

پیش‌بینی و تعیین هوشمند دُز صحیح انسولین در بیماران دیابتی براساس تشخیص فازی بیماری دیابت

محمد فیوضی*^۱، جواد حدادنیا^۲، نسرین ملانیا^۳

چکیده

مقدمه: تشخیص به موقع بیماری دیابت به‌طور چشم‌گیری صدمات و آسیب‌های ناشی از این بیماری را در جامعه کاهش می‌دهد. دیابت بیماری است که علاوه بر پیشگیری، نیاز به مراقبت‌های فراوانی از قبیل پیش‌بینی صحیح میزان نوسانات سطح قند خون دارد. از مهم‌ترین عوارض این بیماری می‌توان به بیهوشی، کما و حتی مرگ اشاره کرد. امروزه در این بیماران، تعیین صحیح دُز انسولین براساس تجربه و دانش پزشکان در کنار تعامل بیماران با آن‌ها مشخص می‌شود، هر چند که وجود خطاهای انسانی اجتناب‌ناپذیر است.

روش‌ها: در این تحقیق ۱۲۴ بیمار و ۱۸۸ فرد سالم مشکوک به بیماری براساس ۲۱ ویژگی (۷ ویژگی به‌منظور تشخیص، ۱۴ ویژگی به‌منظور پیش‌بینی دُز انسولین) مورد مطالعه و بررسی قرار گرفتند، سپس سیستمی ارائه شد که ابتدا شناسایی یا تشخیص بیماری را انجام دهد، بعد برای افراد بیمار، مهم‌ترین داروی این بیماران، یعنی دُز انسولین مشخص کند. سیستم پیشنهادی دارای دو مرحله (تشخیص بیماری و پیش‌بینی) و چندین زیر سیستم می‌باشد. در مرحله‌ی تشخیص بیماری، زیر سیستم‌هایی از قبیل سیستم فازی (Fuzzy) به‌منظور برآورد صحیح پیشرفت بیماری در بیماران، درخت تصمیم‌گیری (D-T) به‌منظور تهیه قوانین در سیستم فازی (فرآیند نگاشت فضای ویژگی (افراد) به خروجی (نتیجه‌ی تشخیص)) استفاده شده است. همچنین در مرحله‌ی پیش‌بینی دُز انسولین از الگوریتم‌های کاوشی (BPSO) به‌منظور انتخاب بهترین ویژگی‌ها، الگوریتم‌های طبقه‌بندی (SVM) به‌منظور طبقه‌بندی ویژگی‌های مؤثر از غیر مؤثر و سیستم‌های انطباقی مصنوعی فازی - عصبی (ANFIS) برای پیش‌بینی نهایی داروی بیماران استفاده شده است.

یافته‌ها: سیستم پیشنهادی براساس بهترین ویژگی‌ها در بانک داده تهیه شده در قالب ترکیب و تعامل موفق شد به دقت ۹۵/۱٪ دست یابد، که البته در هنگام مقایسه با سایر روش‌های معمول به سرعت قابل توجه و عملکرد مناسب آن و البته دقت بالای آن پی می‌بریم. **نتیجه‌گیری:** نتایج به‌دست آمده به‌میزان قابل توجهی نسبت به تحقیقات قبلی بهبود یافت. همچنین در مقایسه با نتایج پزشکان، نشان دهنده‌ی عملکرد خوب در صحت پیش‌بینی سری زمانی غلظت قند خون است، چرا که سیستم پیشنهادی موفق شد، سطح قند خون تا ۴۸ ساعت آینده را پیش‌بینی نماید.

واژگان کلیدی: دیابت، تشخیص، پیش‌بینی، سیستم‌های فازی، الگوریتم طبقه‌بندی، سیستم فازی عصبی انطباقی

۱- مرکز تحقیقات سالمند شناسی، دانشگاه علوم پزشکی سبزوار، سبزوار، ایران

۲- مرکز تحقیقات فناوری‌های نوین پزشکی، دانشکده‌ی مهندسی برق، دانشگاه حکیم سبزواری، سبزوار، ایران

۳- بخش بیوشیمی، گروه زیست شناسی، دانشکده‌ی علوم پایه، دانشگاه حکیم سبزواری، سبزوار، ایران

* **نشانی:** خراسان رضوی، شهرستان سبزوار، توحیدشهر، دانشگاه حکیم سبزواری، دانشکده‌ی مهندسی برق و کامپیوتر، بخش مهندسی پزشکی، کدپستی: ۹۶۱۷۹۷۶۴۸۷، تلفن: ۰۵۷۱۴۴۱۰۱۰۴، نمایر: ۰۵۷۱۴۴۱۰۱۰۴، پست الکترونیک: m mohammad.fiuzy@yahoo.com

مقدمه

دیابت یا مرض قند یکی از شایع‌ترین بیماری‌های شناخته شده در دنیا می‌باشد. براساس گزارش‌ها حدود ۲۲۰ میلیون نفر در جهان از این بیماری رنج می‌برند [۱-۳]. دیابت عمدتاً از طریق آزمایش‌های قند خون (Plasma, Glucose, Fasting) تشخیص داده می‌شود که به ۳ دسته تقسیم می‌شود. دیابت نوع یک یا دیابت وابسته به انسولین (IDDM) بیشتر در سن پایین و کودکان دیده می‌شود. دیابت نوع دو یا دیابت غیر وابسته به انسولین (NIDDM) در ۹۰ تا ۹۵ درصد بیماران دیابتی مشاهده می‌شود [۳]. در این دسته انسولین ترشح می‌شود، ولی بدن در برابر مصرف انسولین از خود مقاومت نشان می‌دهد. دیابت نوع سوم بیشتر در زنان باردار دیده می‌شود که عمدتاً پس از بارداری به دیابت نوع دو تبدیل می‌شود [۴]. در چند سال اخیر توجه قابل ملاحظه‌ای به روش‌های داده کاوی پزشکی شده است. به صورت تئوری ثابت شده است که یک گروه از کلاس‌ها، بندها، نتایج دقیق‌تری از بهترین آن‌ها ارائه می‌کنند [۵-۷]. می‌توان دریافت که هر دسته داده و روش، دارای توانایی‌ها و ضعف‌های مربوط به خود بوده که ممکن است در یک سیستم داده کاوی خوب عمل کند و در جایی دیگر نتیجه‌ی مطلوب را ندهد [۸]. به‌همین خاطر برای رسیدن به جوابی پایدار و مطلوب، میل به ترکیب روش‌های هوشمند رو به رشد است. تاکنون روش‌های بسیاری در غالب تحقیقات متعدد که هر کدام در بهترین حالت از طریق شبکه‌های عصبی^۱ RBF با دقت ۷۳/۳۲٪ [۹]، شبکه‌های عصبی MLP^۲ با دقت ۷۶/۸۹٪ [۱۰]، K^۳ امین نزدیکی همسایگی با دقت ۷۵/۵۵٪ [۱۱]، الگوریتم‌های تکاملی^۴ با دقت ۷۶/۱۷٪ [۳] و الگوریتم SVM با دقت ۸۲/۵٪ [۳]، براساس ویژگی‌های ۸ گانه‌ی سازمان جهانی بهداشت در زمینه‌ی تشخیص بیماری معرفی و بررسی شده‌اند. به‌منظور تشخیص بیماری دیابت علاوه بر تحقیقات متعدد، توسط همین مؤلفین نیز تحقیقاتی [۱۲، ۱۳] براساس ویژگی‌های ۸ گانه سازمان جهانی بهداشت یعنی الف: سن، ب: سابقه‌ی بیماری دیابت، ج: شاخص جرم بدن (kg/m²)، د: انسولین سرم دو ساعته (mu U/ml)، ه: ضخامت پوست ماهیچه سه سر بازویی (mm)، و: فشار خون (mm gh)، ز: غلظت گلوکز پلاسما‌ی خون در دو ساعت، ک:

دفعات بارداری صورت گرفته، که با توجه به نتایج؛ استفاده از تکنیک‌های داده‌کاوی و هوش مصنوعی توصیه شده است. مشکلی که در رابطه با این بیماری و مخصوصاً در نوع یک آن وجود دارد، افزایش و یا کاهش بیش از حد سطح قند خون است که می‌تواند به بیهوشی، اغماء و حتی مرگ بیمار منجر شود [۱۴]. بنابراین استفاده از روشی مناسب به‌منظور پیش‌بینی و تعیین میزان صحیح دُز انسولین و در نهایت، پیش‌گیری از این عوارض گام مهمی در جهت کنترل بهتر این بیماری محسوب می‌شود. در همین راستا تحقیقاتی انجام گرفته است؛ از آن جمله می‌توان به مدل‌های ریاضی و مینیمال [۱۷-۱۵] براساس ویژگی‌های حساسیت بدن بیمار در پاسخ به گلوکز تزریقی، دُز انسولین مصرفی و سطح فعلی قند خون اشاره کرد که برای تعیین حساسیت بدن بیمار در پاسخ به گلوکز تزریقی و نیز پیش‌بینی کوتاه مدت سطح قند و انسولین است. در سال ۲۰۱۰ با استفاده از شبکه‌های عصبی، پیش‌بینی سطح گلوکز خون برای بیماران براساس ویژگی‌های اطلاعات CGMS^۵ در هر بیمار شامل؛ وزن، زمان اولیه‌ی تزریق، زمان انتهایی تزریق و کربوهیدرات مصرفی، انجام شده است [۱۸]. با استفاده از ویژگی‌های ۱۳ گانه‌ی بانک اطلاعاتی NASA^۶ [۱۹] توسط روش برنامه‌ریزی LSSVM^۷ پیش‌بینی دُز انسولین صورت گرفته است [۲۰]. یک سیستم حلقه بسته‌ی عملی آزمایشگاهی به‌صورت آنالین برای کنترل گلوکز خون براساس جنسیت، سن، مدت مصرف، هموگلوبین HbA1c، دُز انسولین و کربوهیدرات مصرفی و ویژگی‌های آماری شامل میانگین و انحراف از معیار در ویژگی‌های بیان شده، تهیه شده است [۲۱]. Scott و همکاران [۱۵] در سال ۲۰۰۸ تنها برای ۱۸۰-۵۰ دقیقه توانستند براساس ویژگی‌های نحوه‌ی زندگی، وضعیت روحی، دُز انسولین، وعده‌ی غذایی و کربوهیدرات مصرفی، پیش‌بینی سطح انسولین برای بیماران را انجام دهند. Stavroula و همکاران [۱۶] با استفاده از ۴ پارامتر شامل، عملکرد کوتاه، متوسط، سریع، بلندمدت و کربوهیدرات مصرفی و غلظت قند خون که تابعی از زمان بعد از تزریق انسولین هستند، توسط شبکه‌های عصبی میزان دُز انسولین را برای بیماران مشخص کردند. Wahab و همکاران [۱۷] با استفاده از کنترل کننده‌ی تطبیقی براساس ۲ معادله‌ی دیفرانسل مرتبه‌ی

⁵ Continues Glucose Monitoring System for diabetic

⁶ None Adaptive And Aplication for patient

⁷ Least Square Support Vector Machine

¹ Radial Basis Function

² Multi Layer Preceptron

³ K-Near Neighbor

⁴ Evolutionary Algorithm

الگوریتم ماشین بردار پشتیبان (SVM^۲) برای طبقه‌بندی نمونه‌ها و همین‌طور از شبکه‌های فازی - عصبی برای طراحی طبقه‌بندی کننده و آزمایش نهایی استفاده می‌گردد. الگوریتم‌های انتخاب ویژگی بسته به روند ارزیابی آن‌ها به دو دسته‌ی عمده تقسیم می‌شوند. اگر انتخاب ویژگی مستقل از هرگونه الگوریتم یادگیری انجام شود (یعنی به‌صورت یک پیش‌پردازنده‌ی کاملاً مجزا)، آن را روش فیلتر یا حلقه‌ی باز گویند. در این مورد ویژگی‌های نامطلوب الگوریتم طبقه‌بندی در ارتباط باشد، روش انتخاب ویژگی را Wrapper یا حلقه‌ی بسته می‌نامند. این روش، جستجو در فضای زیر مجموعه‌ها را براساس تخمین دقت ناشی از انتخاب یک زیر مجموعه‌ی خاص تحت شرایط الگوریتم طبقه‌بندی مورد استفاده (به‌عنوان معیاری از بهینگی آن زیرمجموعه) انجام می‌دهد [۲۶]. الگوریتم‌های دسته‌ی دوم معمولاً نتایج بهتری به‌دست می‌دهند. مهم‌ترین بخش هر روش انتخاب ویژگی حلقه‌ی بسته، الگوریتم جستجویی است که در آن به‌کار رفته است. طی دهه‌ی گذشته محققان روی الگوریتم‌های جستجوی تکاملی مثل الگوریتم ژنتیک [۲۷]، ACO^۳ [۲۶] و ترکیب ACO/PSO [۲۸] تمرکز کرده‌اند.

روش‌ها

به‌منظور تعیین و پیش‌بینی دُز انسولین برای بیماران، از مرکز تحقیقات دیابت شهرستان سبزواز یاری گرفته شد. به این ترتیب که با هماهنگی و همکاری تیم پزشکی و متخصص غدد، پرسش‌نامه‌هایی به‌منظور مشخص کردن ویژگی‌های مؤثر توسط کارشناسان مرکز و تیم اجرایی طراحی شد. ۳۱۲ بیمار که به‌صورت کاملاً تصادفی (سیستماتیک) انتخاب شده‌اند، این پرسش‌نامه‌ها تکمیل شد، براساس نظر پزشکان سعی شد تا ویژگی‌هایی تعریف شوند که در روند درمان، مدل سازی رفتار بیمار و تصمیمات پزشکان بیشترین تأثیر را داشته باشند. بعد از مشخص شدن ویژگی‌ها، کُدگذاری این ویژگی‌ها مهم می‌باشد، که در پاره‌ای از موارد مقدار دقیق عددی آزمایش را قرار می‌دهیم و در قسمت‌هایی دیگر لازم است تا مقدار کُد شده قرار داده شود. به‌منظور کُدگذاری از نظر تیم پزشکی بر اساس راهنمایی‌های بیماران، استفاده شد. در جدول ۱ نمونه‌ای از ویژگی‌ها به‌همراه نحوه‌ی کُدگذاری پیشنهادی مشخص شده است.

اول برای انسولین و تنظیم گلوکز در ۲ زیر سیستم، تعیین دُز انسولین را انجام داده‌اند. در ابتدای روش پیشنهادی، ابتدا داده‌های بیماران با همکاری مرکز تحقیقات بیماری‌های دیابت دانشگاه علوم پزشکی سبزوار جمع‌آوری شد، سپس به‌علت ازدیاد ویژگی‌ها یا بُعد داده‌های استخراج شده، لازم است تا ویژگی‌های مطلوب و اصطلاحاً بهینه از سایر ویژگی‌های جدا شوند، در قسمت پیش‌بینی صحیح دُز انسولین به‌همین خاطر روش ترکیبی الگوریتم پرندگان - ماشین بردار پشتیبان کاهش بعد استفاده شد، بنابراین توسط ویژگی‌های انتخاب شده به‌منظور تعیین درصد (بیان نسبی) سلامتی و دیابتی افراد و همین‌طور ویژگی‌های انتخاب شده توسط الگوریتم پیشنهادی، میزان صحیح دُز انسولین برای بیماران پیش‌بینی می‌شود. ما سعی کردیم ابتدا مسأله و وخامت موضوع در مورد بیماری دیابت بیان، سپس روش کلی شرح داده شود. بعد از آن، با توجه به مطالب فلوجارت ۱ به بحث و بررسی رئوس مطالب می‌پردازیم. امید است تا این تحقیق گامی هرچند اندک ولی مؤثر در جهت بهبود روش‌ها، معالجه و تشخیص این بیماری داشته باشد.

مسأله انتخاب ویژگی و بازشناسی الگو

ابتدا باید به مسأله انتخاب ویژگی‌های مناسب در فضای تبدیل پرداخت. انتخاب ویژگی در واقع برگزیدن ویژگی‌هایی است که حداکثر توان را در پیشگویی خروجی دارا باشند [۲۲]. تعریف این که زیر مجموعه‌ی بهینه چه می‌تواند باشد، به مسائلی که قصد داریم حل کنیم، وابسته است [۲۳]. در روش‌هایی که بحث شده‌اند، مقادیر ویژه بردارها اهمیت آن‌ها را نشان می‌دهند. اهمیت دادن به تمام ویژگی‌ها بر این اساس همیشه مناسب نمی‌باشد. علاوه بر این بردارهای ویژگی نیز همیشه شامل اطلاعات مفید نمی‌باشند. همان‌گونه که به اثبات رسیده است [۲۴]، با اینکه انتظار می‌رود با داشتن اطلاعات بیشتر نتایج بهتر شود، اما در عمل مشاهده می‌شود که این مسأله باعث افت دقت می‌گردد [۲۵]. بنابراین برای رسیدن به یک بازنمایی مناسب، انتخاب بهترین بردار ویژگی یا ساده‌سازی بردار ویژگی امری ضروری است. یک سیستم شناسایی الگوی متداول شامل ۴ بخش، استخراج ویژگی، انتخاب ویژگی، طراحی و آموزش طبقه بندی کننده و سرانجام آزمایش است. در قسمت پیش‌بینی دُز انسولین از الگوریتم پرندگان دودویی (BPSO^۱) برای کاوش و انتخاب ویژگی‌های مؤثر،

^۲ Support Vector Machine

^۳ Ant Colony Optimization

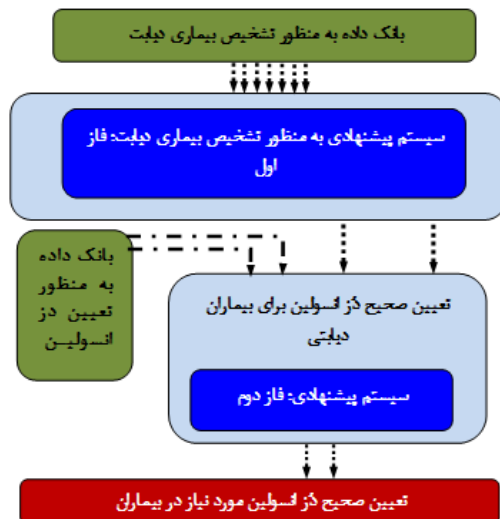
^۱ Binary Particle Swarm Optimization

جدول ۱- ویژگی‌های مورد نیاز و ثبت شده

ویژگی‌های مورد نیاز در تشخیص بیماری دیابت						
شماره	ویژگی	نوع کُدگذاری	پیشینه-کمینه	شماره	ویژگی	نوع کُدگذاری
۱	بارداری	-----	[۰-۱]	۵	گلوکز خون دو ساعته	-----
۲	میزان گلوکز خون	-----	[۵۰-۳۰۰]	۶	ارثی بودن	۰ یا ۱ یا ۲ یا ... ۵
۳	فشار خون	-----	[۴۰-۱۳۰]	۷	شاخص جرم توده بدن	-----
۴	قطر ماهیچه سه سر بازویی	-----	[۰-۱۰۰]	۸	سن	-----
ویژگی‌های منتج از مرحله تشخیص در تعیین صحیح دُز انسولین در بیماران						
شماره	ویژگی	نوع کُدگذاری	پیشینه-کمینه	شماره	ویژگی	نوع کُدگذاری
۱	نسبت دیابتی	-----	[۰-۱]	۲	نسبت سلامت	-----
ویژگی‌های مورد نیاز در تعیین صحیح دُز انسولین در بیماران						
شماره	ویژگی	نوع کُدگذاری	پیشینه-کمینه	شماره	ویژگی	نوع کُدگذاری
۱	سطح غلظت قند خون در ابتدای بازه‌ی زمانی (mg/d lit)	مقدار اندزه‌گیری	[۰-۱۰۰۰]	۷	غلظت انسولین ۶ ساعته تزریقی به بدن (unit)	مقدار اندزه‌گیری
۲	دز انسولین ۱۲ ساعته تزریقی به بدن (در مقیاس واحد)	مقدار اندزه‌گیری	[۰-۱۰۰۰]	۸	سطح استرس (unit)	۰ یا ۱ یا ... ۵
۳	سطح فعالیت بدنی (در مقیاس واحد)	۵، ۰۰۰، ۲، ۱، ۰	[۰-۵]	۹	Total Cholesterol (mg/dl)	مقدار اندزه‌گیری
۴	چگالی بالای لیپوپروتین (HDL)	مقدار اندزه‌گیری	[۱۰-۴۰۰]	۱۰	Fast Blood Sugar (mg/dl)	مقدار اندزه‌گیری
۵	سابقه‌ی بیماری	۱۱ و ۱ یا ۱۱	[۰-۱۱]	۱۱	Blood Sugar (mg/dl)	مقدار اندزه‌گیری
۶	سابقه‌ی دارو	۰ یا ۱ یا ... ۵	[۰-۵]	۱۲	چگالی پایین لیپوپروتین (LDL)	مقدار اندزه‌گیری

الگوریتم پیشنهادی

ابتدا بانک اطلاعاتی جامعی که در قسمت قبل توضیح داده شد، تهیه می‌شود، سپس با استفاده از سیستم پیشنهادی، بیماران تشخیص داده می‌شوند (فاز اول)، در مرحله بعد با استفاده از ۱۲ ویژگی ثبت شده به همراه ۲ ویژگی (بیان فازی (پیشرفت) بیماری یا سالم بودن) مجموعاً ۱۴ ویژگی، پیش بینی صحیح دُز انسولین برای بیماران انجام می‌شود. تمام مراحل پیشنهادی در فلوچارت ۱ بیان شده است.



فلوچارت ۱- روند الگوریتم پیشنهادی

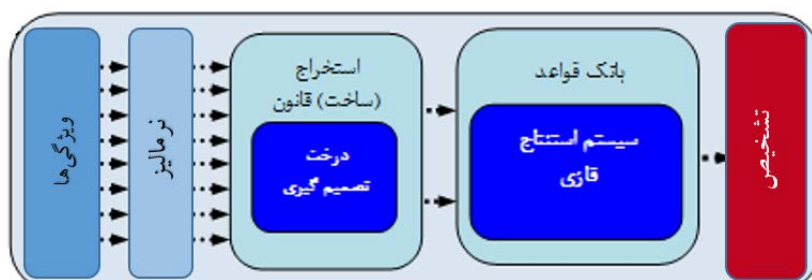
درخت فراهم شود. با توجه به مسیرهای ترسیمی درخت، یک سری شروط منطقی برای رسیدن از مبدا درخت به انتهای درخت (درصد بیمار یا سالم بودن فرد) به دست می‌آید که از این شروط می‌توان به‌عنوان منشاء بانک قواعد سیستم فازی (Fuzzy) استفاده کرد. سپس سیستم فازی با توجه به بانک قواعد مبتنی بر درخت تصمیم‌گیری و ورودی‌ها (ویژگی‌های هر فرد)، میزان سالم یا ناسالم بودن برای هر فرد را مشخص کرده تا تعیین صحیح دُز انسولین برای بیماری از طریق ورود این اطلاعات و همین‌طور ویژگی‌های دیگر بانک دوم اطلاعاتی، مشخص شود. فلوجارت ۲ بیانگر تمام مراحل کار در این فاز به‌منظور تشخیص است.

مراحل اجرایی در تشخیص بیماری دیابت

در ابتدا سعی شد تا داده‌ها مطابق مرجع (Pima Indian Diabet) (data base) از پایگاه اطلاعاتی تهیه شوند [۲۹]. با این تفاوت که ویژگی شماره‌ی ۴ در مرجع PID، ضخامت پوست ماهیچه دو سر بازویی می‌باشد که در اینجا سعی کردیم تا ویژگی را حذف نماییم، چرا که مشخص کردن این ویژگی ممکن نبود، سپس ویژگی‌ها را توسط رابطه‌ی ۱ در بازه‌ی [0 1] نرمالیزه می‌کنیم.

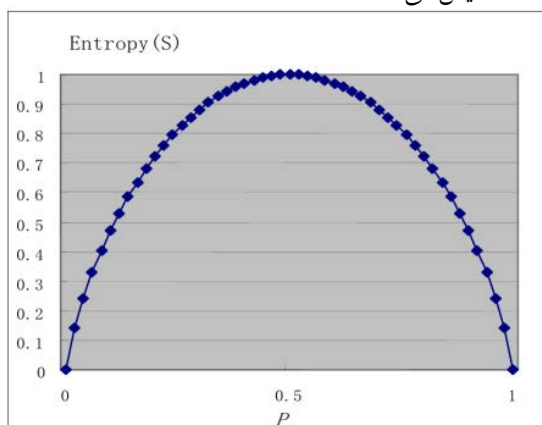
$$\text{normalize}(x_i) = \frac{x_i}{\max(x_i)} \quad \text{رابطه‌ی ۱}$$

بعد از اینکه ویژگی‌ها مشخص شدند، توسط فرآیند درخت تصمیم‌گیری (Decision-Tree) مورد کاوش قرار می‌گیرند تا فضایی برای تحقیق و مسیر رسیدن از فضای ورودی (ویژگی‌های بیماران) به خروجی (میزان بیمار بودن (۱) یا سالم بودن افراد (۰)) توسط ترسیم نمادین یک



فلوجارت ۲- روند اجرایی فاز اول الگوریتم پیشنهادی در جهت تشخیص بیماری

براساس احتمال داده‌های یک کلاس برای یک مسئله دو کلاسه نمایش می‌دهد.



شکل ۱- تغییرات آنتروپی براساس احتمال رخداد [۳۱]

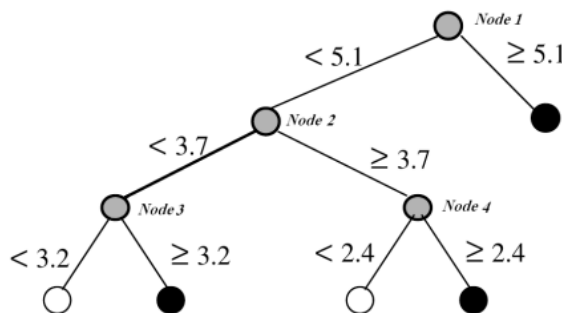
درخت تصمیم‌گیری

درخت‌های تصمیم‌گیری اولین بار در سال ۱۹۸۶ توسط کوینلان با معرفی الگوریتم ID3^۱ ایجاد گردیدند که از معیار آنتروپی برای ارزیابی قدرت تفکیک هر کدام از ویژگی‌های داده‌ها استفاده می‌کنند [۳۰]. مهم‌ترین خصوصیت درخت‌های تصمیم‌گیری قدرت بالای آن‌ها در خرد کردن یک مسئله پیچیده به مسایل کوچک‌تر و در نتیجه ارائه یک راه حل قابل درک می‌باشد [۳۰، ۳۱]. هرچه تنوع داده‌ها بیشتر و اعضای دو کلاس برابر شوند، معیار آنتروپی به سمت یک و هر چه تعداد اعضای یک کلاس نسبت به دیگری کمتر شوند به سمت صفر میل می‌کند [۳۱]. شکل ۱ تغییرات آنتروپی را

^۱ Iterative Dichotomies 3

داده‌های ورودی گسسته نمایش داده شده با L مشخص شود. آنگاه 2L-1 حالت برای انتخاب دسته‌ها وجود خواهد داشت. در صورتی که داده‌ها به صورت پیوسته باشند و در مجموع K دسته داده‌ی متفاوت وجود داشته باشد آنگاه K-1 امکان برای تقسیم داده‌های ورودی وجود خواهد داشت. محاسبه خلوص داده‌ها از معیارهای توینگ و توینگ مرتب شده انجام می‌شود. شکل ۲ یک نمونه درخت با ۴ گره می‌باشد [۲۹].

درخت (Classification And Regression Tree) CART (درخت استفاده شده در این کار) در سال ۱۹۸۴ توسط بریمن ارائه گردید. ایده‌ی اصلی این روش در هر مرحله، از میان ترکیب تمام حالت‌های ممکن برای تقسیم داده‌ها به دو قسمت بوده که آن تقسیمی انتخاب می‌شود که منجر به تولید اطلاعاتی خالص‌تر گردد [۲۹]. مناسب بودن تقسیم تنها به یک مرحله قبل وابسته بوده و تقسیم بهینه براساس مشاهدات قبلی بررسی نمی‌شود. اگر

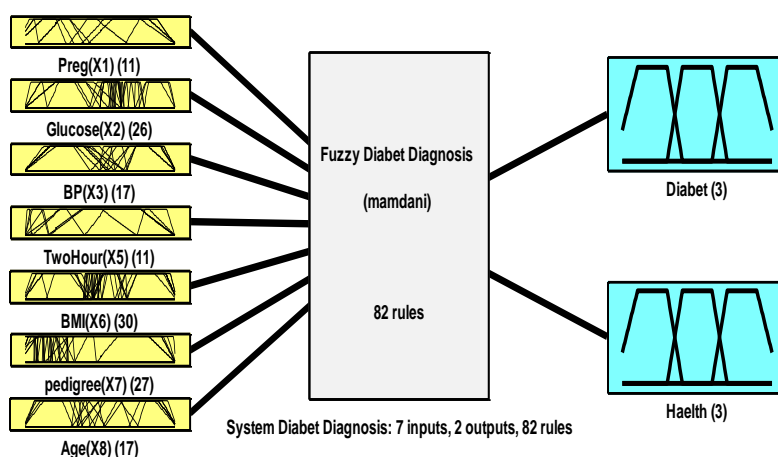


شکل ۲- شمای یک نمونه درخت CART [۲۹]

درجه‌ی عضویت (عددی بین ۰ یا ۱) می‌باشد [۳۲، ۳۳]. هر سیستم فازی قوانینی دارد به صورت عبارات "اگر - آنگاه" در بانک قواعد. در این تحقیق برای به دست آوردن قوانین منطقی و صحیح از تصمیمات درخت استفاده کرده‌ایم. مرحله پردازش که به آن موتور استخراج تصمیم‌گیری گفته می‌شود براساس بانک قواعد عمل می‌کند. دو نوع سیستم استخراج فازی ممدانی و سوگونو وجود دارند که در روش تعیین خروجی با هم تفاوت دارند [۳۴].

سیستم‌های فازی

منطق فازی تکنیکی نسبتاً جدیدی است که در مقابل روش‌های مرسوم برای طراحی و مدل‌سازی سیستم‌هایی که نیازمند ریاضیات و احتمالات نسبتاً پیشرفته هستند به کار می‌رود. منطق فازی گزاره‌های صحیح بولی را با درجه‌ی درستی بیان می‌کند. تابع عضویت یک مجموعه فازی، در حالت خاص یک منحنی است که نشان دهنده‌ی چگونگی نگاشت هر یک از نقاط فضا به



System Diabet Diagnosis: 7 inputs, 2 outputs, 82 rules

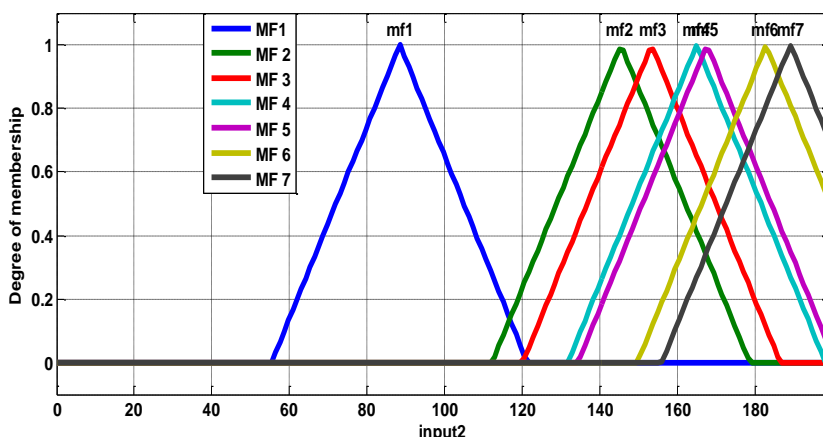
شکل ۱- سیستم فازی طراحی شده

در سیستم مورد نظر، از استنتاج ممدانی استفاده کرده‌ایم، شکل ۳ سیستم فازی طراحی شده با ۷ ورودی و ۲ خروجی است. متغیرهای ورودی (برای هر فرد) در این سیستم همان ویژگی‌های ۷ گانه‌ی بانک اطلاعاتی می‌باشند. برای هر ورودی باید توابع عضویت و بازه‌ی آن‌ها مشخص شود، تعداد توابع عضویت از روی محدوده‌ی آن‌ها مشخص می‌شود، و محدوده‌ی توابع عضویت توسط تصمیمات درخت مشخص می‌شود. نوع توابع عضویت با توجه به دامنه و محدوده توابع عضویت مشخص می‌شود، مثلاً توابع عضویتی (در متغیرهای ورودی) که دارای دامنه‌ی مناسب هستند از نوع "مثلثی" (مانند شکل ۴ برای ویژگی

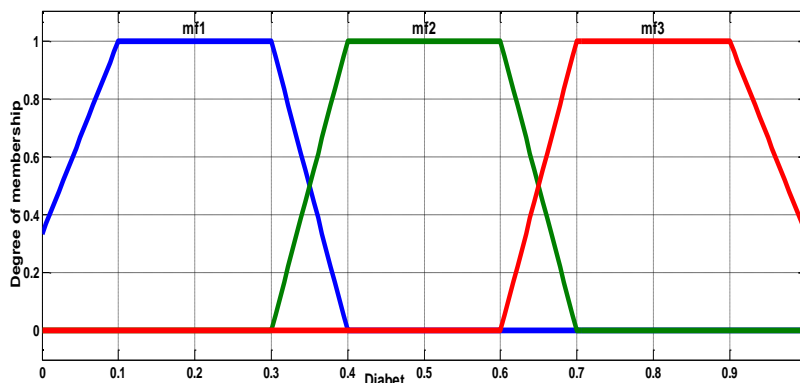
ورودی) دوم و معادله ۲) و توابعی که دارای دامنه‌ی خیلی کم می‌باشند از نوع "دوزنقه‌ای" (مانند شکل ۵ برای خروجی اول و رابطه ۳) در نظر گرفته شده است. رابطه‌ی ۲ و ۳:

$$f(x, a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right);$$

$$f(x, a, b, c, d) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right);$$



شکل ۴- توابع عضویت ورودی (ورودی دوم)



شکل ۵- توابع عضویت خروجی (خروجی اول)

متغیرهای خروجی بیانگر میزان عضویت هر فرد (به بیان فازی) به ۲ متغیر خروجی سالم و بیمار که توسط ۳ تابع عضویت دوزنقه‌ای بیان شدند، می‌باشند. مهم‌ترین قسمت از سیستم فازی بانک قواعد می‌باشد. در بانک قواعد با توجه به ورودی، خروجی، توابع عضویت و سیستم استنتاج، تصمیمات صحیح برای

مشخص شدن مقدار خروجی Crisp گرفته می‌شود. در جدول ۲ تمام پارامترهای سیستم فازی پیشنهادی بیان شده است.

جدول ۲- مشخصات سیستم فازی

7 in / 2out	تعداد متغیرها	ممدانی	استلزام
Variable	تعداد توابع عضویت	min	اشتراک
Trap&Gbell	نوع توابع عضویت	max	اجتماع
Centroid	دی فازی‌ساز	min	استنتاج

(توابع عضویت مجموعه‌ها) V قانون ایجاد می‌شود، که برقراری ارتباط بین این ورودی‌ها و خروجی‌ها مشکل است، لذا با توجه به تعداد زیاد ویژگی‌ها از یک طرف و خروجی‌های تعریف شده، از درخت تصمیم‌گیری به منظور اجتناب از این مشکل استفاده شده است. جدول ۳ شمایی از قوانین یک شاخه از درخت تصمیم‌گیری و خروجی افراد سالم را نمایش می‌دهد. قسمتی از قوانین مستخرج از جدول ۳ در جدول ۴ بیان شده است.

شاخه‌ی دوم در جدول ۵ درخت، رابطه‌ی بین ویژگی‌ها و بیمار بودن هر فرد مشخص شده است.

درختان تصمیم‌گیری قادر به تولید توصیفات قابل درک، از روابط موجود در یک مجموعه داده‌ای هستند و می‌توانند برای وظایف دسته‌بندی و پیش‌بینی به کار روند [۳۵]. ساختار تصمیم‌گیری می‌تواند به شکل تکنیک‌های محاسبات که به توصیف، دسته‌بندی و عام‌سازی یک مجموعه از داده‌ها کمک می‌کنند نیز معرفی شوند [۳۶]. در فلوچارت ۲ ارتباط بین درخت تصمیم‌گیری و سیستم فازی بیان شد. منشاء اصلی نگارش قوانین فازی، با تصمیمات درخت می‌باشد. یکی از روش‌ها برای به دست آوردن تعداد صحیح قوانین قاعده‌ی Decouple است [۳۷]. حداقل ۷ مجموعه تحت عنوان بانک قواعد، هر کدام در هفت به توان توابع عضویت مجموعه‌ها

جدول ۳- قسمتی از شاخه درخت سلامتی

قسمتی از شاخه‌ی درخت به منظور بیان سلامتی افراد
$X_2 \geq 143.5$ $ X_2 \geq 143.5 \& X_2 \leq 154.5: \text{Health} = 0.5$ $ X_2 \geq 143.5 \& X_2 > 154.5 \& X_6 < 29.85: \text{Health} = 0.47619$ $ X_2 \geq 143.5 \& X_2 > 154.5 \& X_6 \geq 29.85 \text{ and } X_5 \geq 544: \text{Health} = 0.6$ $ X_2 \geq 143.5 \& X_2 > 154.5 \& X_6 \geq 29.85 \text{ and } X_5 < 544 \& X_7 < 0.3005: \text{Health} = 0.25$ $ X_2 \geq 143.5 \& X_2 > 154.5 \& X_6 \geq 29.85 \& X_5 < 544 \& X_7 \geq 0.3005 \& X_4 \geq 44 \text{ Health} = 0.21063$ $ X_2 \geq 143.5 \& X_2 > 154.5 \& X_6 \geq 29.85 \& X_5 < 544 \& X_7 \geq 0.3005 \& X_4 < 44 \text{ Health} = 0$

جدول ۴- قوانین مستخرج از جدول ۳

قسمتی از قوانین شاخه درخت به منظور بیان سلامتی افراد
$\text{if } x_2 \geq 143.5 \text{ and } x_2 \leq 154.5 \text{ then Health} = 0.5$ $\text{if } x_2 \geq 143.5 \text{ and } x_2 > 154.5 \text{ and } x_6 < 29.85 \text{ then Health} = 0.47619$ $\text{if } x_2 \geq 143.5 \text{ and } x_2 > 154.5 \text{ and } x_6 \geq 29.85 \text{ and } x_5 \geq 544 \text{ then Health} = 0.6$ $\text{if } x_2 \geq 143.5 \text{ and } x_2 > 154.5 \text{ and } x_6 \geq 29.85 \text{ and } x_5 < 544 \text{ and } x_7 < 0.3005 \text{ then Health} = 0.25$ $\text{if } x_2 \geq 143.5 \text{ and } x_2 > 154.5 \text{ and } x_6 \geq 29.85 \text{ and } x_5 < 544 \text{ and } x_7 \geq 0.3005 \text{ and } x_4 \geq 44 \text{ then Health} = 0.21063$ $\text{if } x_2 \geq 143.5 \text{ and } x_2 > 154.5 \text{ and } x_6 \geq 29.85 \text{ and } x_5 < 544 \text{ and } x_7 \geq 0.3005 \text{ and } x_4 < 44 \text{ then Health} = 0.0$

جدول ۵- قسمتی از شاخه درخت دیابتی

قسمتی از شاخه درخت به منظور بیان دیابتی بودن افراد
$X_2 \geq 143.5$
$ X_2 \geq 143.5 \ \& \ X_2 \leq 154.5: \text{Diabetic} = 0.5$
$ \ \ X_2 \geq 143.5 \ \& \ X_2 > 154.5 \ \& \ X_6 < 29.85: \text{Diabetic} = 0.52381$
$ \ \ \ X_2 \geq 143.5 \ \& \ X_2 > 154.5 \ \& \ X_6 \geq 29.85 \ \text{and} \ X_5 \geq 544: \text{Diabetic} = 0.4$
$ \ \ \ \ X_2 \geq 143.5 \ \& \ X_2 > 154.5 \ \& \ X_6 \geq 29.85 \ \text{and} \ X_5 < 544 \ \& \ X_7 < 0.3005: \text{Diabetic} = 0.75$
$ \ \ \ \ \ X_2 \geq 143.5 \ \& \ X_2 > 154.5 \ \& \ X_6 \geq 29.85 \ \& \ X_5 < 544 \ \& \ X_7 \geq 0.3005 \ \& \ X_4 \geq 44: \text{Diabetic} = 0.78937$
$ \ \ \ \ \ \ X_2 \geq 143.5 \ \& \ X_2 > 154.5 \ \& \ X_6 \geq 29.85 \ \& \ X_5 < 544 \ \& \ X_7 \geq 0.3005 \ \& \ X_4 < 44: \text{Diabetic} = 1$

براساس جدول (۵) قوانینی بدست می آید که در جدول (۶) بخشی از این قوانین نیز مشخص شده اند.

جدول ۶- قوانین مستخرج از شاخه دیابتی یا جدول (۵)

قسمتی از قوانین شاخه درخت به منظور بیان دیابتی بودن افراد
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 \leq 154.5 \ \text{then} \ \text{Patient} = 0.5$
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 > 154.5 \ \text{and} \ x_6 < 29.85 \ \text{then} \ \text{Patient} = 0.52381$
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 > 154.5 \ \text{and} \ x_6 \geq 29.85 \ \text{and} \ x_5 \geq 544 \ \text{then} \ \text{Patient} = 0.4$
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 > 154.5 \ \text{and} \ x_6 \geq 29.85 \ \text{and} \ x_5 < 544 \ \text{and} \ x_7 < 0.3005$
$\text{then} \ \text{Patient} = 0.75$
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 > 154.5 \ \text{and} \ x_6 \geq 29.85 \ \text{and} \ x_5 < 544 \ \text{and} \ x_7 \geq 0.3005$
$\text{and} \ x_4 \geq 44 \ \text{then} \ \text{Patient} = 0.78937$
$\text{if } x_2 \geq 143.5 \ \text{and} \ x_2 > 154.5 \ \text{and} \ x_6 \geq 29.85 \ \text{and} \ x_5 < 544 \ \text{and} \ x_7 \geq 0.3005$
$\text{and} \ x_4 < 44 \ \text{then} \ \text{Patient} = 1$

بیماران لازم است تا، ارزیابی از سیستم پیشنهادی صورت گیرد. در نهایت ۸۲ قانون براساس رفتار درخت تصمیم گیری به دست می آید که البته ۲۰ قانون ابتدایی آن در جدول ۷ براساس ساختار فهرستی نمایه ای مشخص شده است.

با ورود متغیرهای ورودی (ویژگی های هر بیمار) و در نظر داشتن توابع عضویت و اصطلاحاً آتش شدن (فعال شدن قوانین)، اعمال استنتاج و در نهایت استفاده از دفازی ساز (مرکز ثقل) برای ساخت مقادیر فازی زدا، خروجی به دست می آید. قبل از استفاده از خروجی این مرحله به منظور تعیین صحیح دُز انسولین در

جدول ۷-۲۰ قانون ابتدایی در سیستم فازی

قسمتی از قوانین برگرفته از درخت در سیستم فازی
بر اساس ساختار فهرستی (Index)، ۲۰ قانون اول

1 1 0 0 1 1 1, 1 0	0 5 3 0 3 5 1, 1 0
(1): 1	(1): 1
1 1 0 0 1 2 1, 3 0	0 5 3 0 4 5 1, 3 0
(1): 1	(1): 1
1 1 0 0 1 3 1, 1 0	0 6 3 0 2 5 1, 1 0
(1): 1	(1): 1
2 1 0 0 1 0 1, 3 0	0 7 4 0 2 5 1, 3 0
(1): 1	(1): 1
0 1 1 0 2 4 1, 2 0	0 8 3 0 2 5 1, 1 0
(1): 1	(1): 1
0 2 2 1 2 4 1, 1 0	0 8 3 0 2 5 1, 1 0
(1): 1	(1): 1
0 3 2 1 2 4 1, 1 0	0 1 0 0 5 0 1, 3 0
(1): 1	(1): 1
0 3 2 2 2 4 1, 3 0	0 1 0 0 6 0 1, 3 0
(1): 1	(1): 1
0 3 2 3 2 4 1, 1 0	0 1 0 0 7 0 2, 1 0
(1): 1	(1): 1
0 4 3 0 2 5 1, 1 0	0 25 0 0 8 0 2, 3 0
(1): 1	(1): 1

تعداد اشیاء یا داده است)، مشخص می‌شود. در این روش دو محدودیت اصلی وضع شده است؛ محدودیت اول این است که هیچ خوشه‌ای نباید تهی باشد، $(\sum_{j=1}^n u_{ij} > 0 \quad \forall i \in \{1, \dots, c\})$ و محدودیت دوم که محدودیت نرمال سازی نامیده می‌شود، بیان می‌کند که مجموع درجه‌ی عضویت هر داده به همه‌ی خوشه‌ها باید برابر یک باشد $(\sum_{j=1}^n u_{ij} = 1 \quad \forall j \in \{1, \dots, n\})$.

روش دسته‌بندی میانگین فازی سعی می‌کند تا برای یک مجموعه نقاط داده، افزازهایی (c عدد خوشه به صورت فازی) بیابد که تابع هدف یا هزینه زیر را کمینه نماید (رابطه‌ی ۴).

رابطه‌ی ۴:

$$J_f(X, U_f, C) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2$$

d_{ij} فاصله‌ی بین داده‌ی X_j و مرکز خوشه i است و $m \in [1, \infty)$ میزان فازی بودن است (معمولاً $m = 2$). پس اگر m به سمت یک میل کند، $(m \rightarrow 1)$ خوشه‌بندی کریسپ یا سخت‌تر می‌شود و برعکس اگر m به سمت بی‌نهایت میل کند $(m \in \infty)$ خوشه‌بندی فازی‌تر خواهد شد. البته تابع J_f را نمی‌توان مستقیماً کمینه نمود، بنابراین باید از

ارزیابی فاز اول سیستم پیشنهادی در بیان پیشرفت و سلامت افراد

در ارزیابی سیستم پیشنهادی که به‌منظور بیان پیشرفت بیماری و سلامت افراد طراحی شده است، از الگوریتم خوشه‌بندی میانگین فازی استفاده می‌نماییم. این الگوریتم سعی در خوشه‌بندی یا کلاس‌بندی داده دارد، علت استفاده از این الگوریتم به این خاطر است که ارزیابی این الگوریتم بدون ناظر و بردار هدف است، یعنی فقط انتظار ما از این الگوریتم خوشه‌بندی و یا کلاس‌بندی بیماران به ۲ دسته است، دسته بیماران و افراد سالم، که البته اگر سیستم پیشنهادی عملکرد صحیحی داشته باشد، داده‌ها به‌طور ۱۰۰ درصد به ۲ دسته تقسیم می‌شوند.

الگوریتم میانگین فازی FCM^۱

Baraldi در سال ۱۹۶۹ اولین مدل خوشه‌بندی با ایده فازی را مطرح کرد [۳۸]. در این روش میزان عضویت یا تعلق هر شی یا داده به هر خوشه، در ماتریس عضویت $(U = [u_{i,j}]_{c \times n} = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n))$ که تعداد خوشه‌ها و n

^۱ Fuzzy C-Means

گام آخر، ماتریس مرکز خوشه‌ها را با درجه‌ی عضویت‌های بهینه شده، به‌روزرسانی می‌نماییم.

$$C_i^{(t+1)} = \frac{\sum_{j=1}^n (u_{ij}^{(t+1)})^m \bar{x}_j}{\sum_{j=1}^n (u_{ij}^{(t+1)})} \quad \text{رابطه‌ی ۲:}$$

تکرار گام ۲ و ۳ تا زمانی که یا $\|C^{(t+1)} - C^{(t)}\| < \epsilon$ یا $\|U^{(t+1)} - U^{(t)}\| < \epsilon$ برقرار باشند.

با آزمایش تمام مراحل فوق بر روی داده‌های حاصل از سیستم فازی (مرحله‌ی تشخیص) شکل ۶ حاصل می‌شود. با کمی دقت مشخص می‌شود که داده‌ها به ۲ دسته تقسیم شدند (تعدادی از داده‌ها در هیچ دسته‌ای نرفتند که مسلماً ضعف سیستم کلاس‌بندی و عملکرد سیستم پیشنهادی می‌باشد).

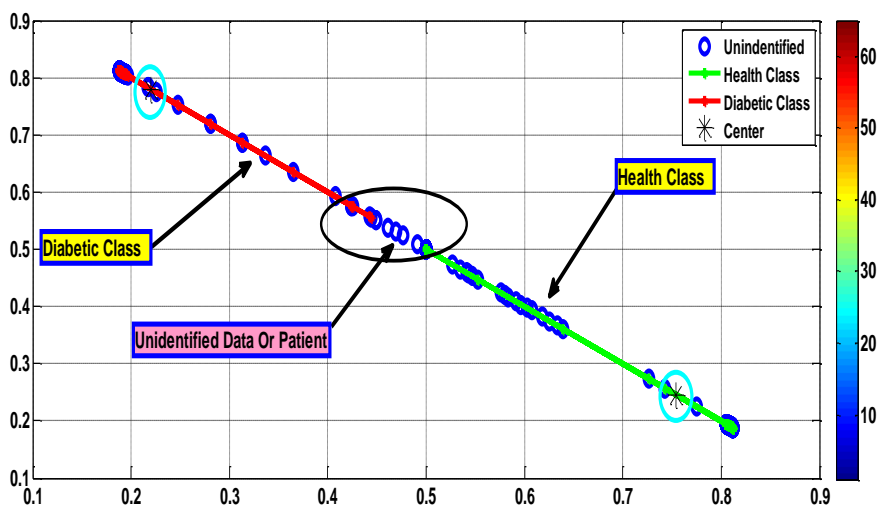
الگوریتم‌های تکراری استفاده کرد که برای حل این مشکل از شمای جایگزینی بهینه به‌صورت زیر استفاده شد.

(۱) انتخاب مقادیر مناسبی برای c, m و یک عدد مثبت کوچک برای ϵ ؛ ماتریس C (مرکز یا میانه‌ی خوشه‌ها) را نیز به‌صورت تصادفی پُر می‌نماییم. در نهایت مقدار $t=0$ را قرار می‌دهیم.

(۲) ماتریس عضویت را (در $t=0$) محاسبه یا (در $t > 0$) به‌روزرسانی می‌کنیم. یعنی درجه‌ی عضویت برای پارامترهای ثابتی از خوشه‌ها به‌صورت زیر بهینه‌سازی می‌شوند (رابطه‌ی ۵).

$$u_{ij}^{(t+1)} = \frac{d_{ij}^{-2/(m-1)}}{\sum_{i=1}^c d_{ij}^{-2/(m-1)}} = \frac{1}{\sum_{l=1}^c \left(\frac{d_{lj}}{d_{ij}}\right)^{1/(1-m)}} \quad \text{رابطه‌ی ۵:}$$

for $i=1, \dots, c$ and $j=1, \dots, N$



شکل ۶ - نمایشی از عملکرد سیستم در کلاس‌بندی داده‌ها

کلاس‌بندی میانگین فازی، در جدول ۹ مشخص می‌شوند.

براساس پارامترهای جدول ۸ نتایج پارامترهای دقت، حساسیت و ویژگی سیستم پیشنهادی، با توجه به کلاس‌بندی الگوریتم

جدول ۸- پارامترهای ۳ گانه در ارزیابی سیستم پیشنهادی

مقادیر واقعی	مقادیر پیش‌بینی شده
Tn (منفی درست)	Fp (مثبت نادرست)
Fn (منفی نادرست)	Tp (مثبت درست)
$Sen = \frac{Tp}{Tp + Fn}$	$Spe = \frac{Tn}{Tn + Fp}$
کسر حساسیت	کسر ویژگی
صحت (درستی)	$Accuracy = \frac{Tp + Tn}{Tp + Fn + Tn + Fp}$

نتایج اعمال این پارامترها بر روی نتایج الگوریتم دسته‌بندی میانگین فازی در جدول ۹ مشخص می‌شود.

جدول ۶- نتایج حاصل از اعمال پارامترهای جدول ۸ در دسته‌بندی داده‌ها

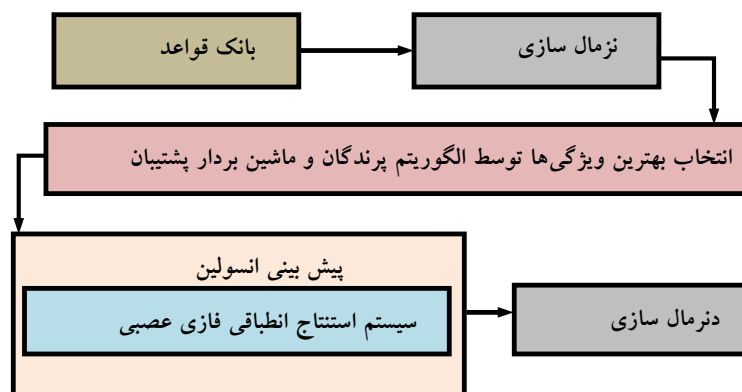
گزینه	الگوریتم پیشنهادی
حساسیت (%)	۹۲/۴
ویژگی (%)	۹۱/۱
دقت نهایی (%)	۸۸/۲

منظور فلوچارت ۳ را برای پیش‌بینی صحیح دُز انسولین در بیماران پیشنهاد می‌نماییم که شرح آن به این صورت است که ابتدا داده‌ها توسط رابطه‌ی ۷ در بازه‌ی [0 1] نرمالیزه می‌شوند، سپس در الگوریتم نهایی ترکیبی از بهترین ویژگی‌ها استخراج شده و دو ویژگی مرحله‌ی قبل براساس بردار هدف توسط سیستم انطباقی فازی - عصبی، مدل‌سازی دقیقی به منظور پیش‌بینی تهیه می‌شود. تمام مطالب در فلوچارت ۳ مشخص شده است.

$$y_i = \frac{x_i - m}{\sigma_i^2} \quad \text{رابطه‌ی ۷:}$$

تا این قسمت از کار بیان فازی پیشرفت بیماری و یا سلامت در افراد مشخص شده است. از این ۲ ویژگی در تعیین دُز صحیح انسولین برای بیماران استفاده می‌نماییم.

سیستم پیشنهادی در پیش‌بینی دُز انسولین برای بیماران سیستم پیشنهادی که به منظور تعیین دُز صحیح انسولین برای بیماران در نظر گرفته شده است، به این ترتیب می‌باشد: با استفاده از ۱۲ ویژگی ثبت شده و ۲ ویژگی منتج از مرحله‌ی قبل به کمک الگوریتم‌های داده کاوی و سیستم‌های هوشمند سعی می‌شود تا پیش‌بینی صحیح و درستی از دُز انسولین برای بیماران به دست آید. به ترتیب به مرور و بررسی این قسمت از کار می‌پردازیم. به همین

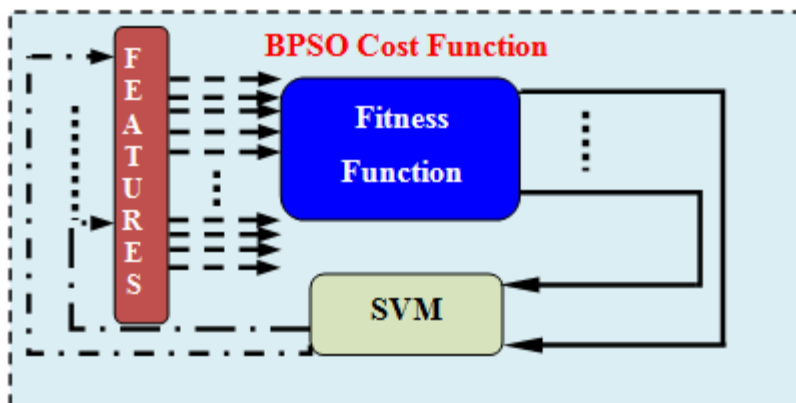


فلوچارت ۳- فلوچارت پیشنهادی در قسمت دوم الگوریتم به منظور پیش‌بینی دُز انسولین

به ترتیب معرف موقعیت و سرعت فعلی در لحظه (t) از ذره i هستند مشخص می‌شود. در هر مرحله از حرکت جمعیت، مکان هر پرنده با دو مقدار بهترین (p_best, g_best) به روز می‌شود. در هر تکرار، الگوریتم بعد از یافتن سرعت و موقعیت جدید پرنده را براساس روابط ۸ و ۹ به‌روزرسانی می‌کند. روند روش پیشنهادی در الگوریتم BPSO در فلوچارت ۴ می‌باشد.

الگوریتم حرکت دسته جمعی ذرات

الگوریتم حرکت دسته جمعی ذرات از جمله الگوریتم‌های جستجوی موازی مبتنی بر جمعیت است که توسط Eberhart, Kennedy در سال ۱۹۹۵ مطرح شد [۳۹، ۴۰] که با یک گروه از جواب‌های تصادفی (پرنده‌گان یا ذرات) شروع به کار می‌کند. هر ذره (پرنده) به صورت چند بعدی با دو بردار $V_i[t]$ و $X_i[t]$ که



فلوچارت ۴- تابع هزینه‌ی الگوریتم حرکت دسته جمعی ذرات

$$V_i[t+1] = w \times v_i[t] + C_1 r_1 (X^{i, pbest}[t] - X_i[t]) + C_2 r_2 (X^{gbest}[t] - X_i[t]) \quad \text{رابطه‌ی ۸:}$$

$$X_i[t+1] = X_i[t] + V_i[t+1] \quad \text{رابطه‌ی ۹:}$$

برای تغییر سرعت و ب: اعمال تابع سیگموئید در معادله‌ی ۱۰ برای محدود کردن سرعت بین [0,1] رابطه‌ی ۱۰:

$$X_{ij}[t+1] = 0 \quad S(v_{ij}[t+1]) = \frac{1}{1 + e^{-v_{ij}[t+1]}}$$

در اینجا به منظور بهینه‌سازی فضای جواب‌های مسأله از مدل الگوریتم باینری حرکت دسته جمعی ذرات (BPSO) استفاده می‌کنیم. با در نظر گرفتن یک رشته‌ی دودویی ۱۲ بیتی برای هر ذره، الگوریتم BPSO شروع می‌شود مانند زیر:

1 0 1 0 0 1 0 1 1 0 1 0

می‌باشد. البته در اینجا دقت شناسایی مهم‌تر از کوچک بودن زیر

در معادله‌ی ۸، W ضریب اینرسی است (با در نظر گرفتن رفرنس شماره [۴۱])؛ که در بازه‌ی [0-1.25] قرار دارد. C_1 و C_2 میزان تجربه‌ی شخصی و گروهی است که به طور معمول برای هر دوی آن‌ها از مقدار ۱/۵ و یا ۱/۲ استفاده شده است [۱۰، ۴۲]. r_1 و r_2 نیز اعداد تصادفی در بازه‌ی [0-1] می‌باشند. برای جلوگیری از واگرایی الگوریتم، سرعت پرنده به یک بازه محدود می‌شود، که V عضوی از دامنه $[Vmin, Vmax]$ می‌باشد. معادله‌ی ۹ نیز بردار موقعیت فعلی پرنده را با توجه به سرعت جدید آن به‌روزرسانی می‌کند. تغییر معادله‌ی به‌روزرسانی الگوریتم باینری در ۲ گام انجام می‌شود، الف: استفاده از رابطه‌ی ۸

"0" نمایانگر حذف ویژگی و "1" نمایانگر انتخاب آن ویژگی

از β خواهد بود، Accuracy یا میزان صحت، از مقایسه‌ی بین بردار بیماران شناسایی شده توسط سیستم و بیماران واقعی به‌دست می‌آید.

مجموعه‌ی منتخب است، بنابراین با توجه به نتایج مطلوب [۲۹، ۱۰] تابع رابطه‌ی ۱۱ را پیشنهاد می‌کنیم.

$$Fitness = \alpha \cdot Accuracy + \beta \cdot \frac{|n| - |s|}{|n|}$$

رابطه‌ی ۳: $|n|$ کل ویژگی‌ها و $|s|$ ویژگی‌های منتخب است. جمله‌ی اول ضریبی از دقت شناسایی و جمله‌ی دوم ضریبی از نرخ کاهش ویژگی‌هاست. مجموع ضرایب α و β را ثابت و برابر ۱۰۰ در نظر می‌گیریم. اینجا دقت شناسایی مهم‌تر است پس α بسیار بزرگ‌تر

جدول ۷- پارامترها و مشخصات الگوریتم بهینه‌سازی (کاوشی) BPSO

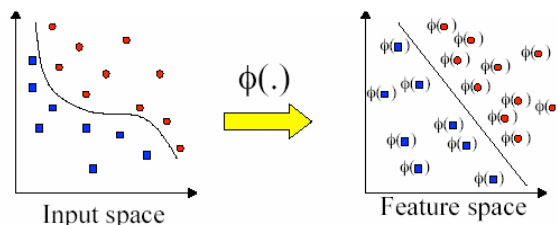
الگوریتم بهینه‌سازی	الگوریتم BPSO
جمعیت اولیه	۲۰ ذره
Maxiter	۱۰۰ تلاش
C1 یا ضریب تجربه شخصی	1 [23, 43]
C2 یا ضریب تجربه گروهی	20 [23, 43]
W یا ضریب اینرسی	بین ۰ تا ۱ بصورت نزولی [43]
Vmax	[23, 43]+4V
Vmin	[23, 43]-4V

قرار گیرد [۳]. این الگوریتم می‌تواند داده‌های آموزش را به دو دسته تقسیم کند، اصولاً داده‌ها توسط یک خط قابل جداسازی هستند، و در غیر این‌صورت الگوریتم با نگاشت ویژگی‌ها به‌وسیله تابع ϕ به ابعاد بالاتر سعی می‌کند توسط یک فرا صفحه با دقت بالاتری داده‌ها را تفکیک نماید در شکل ۷ مشخص است.

BPSO با چند زیر مجموعه‌ی تصادفی شروع به کار می‌کند و سعی دارد تعدادی زیر مجموعه را که حاوی مفیدترین اطلاعات هستند، به ما معرفی کند. بردارهای ویژگی معرفی شده توسط BPSO برای آموزش طبقه‌بندی کننده‌ی SVM به کار برده می‌شوند.

الگوریتم طبقه‌بندی SVM

SVM یک الگوریتم یادگیری با سرپرستی است که می‌تواند در کاربردهایی نظیر جداسازی و طبقه‌بندی داده مورد استفاده



شکل ۷- نگاشت داده‌ها از فضای ورودی به فضای ویژگی [۳]

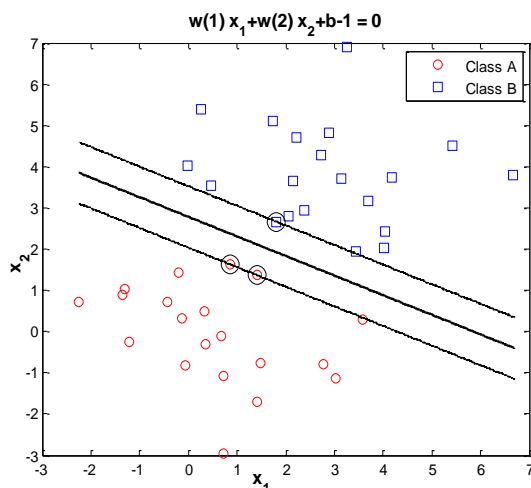
بین دو صفحه حاشیه برابر $\frac{2}{\|w\|}$ می‌باشد.

رابطه‌ی ۱۲ رابطه‌ی فرا صفحه طبقه‌بندی کننده داده‌ها است. روابط ۱۳ و ۱۴ روابط فرا صفحه‌های موازی براساس شرایط حداکثر حاشیه می‌باشند. در صورت ترسیم این توابع ساده فاصله

- رابطه‌ی ۴ $w \times x - b = 0$
- رابطه‌ی ۵ $w \times x - b = -1$
- رابطه‌ی ۶ $w \times x - b = 1$

بیماران از سالم‌ها، از تابع خطی استفاده شده است. در اثر اعمال این الگوریتم در ویژگی سوم، شکل ۸ به دست می‌آید.

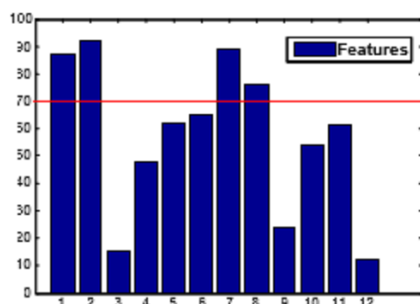
x متغیر ورودی، w بردار نرمال خط جدا کننده و b عرض از مبدا خط جدا کننده است. کلاس‌بندی نمونه‌های خروجی از الگوریتم BPSO با SVM است. به منظور کاهش محاسبات در جداسازی



شکل ۸- طبقه‌بندی داده‌ها به ۲ گروه در ویژگی سوم

می‌شوند؛ ویژگی‌های شماره ۱ و ۲ و ۷ و ۸ (ویژگی‌های بهینه و مؤثر) که البته در شکل ۹ در ۱۰۰ اجرای سیستم پیشنهادی، قابل تشخیص هستند.

پس از اجرای مراحل مختلف الگوریتم پیشنهادی به منظور مشخص شدن بهترین و مؤثرترین ویژگی‌ها، یعنی اعمال الگوریتم‌های بهینه‌سازی (کاوشی) و الگوریتم بردار ماشین پشتیبان (SVM)، ویژگی‌های مؤثر مشخص



شکل ۹- مشخص شدن بهترین و مؤثرترین ویژگی‌ها از غیر مؤثرترین ویژگی‌ها

به همراه بردار هدف به منظور پیش‌بینی صحیح میزان دُز انسولین براساس الگوی معرفی شده توسط ویژگی‌ها باید به مرحله بعد، یعنی شبکه فازی - عصبی، اعمال شود. در

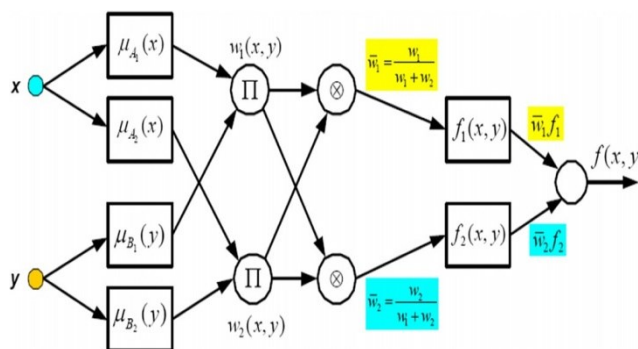
بعد از اینکه ویژگی‌های مهم مشخص شدند، بانک اطلاعاتی با ابعاد ۳۱۲ سطر (افراد مورد مطالعه) و ۶ ستون (۴ ویژگی بهینه به همراه ۲ ویژگی در بیان فازی بیماری یا سلامت افراد) مرتب می‌شود، سپس بانک اطلاعاتی جدید

تطبیقی می‌باشد. ساختار ANFIS مورد استفاده در شکل ۴ دیده می‌شود که دایره‌ها بیان‌کننده‌ی گره‌های ثابت و مربع‌ها بیان‌کننده‌ی گره‌های تطبیقی هستند. در لایه‌ی اول، تمامی گره‌ها تطبیقی هستند و خروجی این لایه، درجه‌ی عضویت فازی ورودی‌هاست.

قسمت بعد ساختار، طرز کار و نتایج حاصل از این الگوریتم معرفی می‌شود.

شبکه‌های انطباقی فازی عصبی (ANFIS)

الگوریتم ANFIS (Adaptive Neuro Fuzzy Inference System) مورد استفاده یک مدل فازی سوگنو در قالب یک سیستم



تنظیم می‌گردد [۴۵]. اگر یک سیستم فازی سوگنو با دو ورودی x و y یک خروجی F داشته باشد، برای یک مدل فازی سوگنو مرتبه‌ی اول، مجموعه قوانین مشترک با دو قانون If-then به صورت روابط ۱۵ و ۱۶ است. رابطه‌ی ۷ و ۸:

در این سیستم‌ها با استفاده از مجموعه‌ی داده‌های ورودی-خروجی به‌عنوان داده‌های آموزشی، یک سیستم استنباط فازی از نوع سوگنو ایجاد می‌شود که پارامترهای تابع عضویت آن با استفاده از الگوریتم پس انتشار و روش کوچک‌ترین مربعات

$$\text{Rule1 : If } X \text{ is } A_1 \text{ and } Y \text{ is } B_1 \text{, then } f_1 = \mu_1 x + q_1 y + r_1$$

$$\text{Rule2 : If } X \text{ is } A_2 \text{ and } Y \text{ is } B_2 \text{, then } f_2 = \mu_2 x + q_2 y + r_2$$

ساختار الگوریتم ANFIS استفاده شده در این تحقیق با اقتباس از [۴۵] در جدول ۱۱ بیان شده است.

جدول ۸- پارامترهای ساختاری در الگوریتم ANFIS

روش‌های ساخت سیستم استنتاج فازی	
FCM	Sub. Clustering
روش‌های آموزش	
Back Propagation	
Hybrid	Least Square-Back Propagation
Grid Partitioning	n Input MFs = 10
Sub. Clustering	Rad = 0.3
FCM	N Cluster = 10
مشخصات و تنظیمات روش آموزش	
Max Epoch	50
Step Size Dec Rate	0.9
Step Size Inc Rate	1.1
Error Goal	0
Initial Step Size	0.01

در قسمت‌هایی به‌منظور ترسیم نمودارها از نرم‌افزار ORIGIN استفاده می‌نماییم.

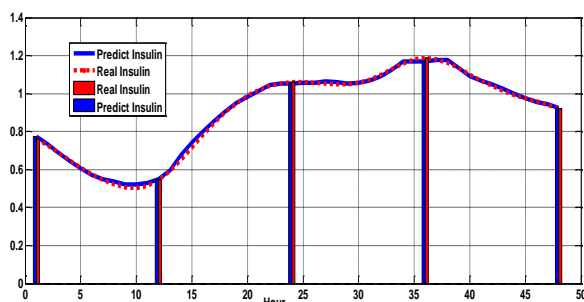
نتایج ارزیابی و آزمایش‌های نهایی

ویژگی‌های اصلی، سرنوشت ساز و تأثیر گذار در تعیین میزان دُز انسولین در بیماران تعیین شدند. سپس شبیه‌سازی نهایی توسط شبکه‌عصبی فازی-مصنوعی (ANFIS) مطابق شکل ۱۷ برای ۴۸ ساعت آینده بیمار به‌صورت ۱۲ ساعتی تهیه شده است. البته مقادیر پس از اینکه توسط شبکه فازی - عصبی پیش‌بینی شدند، توسط رابطه‌ی ۱۷ دنرمالیزه می‌شوند. پس از دنرمالیزه شدن، نتایج شبیه سازی در شکل ۱۱ و ۱۲ برای بیمار تصادفی شماره‌های ۸ مشخص شده است. رابطه‌ی ۹:

$$x_i = y_i \times \sigma_i + m$$

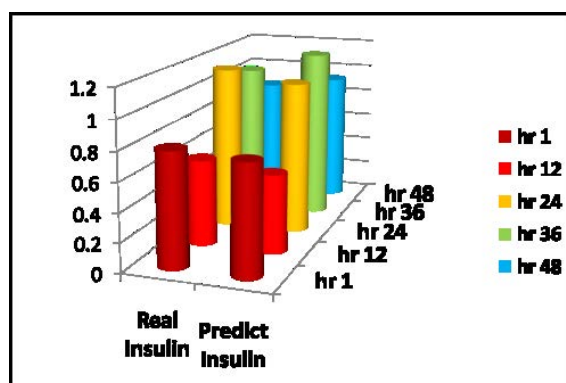
بررسی صحت آزمایش و ارزیابی نهایی

فرآیند انتخاب ویژگی که منجر به پیش‌بینی می‌شود اینگونه است: توسط الگوریتم باینری حرکت دسته جمعی ذرات براساس تابع هزینه تعریف شده (کاوش به‌منظور بررسی ویژگی‌ها) و الگوریتم طبقه‌بندی ماشین بردار پشتیبان (SVM) مؤثرترین ویژگی‌ها مشخص می‌شوند، سپس شبکه فازی عصبی و بانک داده مبتنی بر ویژگی‌های مؤثر، پیش‌بینی صحیحی از میزان دُز انسولین برای بیماران تهیه می‌کند. برای اثبات قابلیت‌های سیستم پیشنهادی، اعتبار سنجی عرضی ۱۰-دسته‌ای (توسط K-fold) بر روی پایگاه داده ۳۱۲ نمونه‌ای موجود پیاده شد که به ۱۰ دسته تقسیم شدند. ۸ دسته ۳۱ داده‌ای و ۲ دسته ۳۲ داده‌ای. در این اعتبار سنجی، ۹ قسمت از داده‌ها برای آموزش و ۱ قسمت برای آزمایش به‌کار می‌روند، آموزش و آزمایش ۱۰ بار به‌صورت متوالی چرخشی انجام می‌شود به‌طوری که هر بار یک قسمت برای آزمایش کنار گذاشته می‌شود. به‌منظور انجا محاسبات از نرم‌افزار MATLAB و



شکل ۱۱- پیش‌بینی دُز انسولین در بیماران براساس سیستم پیشنهادی بر اساس مقادیر واقعی

در شکل ۱۲ نتایج به‌دست آمده برای پیش‌بینی دُز انسولین در ترسیمی ستونی برای همان بیمار مشخص شده است.

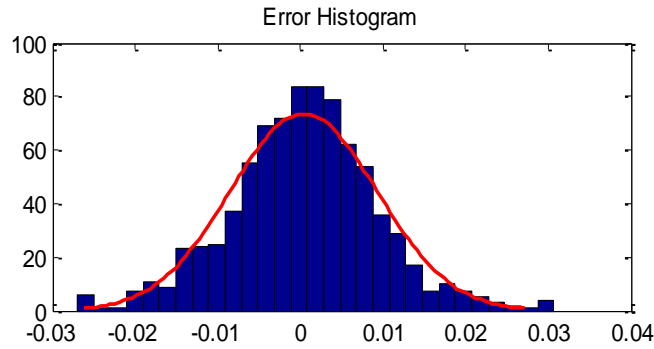


شکل ۱۲- مقادیر پیش‌بینی شده و واقعی دُز انسولین در بیمار شماره‌ی ۹

به‌منظور بررسی و مقایسه‌ی عملکرد سیستم پیشنهادی با تجربیات و نظرات پزشک (سیستم پزشک) از منحنی ROC (Receiver Operating Characteristics) استفاده می‌نمایم.

به‌منظور بررسی خطای عملکرد شبکه‌ی فازی عصبی، رابطه‌ی ۱۸ تعریف می‌شود، نمودار آن در شکل ۱۳ رسم شده است.

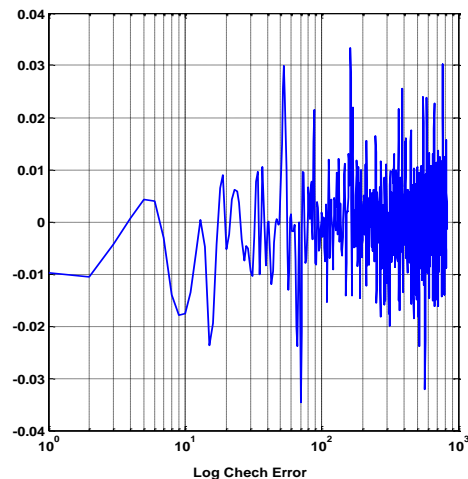
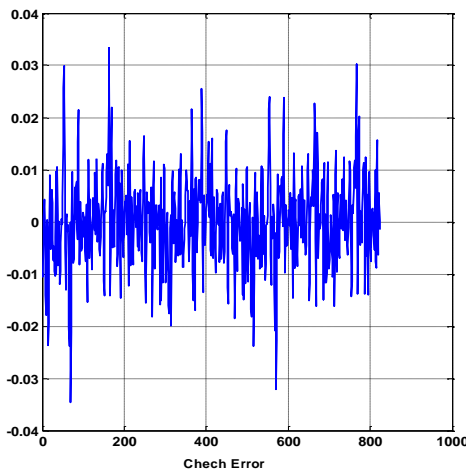
$$e = \sum_{i=1}^n \frac{\|y_i - \hat{y}_i\|^2}{n} \quad \text{رابطه‌ی ۱۰}$$



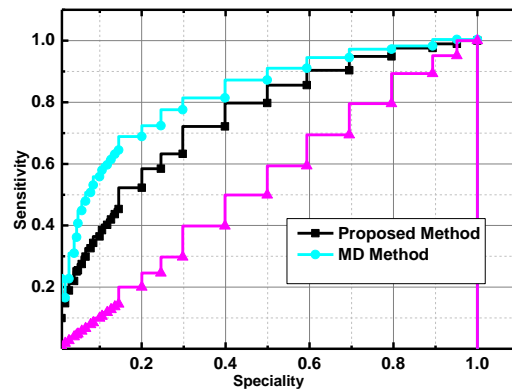
شکل ۱۳- خطای عملکرد شبکه‌ی فازی عصبی در پیش‌بینی طبقه بندی صحیح و بیان عملکرد بین سیستم‌ها در شرایط کاری یکسان.

که بیان‌کننده‌ی احتمال مثبت یک تست است که به بیماران مبتلا نسبت داده شده است. ویژگی در یک آزمون، بیان‌کننده احتمال منفی یک تست می‌باشد که به بیماران نسبت داده شده است [۴۶]. نمودار ROC هر ۲ سیستم در شکل ۱۴ برای ۳۰ مورد بیمار تصادفی آزمایشی است که میزان حساسیت و ویژگی آن‌ها براساس رفتار سیستم پیشنهادی و عملکرد پزشک مشخص شده است.

به‌منظور بررسی خطای سیستم، از شکل ۱۴ یعنی بیان لگاریتمی استفاده می‌نمایم، این شکل به‌خوبی مشخص می‌نماید که نرخ تغییرات خطای سیستم آهسته و باند یا محدوده‌ی آن نیز بسیار کوچک می‌باشد و همان‌طور که در شکل ۱۵ می‌بینید، مشاهده‌ی حساسیت و ویژگی در هر سیستم بر حسب دیگری بیان شده است. در موارد و شرایط ایده‌آل کاری، با افزایش حساسیت، کاهش کمی در ویژگی وجود دارد، تا زمانی که به سطح بسیار بالایی از حساسیت رسیده باشیم. حساسیت یک آزمون می‌باشد



شکل ۱۴- نمایی از بیان لگاریتمی خطا



شکل ۱۵- منحنی ROC برای مقایسه‌ی سیستم واقعی و پیشنهادی در داده‌های آزمایشی

مقادیر پارامترهای حساسیت و ویژگی برای ۳۰ بیمار آزمایش در جدول ۱۲ براساس جدول ۸ مشخص شده‌اند.

جدول ۱۲- پارامترهای ویژگی و حساسیت در ۳۰ بیمار

آزمایشی در سیستم پیش بینی دُر انسولین

Item	Proposed ALG
Sensitivity (%)	95.1
Specialty (%)	90
Total Accuracy (%)	83.8

قند خون تا ۴۸ ساعت آینده را پیش‌بینی نماید. از شبکه‌های تطبیقی فازی عصبی برای انطباق، یادگیری استفاده شد. چالش آموزش این شبکه‌ها، مهم‌ترین دغدغه می‌باشد. ترکیب الگوریتم‌های تکاملی و سیستم‌های داده کاوی می‌تواند یک ایده برای افزایش کارایی این شبکه‌ها در جهت پیش‌بینی صحیح باشند. البته هدف در این تحقیقات به‌دست آوردن صحت و دقت ۱۰۰ درصدی می‌باشد، تا بتوان به سیستم‌های هوشمند مصنوعی اطمینان داشت و بعد عملی به آن‌ها بخشید. رسیدن به چنین هدفی تنها در صورت دانش بالا و تسلط کافی از یک طرف به سیستم‌های هوشمند و از طرفی دیگر به بیماری امکان‌پذیر است، که در آینده‌ایی نه چندان دور به آن نایل خواهد شد.

البته، هر چه قدر مساحت زیر نمودار سیستم پیشنهادی به سیستم واقعی نزدیک‌تر باشد نشان دهنده‌ی عملکرد نزدیک سیستم پیشنهادی به سیستم واقعی یا تصمیمات پزشکان است.

نتیجه‌گیری

در این تحقیق با توجه به ضرورت تعیین صحیح دُر انسولین برای بیماران، روشی نوین مبتنی بر ترکیب سیستم‌های هوشمند ارائه شده است، نتایج به‌دست آمده به‌میزان قابل توجهی نسبت به تحقیقات قبلی بهبود یافت. همچنین در مقایسه با نتایج پزشکان، نشان دهنده‌ی عملکرد خوب در صحت پیش‌بینی سری زمانی غلظت قند خون است، چرا که سیستم پیشنهادی موفق شد، سطح

مآخذ

1. World Health Organization, Diabetes Center, Fact Sheet N°312, www.who.int/mediacentre/factsheets/fs312/en
2. Centers for Disease Control and Prevention, National Diabetes, Fact Sheet 2011 www.cdc.gov/diabetes
۳. محمدرضا نعیم آبادی، نوشاز امیر احمدی چماچار، احسان تهامی و حسین ربانی. تشخیص دیابت با استفاده از SVM. چهاردهمین کنفرانس دانشجویی مهندسی برق ایران، ISCEE 14th، تابستان ۱۳۹۰.
4. American Diabetes Association, Diabetes Basics www.diabetes.org/diabetes-basics.
5. Kuncheva LI; Bezdek JC; Duin RPW. Decision templates for multiple classifier fusion: an experimental comparison. *Journal of Pattern Recognition* 2001, 34(2):299-314.
6. Bi Y, Bell D; Wang H; Guo G; Guan J. Combining Multiple Classifiers Using Dempster's rule for text CaregORIZATION. *Journal of Applied Artificial Intelligence* 2007; 21(3):211- 239.
7. Kuncheva LI. *Combining Pattern Classifiers, Methods and Algorithms*, New York, NY: Wiley Interscience, 2005.
8. Tulyakov S, Jaeger S, Govindaraju V, Doermann D. Review of Classifier combination Methods. *Journal of Studies in Computational Inteligence* (SCI) 2008; 90 (4):361-386.
9. Jaafar SFB and Ali DM. Diabetes Mellitus Forecast using artificial neural network (ANN). Sensors and the International Conference on new Techniques in Pharmaceutical and Biomedical Research, 2005 Asian Conference on. IEEE, 2005; 135 – 139.
۱۰. سید احسان تهامی، محمد علی خلیل زاده. تشخیص هوشمند بیماری دیابت با استفاده از شبکه‌های عصبی چند لایه. دوازدهمین کنفرانس مهندسی پزشکی ایران، دانشگاه صنعتی سهند تبریز ۲۵-۲۷، آبان ۱۳۸۴.
11. Jiang Y and Zhou Z. Editing Training Data for kNN Classifiers with Neural Network Ensemble. *Advances in Neural Networks – ISNN 2004* 2004; pp 356-361.
۱۲. محمد فیوضی، اعظم قره‌خانی، جواد حدادنیا. تشخیص بیماری دیابت بر اساس ترکیب و تعامل سیستم‌های فازی، درخت تصمیم‌گیری و سیستم‌های فازی عصبی. چهارمین کنفرانس ملی فناوری اطلاعات و دانش، اردیبهشت ۱۳۹۱، دانشگاه صنعتی بابل.
۱۳. محمد فیوضی، جواد حدادنیا، نسرین ملانیا. تشخیص بیماری دیابت براساس روش ترکیبی هوشمند. مجله‌ی هوش مصنوعی و ابزار دقیق ایران، شماره ۳۲، تابستان ۱۳۹۱.
۱۴. سعید راحتی قوچانی، سیداحسان تهامی. پیش‌بینی نوسانات سطح قندخون در بیماران مبتلا به دیابت نوع ۱ با استفاده از شبکه‌های عصبی خود بازگشتی المن. مجله‌ی فنی و مهندسی دانشگاه آزاد اسلامی مشهد، دوره‌ی دوم، شماره‌ی اول، پاییز ۱۳۸۷.
15. Scott MP, Brent DC, Paul MR. Development of a Neural Network for Prediction of Glucose Concentration in Type 1 Diabetes Patients. *Journal of Diabetes Science and Technology* 2008; 2 (5).
16. Mougiakakou SG, Prountzou A, Iliopoulou D, Nikita KS, Vazeou A, Bartsocas CS. Neural Network based Glucose – Insulin Metabolism Models for Children with Type 1 Diabetes. *Conf Proc IEEE Eng Med Biol Soc* 2006; 1:3545-8.
17. Wahab A, Kong YK, and Quek C. Model Reference Adaptive Control on Glucose for the Treatment of Diabetes Mellitus. *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems* 2006.
18. Hsiao-Ping Huang, Shih-Wei Liu, I-Lung Chien, Chia-Hung Lin. A Dynamic Model with Structured Recurrent Neural Network to Predict Glucose-Insulin Regulation of Type 1 Diabetes Mellitus. *9th International Symposium on Dynamics and Control of Process Systems* (DYCOPS 2010), Leuven, Belgium, July 5-7, 2010.
19. Lean Yu. An evolutionary programming based asymmetric weighted least squares support vector machine ensemble learning methodology for software repository mining. *Journal of Information Sciences* 2012; 191 (4), 31-46.
20. NASA MDP repository, 2004. <http://mdp.ivv.nasa.gov/>.

21. Fernandez de Canete J, Gonzalez-Perez S, Ramos-Diaz JC. Artificial neural networks for closed loop control of in silico and ad hoc type 1 diabetes. *Comput Methods Programs Biomed.* 2012; 106(1):55-66.
22. Jensen R. *Combining rough and fuzzy sets for feature selection*. Ph.D. Thesis, School of informatics, Univ. Edinburgh, 2005
۲۳. محمد علیپور، جواد حدادنیا. معرفی یک سیستم دقیق تشخیص سرطان پستان. فصلنامه‌ی بیماری‌های پستان ایران، شماره‌ی دوم، تابستان ۱۳۸۸.
24. Xiang T, and Gong S. Spectral clustering with eigenvector selection. *Journal of Pattern Recognition* 2008; 41(3):1012-1029.
۲۵. سهیلا اشک زری طوسی، هادی صدوقی یزدی. خوشه‌بندی طیفی با انتخاب ویژگی Kernel PCA. شانزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران، دانشگاه صنعتی شریف، تهران اسفند ۱۳۸۹.
26. Kanan H, Faez K, Hosseinzadeh M. Face recognition system using ant colony optimization-based selected features. *Proceeding IEEE Symp. Computational Intelligence in Security and Defense Applications: 2007*; 57-62.
27. Siedlecki W, Sklansky J. A note on genetic algorithms for large-scale feature selection. *Journal of Pattern Recognition Letters* 1989; 10(5): 335-47.
28. Meng Y. A swarm intelligence based algorithm for proteomic pattern detection of ovarian cancer. *IEEE Symp. Computational Intelligence and Bioinformatics and Computational Biology (CIBCB)* 2006; 1-7.
29. Newman DJ, Hettich S, Blake CLS, & Merz CJ. UCI "repository of machine learning database", Irvine, CA: University of California, 1998.
30. Mitra S, Konwar KM, and Pal SK. Fuzzy Decision Tree, Linguistic Rules and Fuzzy Knowledge-Based Network: Generation and Evaluation. *Journal of IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 2002; 32: 328-339.
۳۱. سید محمد محمد زاده ضیابری، دکتر عبادزاده. پروژه‌ی پایانی سیستم‌ها و روش‌های فازی. دانشگاه صنعتی امیر کبیر، تیرماه ۱۳۸۸
32. Sugeno M. *Industrial Applications of Fuzzy Control*. Elsevier, Book, New York, 1985.
33. Bergh F. and Engelbrecht A. A new locallyconvergent particle swarm optimizer. *Fundamenta Informatic Journal IEEE Transaction* 2002; 8 (2).
34. Herrera F, and Lozano M. Fuzzy adaptive genetic algorithm: for design, taxonomy, and future direction. *Journal of Soft Computing* 2003; 7(8):545-562.
۳۵. جواد محجوبی، اف.ا. شهیدی. تخمین ارتفاع امواج ناشی از باد در نكء به کمک درختان تصمیم‌گیری. ۷ کنفرانس JDMC، 20-21 November 2007، دانشگاه صنعتی امیر کبیر.
36. Kantardzic M. *Data Mining: Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, 2003.
۳۷. محمد فیوضی، خسرو رضایی، جواد حدادنیا. بهبود همگرایی در الگوریتم حرکت دسته جمعی ذرات براساس منطق فازی. سومین کنفرانس ملی برق و الکترونیک ایران، گناباد، تیرماه ۱۳۹۰.
38. Baraldi A. and Blonda P. A Survey of Fuzzy Clustering, Algorithms for Pattern Recognition—Part I and II. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 1999; 29(6).
39. Pan Hongxia, Ma Qingfeng, Wei Xiuye. Research on Fault Diagnosis of Gearbox Based on Particle Swarm Optimization Algorithm. *IEEE 3rd International Conference on Mechatronics, ICM Conference* 3-5 July 2006.
40. Sheybani M, Meibody M. PSO-LA: A Novel of Optimization. *12th Conference on Computer Engineering Iran*, 2006; pp: 1162.
41. Feng X, Zhang J, Yang Z. Adaptive Particle Swarm Optimization on Individual Level. *IEEE confrence, China*, 2002; pp: 1215- 1218.
42. Norozie Beyrami M. *Improve the convergence of the algorithm of particle swarm*. *Conference on industrial engineering Osco IAU Branch*, 2003.
43. Kennedy J. and Eberhart R. Particle swarm optimization. *Conference on Evolutionary Computation Proceeding IEEE* 1995; pp. 1942-1948, USA.
۴۴. سید احسان تهامی، سید محمد بامشکی، محمد علی خلیل زاده. تشخیص بیماری دیابت نوع ۱ با استفاده از الگوریتم ANFIS، GA-NN. اولین کنفرانس مشترک

سیستم‌های هوشمند و سیستم‌های فازی، شهریور ۱۳۸۶،

دانشگاه فردوسی مشهد.

45. Rajeswari K, Vaithyanathan V. Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network. *International Journal of Computer Science and Network Security*, 2011;11(4).

PREDICTION AND DETERMINATION OF THE CORRECT DOSE OF INSULIN IN DIABETIC PATIENTS BASED ON DIABETES FUZZY DIAGNOSIS

Mohammad Fiuzy^{*1}, Javad Haddania², Nasrin Mollania³

1. *Elderology Research Center, Sabzevar University of Medical Sciences, Sabzevar, Iran*

2. *Biomedical Engineering division, Electrical Engineering Faculty, Hakim Sabzevari University, Sabzevar, Iran*

3. *Biochemistry division, Basic Science Faculty, Hakim Sabzevari University, Sabzevar, Iran*

ABSTRACT

Background: On time diabetes diagnosis dramatically reduces the many injuries and damage in the community. Diabetes is a disease that requires a lot of care in addition to prevention, such as prediction the correct level of blood sugar fluctuations. The most important complications of such disease are anesthesia, coma and even death at final. Today, in these patients, determining the correct dose of insulin is based on the experience and knowledge of physicians along with the interaction of patients with them, although human error is inevitable.

Methods: This study includes 124 patients and 188 healthy suspects were examined based on 21 features which hold by 7 features for diagnosis and 14 features for predicting insulin dose. The proposed system was presented to identify or diagnose the disease at first, and finally the correct doses of insulin for patients have been determine. The proposed system has two stages (which include diagnosis and prediction) and several subsystems. In the diagnosis phase, some sub systems such as the Fuzzy system for the purpose of accurately estimating the disease progression in patients and the decision tree (DT) for the preparation of rules in the fuzzy system (the process of mapping the attribute space (individuals) to the output (the diagnostic result)) have used. Also, in the prediction phase of insulin dose, the BPSO algorithms are used to select the best features. Classification algorithms (SVMs) are used to categorize effective to non-effective and adaptive artificial neuropsychological (ANFIS) systems for ultimate patient prediction have used.

Results: The proposed system, based on the best features in the provided data base in the form of the combination and interaction, succeeded to achieve a 95.1% precision, of course due to comparing by other commonly used methods and its performance the proposed method have high precision.

Conclusion: The results were significantly improved compared to previous studies. Also, in comparison with the results of physicians, it is indicative of good performance in predicting the accuracy of the time series of glucose concentration because the proposed system succeeded in predicting blood sugar levels for up to 48 hours.

Keywords: Diabetes, Diagnosis, Prediction, Fuzzy Systems, Classification Algorithm, Adaptive Neuro Fuzzy Inference System

* Medical Engineering Department, Faculty of Electrical and Computer Engineering, Hakim Sabzevar University, Sabzevar, Tohid Shahr, Khorasan Razavi, Iran. Postal Code: 9617976487. Tel-Fax: +98(57)14410104, Email: mohammad.fiuzy@yahoo.com