specificity کارایی بیشتری نسبت به سایر مدل‌های پیشنهادی دارد.

جدول (۱۰): نتایج ارزیابی مدل‌های پیشنهادی برای پیش‌بینی بیماری قلبی

	AUC(%)	CI(95%)	Sensitivity (%)	Specificity (%)	MSE	MSE (train)	MSE (test)
Fuzzy	۹۱/۷۷	[۰/۸۹,۰/۹۵]	۷۷/۷۷	۹۱/۸۶	۰/۱۷	–	–
Fuzzy-GA	۹۳	[۰/۹۷,۰/۹۹]	۹۵/۹۱	۹۰/۴۳	۰/۰۶	۰/۰۷	۰/۰۶
Fuzzy-DE	۹۷/۶۷	[۰/۹۶,۰/۹۹]	۹۵/۹۱	۹۱/۳۹	۰/۰۶	۰/۰۷	۰/۰۴
Fuzzy-GA-DE	۹۷/۳۲	[۰/۹۶,۰/۹۹]	۹۵/۳۲	۹۲/۳۴	۰/۰۶	۰/۰۶	۰/۰۴

۵. مقایسهٔ مدل‌های پیشنهادی و مدل‌های ارائه‌شده در پژوهش‌های قبلی

بنا بر نتایج به‌دست‌آمده در جدول (۱۱)، سیستم خبرهٔ فازی پیشنهادی برای پیش‌بینی بیماری قلبی، بیشترین دقت طبقه‌بندی نسبت به سایر مدل‌های مبتنی بر فازی ارائه‌شده در پژوهش‌های اخیر است. نتایج به‌دست‌آمده از این سیستم دارای قابلیت درک و تفسیر بالاست و همچنین به‌دلیل به‌کارگیری حداقل تعداد ورودی و حداقل تعداد قوانین در طراحی این سیستم، سیستم خبرهٔ فازی ارائه‌شده بسیار ساده و کم‌هزینه است. از آنجایی که قوانین فازی به کار گرفته‌شده در طراحی این سیستم، بر اساس بیشترین میزان سازگاری با مجموعهٔ داده انتخاب شده است، استفاده از آن‌ها در سیستم خبرهٔ فازی پیشنهادی بسیار مؤثر می‌باشد. سیستم خبرهٔ فازی ارائه‌شده برای پیش‌بینی بیماری قلبی از داده‌های کاملاً قابل دسترس و کم‌هزینه استفاده می‌کند. به‌طور مختصر می‌توان از مزایای سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی بدین موارد اشاره کرد:

۱. استفاده از داده‌های قابل دسترس و کم‌هزینه برای بیماران؛

۲. دقت بیشتر نسبت به سایر مدل‌های موجود مبتنی بر فازی؛

۳. قابلیت درک و تفسیر بالا؛

۴. قابلیت سادگی بالاتر با بهره‌مندی از کمترین تعداد ورودی، خروجی و قوانین.

همچنین مدل‌های ترکیبی پیشنهادی مبتنی بر فازی نسبت به مدل‌های ترکیبی فازی- ژنتیک در [۶] و فازی- الگوریتم کرم شب‌تاب در [۷] دارای دقت بیشتر و عملکرد بهتری هستند. ورودی‌های همهٔ این مدل‌ها پارامترهای پزشکی می‌باشند. مدل ترکیبی Fuzzy-GA پیشنهادی نسبت به سایر مدل‌های ترکیبی تکاملی ارائه‌شده در این پژوهش و پژوهش‌های مشابه از عملکرد بهتر و دقت بیشتری برخوردار است. مدل‌های ترکیبی Fuzzy-DE و Fuzzy-GA-DE پیشنهادی دارای روند همگرایی مناسب هستند.

از آنجایی که سیستم‌های تشخیص پزشکی باید به‌عنوان ابزار کمکی در اختیار پزشکان قرار گیرد، مستلزم سهولت در استفاده از آن‌ها می‌باشند و همان طور که مدل‌های پیشنهادی در این مقاله مبتنی بر سیستم استنتاج فازی هستند، از قابلیت درک و تفسیر بالا برخوردارند و در طراحی چنین سیستم‌هایی باید موازنه‌ای بین قابلیت درک و همچنین دقت این سیستم‌ها در نظر گرفت.

مدل‌های ارائه‌شده در پژوهش‌های اخیر که مبتنی بر شبکه‌های عصبی هستند، دقت بیشتری نسبت به مدل‌های فازی دارند اما قابل درک و تفسیر توسط پزشکان نیستند. از جمله قابلیت سیستم‌های فازی قابلیت محاوره با کاربران و درک و فهم نتایج حاصل از استنتاج فازی در قالب عبارت‌های زبانی و پایگاه دانش مبتنی بر عبارت‌های زبانی، توسط پزشکان متخصص است که این قابلیت در ساختار شبکهٔ عصبی موجود نیست. این ویژگی در سیستم‌های فازی منجر به رشد روند استفاده از مدل‌های هوشمند فازی به‌عنوان ابزار هوشمند تصمیم یار در دههٔ اخیر شده است [۲۰ـ۲۳].

جدول (۱۱): مقایسهٔ مدل‌های پیشنهادی و مدل‌های ارائه‌شده در مقالات مشابه

مدل پیش‌بینی بیماری قلبی	ورودی‌ها	خروجی‌ها	مزایا	عملکرد (معیار اندازه‌گیری)	مجموعه دادهٔ ارزیابی
سیستم خبرهٔ فازی با استفاده از مدل استنتاج ممدانی [۱۳]	پارامترهای پزشکی	پیش‌بینی خطر بیماری قلبی در افراد و دسته‌بندی گروه‌های خطر	قابلیت درک آسان قوانین و نیاز به پارامترهای کلینیکی قابل دسترس و کم‌هزینه	٪۸۴/۲۰ (accuracy)	مجموعه داده‌ای شامل ۵۰۰ رکورد که از بیمارستان AMRI تهیه شده است.
سیستم استنتاج فازی ممدانی و به‌کارگیری قوانین فازی وزن‌دار [۱۴]	پارامترهای پزشکی	پیش‌بینی خطر بیماری قلبی در دو دسته	بهبود عملکرد سیستم نسبت به شبکهٔ عصبی، تولید خودکار قوانین فازی و وزن‌دار کردن آن‌ها	٪۵۷/۸۵ (accuracy)	UCI
شبکهٔ عصبی پس‌انتشار [۱۲]	پارامترهای پزشکی	دسته‌بندی بیماران به ۴ دسته	عملکرد بهتر نسبت به روش‌های موجود	٪۹۲ (accuracy)	UCI (Cleveland)
ترکیب الگوریتم بهینه‌سازی گرگ خاکستری و شبکهٔ عصبی (ANN-GWO) [۲]	پارامترهای پزشکی	پیش‌بینی وضعیت بیمار در ۴ کلاس	بهره‌مندی از قابلیت‌های هر دو روش، افزایش سرعت همگرایی و بهبود دقت	شبکهٔ عصبی ۰/۰۰۴۳ (RMSE) ANN-۰/۰۰۳۶GWO (RMSE)	UCI (Cleveland)
ترکیب سیستم استنتاج فازی و الگوریتم ژنتیک [۶]	پارامترهای پزشکی	طبقه‌بندی افراد به دو دستهٔ بیمار و سالم	بهبود دقت طبقه‌بندی و پیاده‌سازی آسان	٪۸۶ (accuracy)	UCI
ترکیب مجموعه‌های ناهموار و سیستم فازی نوع ۲ با الگوریتم کرم شب‌تاب [۷]	پارامترهای پزشکی	دسته‌بندی داده‌ها و بیان خطر بیماری قلبی	مدیریت داده‌های نویزدار و عدم قطعیت و بهبود عملکرد	٪۸۸/۳ (accuracy)	UCI
شبکهٔ عصبی با الگوریتم یادگیری پس‌انتشار و الگوریتم ژنتیک [۸]	پارامترهای پزشکی	پیش‌بینی حضور یا عدم حضور بیماری قلبی	ابزاری مفید برای کمک به پزشکان	٪۹۹/۳۱ (accuracy)	مجموعه تست شامل ۴۳ ویژگی ورودی و یک خروجی از ۵ بیمار
استفاده از شبکهٔ عصبی و الگوریتم ژنتیک [۱۰]	پارامترهای پزشکی	پیش‌بینی حضور یا عدم حضور بیماری قلبی	بهبود دقت پیش‌بینی با به‌کارگیری الگوریتم ژنتیک	٪۹۸ (accuracy)	مجموعه داده‌ای شامل ۵۰ نمونه که از انجمن قلب آمریکا تهیه شده است.

ادامه جدول (۱۱): مقایسهٔ مدل‌های پیشنهادی و مدل‌های ارائه‌شده در مقالات مشابه					
سیستم خبرهٔ فازی با استفاده از مدل استنتاج ممدانی [۱]	پارامترهای پزشکی شامل فشار خون، کلسترول، قند خون و ضربان قلب	پیش‌بینی بیماری قلبی در دو دسته	قابلیت درک و تفسیر بالا، استفاده از حداقل تعداد ورودی و تعداد قانون در طراحی سیستم، انتخاب سازگارترین قوانین با مجموعهٔ داده از بین مجموعه قوانین و نیاز به داده‌های قابل دسترس و کم‌هزینه	(AUC) ٪۹۱/۷۷ [۰/۹۵ و ۰/۸۹] (CI)	مجموعه داده‌ای شامل ۳۸۰ نمونه
مدل ترکیبی Fuzzy- GA (مدل پیشنهادی)	پارامترهای جدول ۲	پیش‌بینی بیماری قلبی در دو دسته	بهبود دقت پیش‌بینی، قابلیت درک و تفسیر بالا و پیاده‌سازی کم‌هزینه	(AUC) ٪۹۷/۹۳ [۰/۹۹ و ۰/۹۷] (CI)	مجموعه داده‌ای شامل ۳۸۰ نمونه
مدل ترکیبی Fuzzy-DE (مدل پیشنهادی)	پارامترهای جدول ۲	پیش‌بینی بیماری قلبی در دو دسته	روند همگرایی مناسب و بهبود کیفیت فرایند اکتشاف و بهره‌برداری	(AUC) ٪۹۷/۶۷ [۰/۹۹ و ۰/۹۶] (CI)	مجموعه داده‌ای شامل ۳۸۰ نمونه
مدل ترکیبی Fuzzy-GA-DE (مدل پیشنهادی)	پارامترهای جدول ۲	پیش‌بینی بیماری قلبی در دو دسته	بهبود دقت پیش‌بینی نسبت به مدل Fuzzy-DE پیشنهادی و روند همگرایی مناسب	(AUC) ٪۹۷/۳۲ [۰/۹۹ و ۰/۹۶] (CI)	مجموعه داده‌ای شامل ۳۸۰ نمونه

۶. نتیجه‌گیری و پیشنهادات آتی

در مقالهٔ حاضر، به‌منظور تنظیم پارامترهای توابع عضویت سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی ارائه‌شده در [۱] و بهبود نتایج آن الگوریتم‌های ژنتیک و تکامل تفاضلی روی این سیستم اعمال شده‌اند. مدل‌های هایبریدی Fuzzy-GA و Fuzzy-DE نتایج سیستم خبرهٔ فازی را به‌طور چشمگیری افزایش داده‌اند. علاوه بر طراحی مدل‌های هایبریدی Fuzzy-GA و Fuzzy-DE به‌منظور بهره‌گیری از قابلیت‌های هر دو الگوریتم تکاملی ژنتیک و تکامل تفاضلی، این دو الگوریتم با یکدیگر ترکیب شده و روی سیستم خبرهٔ فازی اعمال می‌شوند. مدل هایبریدی Fuzzy-GA-DE کارایی مدل‌های پیشنهادی فازی، Fuzzy-GA و Fuzzy-DE را بهبود بخشیده است. برای ارزیابی مدل‌های پیشنهادی از مجموعه دادهٔ واقعی بیماران شامل ۳۸۰ نمونه که از بیمارستان پارسیان تهیه شده است، استفاده می‌شود. به‌عنوان کار آینده می‌توان از الگوریتم‌های تکاملی و بهینه‌سازی به‌منظور استخراج و تنظیم قوانین فازی سیستم خبرهٔ فازی برای پیش‌بینی بیماری قلبی استفاده کرد.

مراجع

[1] Akhoondi, R., and Hosseini, R., "*A fuzzy expert system for prognosis of the risk of development of heart disease*", Journal of Advances in Computer Research, Vol. 7, Issue 2, pp. 101-114, 2016.

[2] Turabieh, H., "*A hybrid ANN-GWO algorithm for prediction of heart disease*", American Journal of Operations Research, Vol. 6, No. 2, pp. 136-146, 2016.

[3] Dey, A., Singh, J. and Singh, N., "*Analysis of supervised machine learning algorithms for heart disease prediction with reduced number of attributes using principal component analysis*", International Journal of Computer Applications, Vol. 140, No. 2, pp. 27-31. 2016.

[4] Narain, R., Saxena, S. and Goyal, AK., "*Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network-based approach*", Patient Preference and Adherence, Vol. 10, pp. 1259-1270, 2016.

[5] Puyalnithi, T. and Viswanatham, VM, "*Preliminary cardiac disease risk prediction based on medical and behavioural data set using supervised machine learning techniques*", Indian Journal of Science and Technology, Vol. 9, No. 31, 2016.

[6] Santhanam, T. and Ephzibah, EP., "*Heart disease prediction using hybrid genetic fuzzy model*", Indian Journal of Science and Technology, Vol. 8, No. 9, pp. 797–803, 2015.

[7] Long, NC., Meesad, P. and Unger, H., "*A highly accurate firefly based algorithm for heart disease prediction*", Journal Expert Systems with Applications, Vol. 42, Issue 21, pp. 8221-8231, 2015.

[8] Verma, T. and Srivastava, DRK., "*Artificial neural networks-based heart disease predictive approach*", International Journal of Application or Innovation in Engineering & Management, Vol. 4, Issue 3, 2015.

[9] Vafaie, MH., Ataei, M. and Koofigar, HR., "*Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals*", Journal Biomedical Signal Processing and Control, Vol. 14, pp. 291-296, 2014.

[10] Waghulde, NP. and Patil, NP., "*Genetic neural approach for heart disease prediction*", International Journal of Advanced Computer Research, Vol. 4, No. 3, pp. 778-784, 2014.

[11] Ishtake, SH. and Sanap, SA., "*Intelligent heart disease prediction system using data mining techniques*", International Journal of Healthcare & Biomedical Research, Vol. 1, No. 3, pp. 94-101, 2013.

[12] Milli, NAl., "*Backpropogation neural network for prediction of heart disease*", Journal of Theoretical and Applied Information Technology, Vol. 56, No. 1, pp. 131-135, 2013.

[13] Pal, D., Mandana, KM., Pal, S., Sarkar, D. and Chakraborty, C., "*Fuzzy expert system approach for coronary artery disease screening using clinical parameters*", Journal Knowledge- Based Systems, Vol. 36, pp. 162-174, 2012.

[14] Anooj, PK., "*Clinical decision support system: risk level prediction of hear disease using weighted fuzzy rules*", Journal of King Saud University-Computer and Information Sciences, Vol. 24, Issue 1, pp. 27-40, 2012.

[15] Cordona, O., Gomideb, F., Herreraa, F., Hoffmann, F. and Magdalenad, L., "*Ten years of genetic fuzzy systems: current framework and new trends*", Fuzzy Sets and Systems, Vol. 141, Issue 1, pp. 5-31, 2004.

[16] Bradley, AP., "*The use of the area under the ROC curve in the evaluation of machine learning algorithms*", Pattern Recognition, Vol. 30, Issue 7, pp. 1145-1159, 1997.

[17] Fawcett, T., *An introduction to ROC analysis. Pattern Recognition Letters*, Vol. 27, Issue 8, 861-874, 2006.

[18] Jiawei, H. and Micheline, K., *Data Mining Concepts and Techniques*, Chapter (5), 2006.

[19] Pang-Ning, T., Michael, S. and Vipin, K., *Introduction to Data Mining, Pearson Addison-Wesley*, Chapter (5), 2006.

[20] Hosseini, R., Qanadli, SD., Barman, S., Mazinani, M., Ellis, T. and Dehmeshki, H., An automatic approach for learning and tuning Gaussian interval type-2 fuzzy membership functions applied to lung CAD classification system, *IEEE Transactions on Fuzzy Systems*, Vol. 20, Issue 2, pp. 224-234, 2001.

[21] Hosseini, R., Dehmeshki, J., Barman, S., Mazinani, M. and Qanadli, S., "*A genetic type-2 fuzzy logic system for pattern recognition in computer aided detection systems*", IEEE International Conference on *Fuzzy Systems:* 1-7, 2010.

[22] Khezri, R., Hosseini, R. and Mazinani, M., "*A Fuzzy Rule-Based Expert System for the Prognosis of The Risk of Development of The Breast Cancer*", International Journal of Engineering Transactions A: Basics, Vol. 27, Issue 10, pp. 1557-1564, 2014.

[23] AmirHosseini, B., Hosseini, R. and Mazinani, M., "*A Hybrid Fuzzy-GA Approach Applied to An Expert System for Diagnosis of Liver Tumo*r", Journal of Soft Computing and Information Technology, Vol. 5, No. 1, pp. 45-52, 2017.