مجله دیابت و لیپید ایران. دو ماهنامه بهمن- اسفند ۱۳۹۱؛ دوره ۱۲ (شماره ۳) ۲۵۹-۲۴۵

پیش بینی صحیح دُز " انسولین و گلیبن‌کلامید" در بیماران دیابتی مبتنی بر ترکیب سیستم‌های هوشمند و سابقه بیماری

محمد فیوضی[۱،۲،*]، جواد حدادنیا[۱]، نسرین ملانیا[۳]، مریم هاشمیان[٤]، کاظم حسن پور[٤]

چکیده

مقدمه: یکی از عوارض خطرناک بیماری دیابت نوع ۱، افزایش و یا کاهش ناگهانی سطح غلظت قند خون می‌باشد که باعث بروز خطراتی مانند اغما و بیهوشی خواهد شد. انسولین و گلیبن‌کلامید، دو دارویی هستند که در اثر تجویز دُز صحیح، سطح غلظت قند خون را در نهایت، به درستی تنظیم می‌کنند تا از بروز چنین عوارضی پیشگیری شود. بنابراین استفاده از روشی مناسب به منظور پیش‌بینی یا تجویز صحیح این داروها و در نهایت پیشگیری از این عوارض، گام مهمی در جهت کنترل بهینه بیماری محسوب می‌شود. در این تحقیق ما بر آن هستیم، تا با استفاده از ترکیب الگوریتم‌های داده‌کاوی و هوش مصنوعی، پیش‌بینی صحیحی از دُز انسولین وگلیبن‌کلامید برای بیماران انجام دهیم.

روش‌ها: مراحل ایجاد سیستمی هوشمند به منظور تعیین و پیش‌بینی میزان صحیح انسولین و گلیبن‌کلامید بدین صورت است؛ که ابتدا لازم است تا شرایط و پارامترهایی که در میزان و دُز دارو (انسولین و گلیبن‌کلامید) برای بیماران دخیل هستند را شناسایی نماییم، سپس بانک جامعی مبتنی بر این ویژگی‌ها تشکیل دهیم، با همکاری تیم پزشکی مرکز تحقیقات دیابت شهرستان سبزوار، بانک جامعی در ۳ مرحله براساس اطلاعات ۱۱۰ بیمار و ۲۵۸ مورد مشکوک آن مرکز طراحی شد، سپس طبق قاعده سیستم‌های هوشمند و شبکه‌های عصبی مصنوعی، مدلی به منظور پیش‌بینی میزان دُز انسولین و گلیبن‌کلامید برای بیماران طراحی شد.

یافته‌ها: سیستم پیشنهادی با استفاده از ترکیب روش‌های مذکور موفق شد با تکیه بر ویژگی‌های پایگاه داده در قالب ترکیب و تعامل به دقت پیش بینی‌بیش از ۹۵٪ دست یابد. در مقایسه با روش‌های رایج از یکطرف و روش‌های مصنوعی از طرف دیگر (غالباً در بهترین حالت دقت و صحت قدرت پیش‌بینی آن‌ها در حدود ۸۵٪ برای مدت زمان کمتر از ۳ ساعت بوده است) به خوبی مشخص شد که، عملکرد سیستم پیشنهادی، مناسب‌تر، سریعتر و تا حدودی مطمئن‌تر از سایر روش‌های ترکیبی هوشمند است.

نتیجه‌گیری: این تحقیق با ارائه یک روش و مدل برگرفته از واقعیت و شرایط بیماران؛ میزان صحیح و تجویز شده دارو برای آنها را مشخص می‌کند، این در حالی است که نسبت به روند طبیعی تجویز توسط پزشکان، علاوه بر افزایش سرعت در تعیین میزان دارو، دارای صحت و مقبولیت و دقت قابل قبولی می‌باشد.

واژگان کلیدی: الگوریتم ژنتیکی، دیابت، دُز صحیح انسولین و گلیبن‌کلامید، شبکه‌های هوشمند مصنوعی

۱- بخش مهندسی پزشکی، دانشکده مهندسی برق و کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران
۲- مرکز تحقیقات فناوری‌های نوین پزشکی، دانشگاه علوم پزشکی سبزوار، سبزوار، ایران
۳- گروه زیست شناسی، دانشکده علوم پایه، دانشگاه حکیم سبزواری، سبزوار، ایران
٤- گروه علوم بالینی، دانشکده پزشکی، دانشگاه علوم پزشکی سبزوار، سبزوار، ایران

*نشانی: خراسان رضوی، شهرستان سبزوار، توحیدشهر، دانشگاه حکیم سبزواری، دانشکده مهندسی برق و کامپیوتر، بخش مهندسی پزشکی، کدپستی: ۹٦۱۷۹۷٦٤۸۷، تلفن: ۲۰۵۷۱٤٤۱۰۱۰٤، نمابر: ۰۵۷۱٤٤۱۰۱۰٤، پست الکترونیک: mohammad.fiuzy@yahoo.com

تاریخ دریافت: ۱۳۹۱/۱۰/۱۵ تاریخ درخواست اصلاح: ۱۳۹۱/۱۱/۲۰ تاریخ پذیرش: ۱۳۹۱/۱۱/۲۹

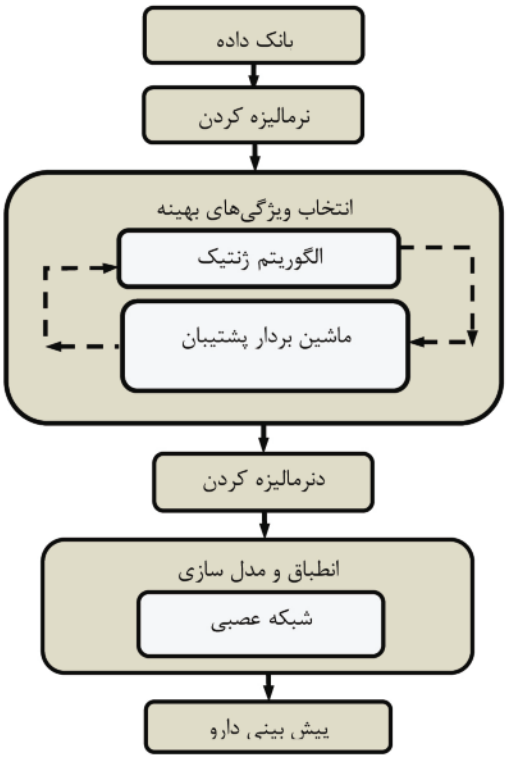
مقدمه

بیماری 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 [۳-۱]. دیابت عمدتاً از طریق آزمایش‌های قند خون (Plasma یا Glucose یا Fasting) تشخیص داده می‌شود که به ۳ دسته تقسیم می‌شود. دیابت نوع ۱ یا دیابت وابسته به انسولین (IDDM) بیشتر در سن پایین و کودکان دیده می‌شود. دیابت نوع ۲ یا دیابت غیر وابسته به انسولین (NIDDM) در ۹۰٪ تا ۹۵٪ بیماران دیابتی مشاهده می‌شود [۳]. در این دسته انسولین ترشح می‌شود، ولی بدن در برابر مصرف انسولین از خود مقاومت نشان می‌دهد. دیابت نوع ۳ بیشتر در زنان باردار دیده می‌شود که عمدتاً پس از بارداری به دیابت نوع ۲ تبدیل می‌شود [٤]. البته در زمینه تشخیص بیماری دیابت توسط همین مولفین تحقیقاتی [٥،٦] انجام شده است، اما در زمینه پیش‌بینی و تعیین صحیح دُز انسولین، تحقیقات چندانی انجام نشده است (تا آنجا که مولفین می‌دانند)، مشکل عمده‌ای که در حال حاضر در رابطه با این بیماری مخرب و خطرناک و به خصوص در نوع ۱ آن وجود دارد، افزایش سطح قند خون (Hyperglycemia) و یا کاهش بیش از حد و ناگهانی آن است که از جمله عوارض بسیار خطرناک آن محسوب می‌گردد، که می‌تواند به بیهوشی، اغماء و گاهی حتی مرگ بیمار منجر شود [۷]. بنابراین استفاده از روشی مناسب به منظور پیش‌بینی و تعیین میزان صحیح دز انسولین و حتی داروهایی دیگر همچون "متفورمین"، "گلیبن‌کلامید " و .... در جهت پیشگیری از این عوارض، گام مهمی برای کنترل بهتر این بیماری محسوب می‌شود.

با توجه به مطالب بیان شده و بررسی کارهای گذشته لزوم به نتیجه رسیدن این‌گونه پژوهش‌ها، بسیار مهم و سرنوشت‌ساز خواهد بود. از آنجا که در تحقیقات علوم پزشکی و اپیدمیولوژی اغلب مسئله سلامت انسان مطرح است، پیش‌بینی درست نتایج اهمیت بیشتری می‌یابد، بنابراین لازم است تا جایی که امکان دارد کمترین خطا و بیشترین اطمینان حاصل شود. با توجه به این که مسائل مرتبط با پاسخ‌های آمیخته و به وفور در مطالعات پزشکی مشاهده می‌شود و با در نظر گرفتن این که روش‌های موجود در آمار کلاسیک برای مدل‌بندی و پیش‌بینی، به دلیل محدودیت در عمل کارایی چندانی ندارد، ارائه روش‌هایی که بتواند راه‌گشای این گونه مسائل باشد، بسیار مفید و ارزنده به نظر می‌رسد. در این تحقیق برآن هستیم تا با ارائه مدلی دقیق و البته با کمترین خطا، میزان صحیح دُز دارو برای بیماران را بدست آوریم. ابتدا روش کلی را بیان می‌کنیم سپس با توجه به مطالب فلوچارت ۱ به بحث و بررسی رئوس مطالب می‌پردازیم. امید است تا این تحقیق گامی هرچند اندک ولی موثر در جهت بهبود روش‌ها و معالجه این بیماری داشته باشد.

روش‌ها

پس از مطالعه مختصر روی سیستم‌ها و تحقیقات انجام شده، سیستم پیشنهادی به گونه‌ایی که در فلوچارت ۱ مشخص شده است، طراحی و معرفی شد. بانک داده‌ایی از بیماران تشکیل شد، سپس براساس داده‌ها، که گونه‌ایی دنباله یا سری‌ایی در واحد زمان با مقادیر متفاوت هستند، توسط شبکه‌های عصبی مصنوعی، پیش‌بینی براساس مدل تشکیل شده صورت گرفت.



فلو چارت ۱- روند الگوریتم پیشنهادی به منظور تعیین دُز صحیح داده‌های مورد استفاده

الف: بانک داده

به منظور تعیین و پیش‌بینی دُز دارو برای بیماران، از مرکز تحقیقات دیابت شهرستان سبزوار، وابسته به دانشگاه علوم پزشکی شهرستان سبزواز یاری گرفته شد. به این ترتیب که با هماهنگی و همکاری تیم پزشکی و متخصص غدد، پرسش‌نامه‌هایی (به منظور مشخص کردن ویژگی‌های موثر) طراحی گردید. این پرسش نامه‌ها توسط ۳۶۸ فرد (۱۱۰ بیمار به همراه ۲۵۸ مورد مشکوک) ، که به صورت کاملاً تصادفی (سیستماتیک) انتخاب شده بودند، در طی ۳ مرحله تکمیل شد (از هر فرد ۳ بار نمونه‌گیری به عمل آمد). براساس نظر پزشکان سعی شد تا ویژگی‌هایی تعریف شوند که در روند درمان، مدل‌سازی رفتار بیمار و تصمیمات پزشکان بیشترین تاثیر را داشته باشند. بعد از مشخص شدن ویژگی‌ها، کدگذاری آن‌ها مهم است، که در پاره‌ایی از موارد مقدار دقیق عددی آزمایش را قرار می‌دهیم و در قسمت‌هایی دیگر لازم است تا مقدار کد شده قرار داده شود. به منظور کدگذاری از نظر تیم پزشکی براساس راهنمای‌های بیماران، استفاده شد. در جدول ۱ نمونه‌ایی از ویژگی‌ها به همراه نحوه کدگذاری پیشنهادی مشخص شده است.

جدول ۱- اطلاعات و داده‌های بانک اطلاعاتی

ویژگی‌ها	کُد	[کیمینه/ بیشینه]
جنسیت	۱+ یا ۱-	[۱+ ۱-]
سن	مقدار مربوطه	[۰ ۱۲۰]
تاهل	۰ یا ۱	[۱+ ۱]
وزن	مقدار مربوطه	[۰ ۲۰۰]
قد	مقدار مربوطه	[۱۴۰ ۲۰۰]
تاریخچه بیماری دیابت	۰ یا ۱ ۲ یا ۵	[۰ ۵]
فشار خون دیاستول	مقدار مربوطه	[۱۰ ۱۰۰]
فشار خون سیستول	مقدار مربوطه	[۴۰ ۱۸۰]
قند خون سریع	مقدار مربوطه	[۵۰ ۵۰۰]
قند خون	مقدار مربوطه	[۵۰ ۵۰۰]
تری‌گلیسرید	مقدار مربوطه	[۵۰ ۵۰۰]
کلسترول	مقدار مربوطه	[۱۰۰ ۴۰۰]
شاخص جرم توده	مقدار مربوطه	[۱۴ ۵۰]
نبض	مقدار مربوطه	[۲۰ ۲۰۰]
نوع تشخیص	۱ یا ۲ یا ۳	[۱ ۳]
نوع درمان	۱ یا ۲ یا ۳	[۱ ۳]
سابقه بیماری	۱ یا ۲ یا ۱۱	[۱ ۱۱]
سابقه دارو	۱+ یا ۱-	[۱+ ۱]

ادامه جدول ۱ در صفحه بعد

ادامه جدول ۱

کلسترول خوب	مقدار مربوطه	[۱۰ ٤۰۰]
کلسترول بد	مقدار مربوطه	[۱۰ ۲۰۰]
اسید اوریک	مقدار مربوطه	[۰ ۱۰۰]
کراتین	مقدار مربوطه	[۰ ۱۰]
هموگلوبین A1c	مقدار مربوطه	[۵ ۲۰]

ب: اهمیت انتخاب ویژگی و بازشناسی الگو

ابتدا باید به مساله انتخاب ویژگی‌های مناسب در فضای تبدیل پرداخت. انتخاب ویژگی در واقع برگزیدن ویژگی‌هایی است که حداکثر توان را در پیشگویی خروجی دارا باشند [۸]. تعریف اینکه زیر مجموعهٔ بهینه چه می‌تواند باشد، به مسئله‌هایی که قرار است حل شوند وابسته است [۹]. در روش‌هایی که بحث شده‌اند مقادیر ویژه بردارها اهمیت آنها را نشان می‌دهد. اهیمت دادن به تمام ویژگی‌ها بر این اساس همیشه مناسب نیست و علاوه بر این بردارهای ویژگی نیز همیشه شامل اطلاعات مفید نیستند. انتظار می‌رود با داشتن اطلاعات بیشتر، نتایج بهتر شود؛ اما در عمل مشاهده می‌شود که این مساله باعث افت دقت می‌شود [۱۰،۱۱]. بنابراین برای رسیدن به یک بازنمایی مناسب، انتخاب بهترین بردار ویژگی یا ساده سازی بردار ویژگی امری ضروری است. یک سیستم شناسایی الگوی متداول شامل ٤ بخش استخراج ویژگی، انتخاب ویژگی، طراحی و آموزش طبقه‌بندی کننده و سرانجام آزمایش است. در این تحقیق از الگوریتم ژنتیک دودویی[1] (GA) برای کاوش و انتخاب ویژگی‌های موثر، الگوریتم ماشین بردار پشتیبان (SVM)[2] برای طبقه‌بندی نمونه‌ها و همین طور از شبکه‌های عصبی مصنوعی برای طراحی طبقه‌بندی کننده و آزمایش نهایی استفاده شده است. الگوریتم‌های انتخاب ویژگی بسته به روند ارزیابی آنها به دو دستهٔ عمده تقسیم می‌شوند. اگر انتخاب ویژگی مستقل از هرگونه الگوریتم یادگیری انجام شود (یعنی به صورت یک پیش پردازندهٔ کاملاً مجزا)، آن روش، فیلتر یا حلقهٔ باز می‌باشد. در این مورد ویژگی‌های نامطلوب قبل از استنتاج دور ریخته می‌شوند. اما، اگر روند ارزیابی با یک الگوریتم طبقه‌بندی در ارتباط باشد، روش انتخاب ویژگی را Wrapper یا حلقهٔ بسته می‌گویند. این روش، جستجو در فضای زیر مجموعه‌ها را براساس تخمین دقت ناشی از انتخاب یک زیر مجموعهٔ خاص تحت شرایط الگوریتم طبقه‌بندی مورد استفاده (به عنوان معیاری از بهینگی آن زیر مجموعه) انجام می‌شود [۱۲]. الگوریتم‌های دستهٔ دوم معمولاً نتایج بهتری به دست می‌آورند. مهمترین بخش هر روش انتخاب ویژگی حلقهٔ بسته، الگوریتم جستجویی است، که در آن به کار رفته است. طی دههٔ گذشته محققان روی الگوریتم‌های جستجوی تکاملی مثل الگوریتم ژنتیک [۱۳،۱٤]، ACO[3] [۱۲،۱۵،۱٦] و ترکیب ACO/PSO[4] [۱۷] تمرکز داشته‌اند.

ج: فرآیند اجرای مراحل در روش پیشنهادی

در الگوریتم پیشنهادی ابتدا داده‌ها در بانک اطلاعاتی تهیه شده قرار گرفتند و سپس توسط رابطه (۱) نرمالیزه شدند.

$$yi = \frac{xi - m}{\sigma i} \quad (۱)$$

در این رابطه xi اعداد و یا داده‌ها؛ m میانگین داده‌های ویژگی iام و واریانس داده‌های ویژگی iام (در هر ویژگی) هستند. توسط الگوریتم ژنتیک باینری موثرترین ویژگی‌هایی که در الگوریتم طبقه‌بندی کننده ماشین بردار پشتیبان (SVM) باعث بهترین و بیشترین طبقه‌بندی می‌شوند، انتخاب شدند و در نهایت براساس بهترین ویژگی‌هایی که انتخاب شدند، عمل مدل‌سازی و انطباق فضای ورودی (ویژگی‌های بیماران) با فضای خروجی (تصمیمات پزشک مبنی بر تجویز دارو) توسط شبکه‌های عصبی مصنوعی صورت گرفت.

د: الگوریتم ژنتیکی دودیی

در این تحقیق از یک نوع الگوریتم ژنتیک ساده توام با الگوریتم طبقه‌بندی ماشین بردار پشتیبان (SVM) برای انتخاب پارامترهای (ویژگی‌های) بهینه استفاده شد. فلوچارت ۲، نمودار بلوکی سیستم مورد استفاده را نمایش می‌دهد؛ البته جایگاه قرارگیری مجموعه پیشنهادی در تابع هزینه الگوریتم ژنتیک می‌باشد. ابتدا یک بردار دودویی به

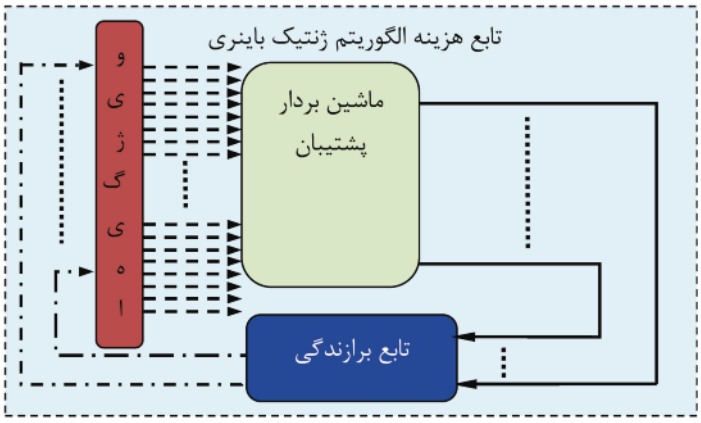
1. Binary Genetic Algorithm
2. Support Vector Machine
3. Ant Colony Optimization
4. Ant Colony Optimization/Particle Swarm Optimization

طول ۶۹ بیت (هر بیمار ۳ بار داده گیری ۲۳ نمونه‌ایی) واحد (به تعداد پارامترها یا ویژگی‌ها) بصورت تصادفی انتخاب می‌شود که هر کدام از بیت‌های این رشته دودویی متناظر با یک پارامتر در ماتریس آموزش شبکه است. اگر بیت متناظر با هر پارامتر صفر باشد، آن پارامتر حذف شده و اگر یک باشد، آن پارامتر حذف نشده و در آموزش دخالت داده می‌شود. پس از اجرای یک نسل از الگوریتم ژنتیک با پارامترهای تصادفی مقدار برازندگی هر رشته دودویی (کروموزوم) برای ۳۰ نمونه تست محاسبه شد و از بهترین کروموزوم‌های نسل قبل، ۲ کروموزوم به همراه ۱۸ کروموزوم جدید که با استفاده از عملگرهای ژنتیک تولید شدند، نسل جدید (دوم) الگوریتم ژنتیک تولید می‌شود. این چرخه اینقدر ادامه پیدا می‌کند، تا به دقت مورد تنظیم دست یابیم و یا تعداد نسل‌ها به تعداد حداکثر تنظیم شده رسیده باشد. در اجرای الگوریتم ژنتیک حداکثر تعداد نسل‌ها را ۵۰۰ نسل تنظیم شدند. سایر پارامترهای الگوریتم ژنتیک نیز مطابق جدول ۲، براساس اقتباس از [۲۰-۱۸] هستند.

جدول ۲- پارامترهای الگوریتم ژنتیک پیشنهادی

تعداد نسل‌ها	۵۰۰
اندازه جمعیت	۲۰
احتمال تقاطع	۰,۸
احتمال جهش	۰,۱
نوع باز ترکیب	دو نقطه‌ایی
نوع انتخاب	چرخ رولت

در این تحقیق به عنوان تابع برازندگی الگوریتم ژنتیک، از حاصل ضرب حساسیت طبقه‌بندی در ویژگی طبقه‌بندی ماشین بردار پشتیبان (SVM) استفاده شدند، و متغیرهای ورودی این شبکه یک رشته دودویی با طول ۶۹ بیت (واحد) می‌باشد که، هریک از بیت‌های این رشته معرف انتخاب یا عدم انتخاب یک پارامتر تشخیصی می‌باشد.



فلوچارت ۲- تابع هزینه الگوریتم ژنتیک

علت استفاده از تابع هزینه (رابطه ۱) در الگوریتم ژنتیک این است که در این حالت مقدار حساسیت (رابطه ۲) و ویژگی (رابطه ۳) را همزمان افزایش می‌یابند. وقتی مقدار حساسیت و ویژگی تشخیصی هر دو افزایش یابند، مقدار دقت و سطح زیر منحنی ROC[1] (منحنی جهت مقایسه عملکرد و مقایسه ۲ سیستم) نیز افزایش می‌یابند.

$$Fitness = Se \times Sp \qquad (۱)$$

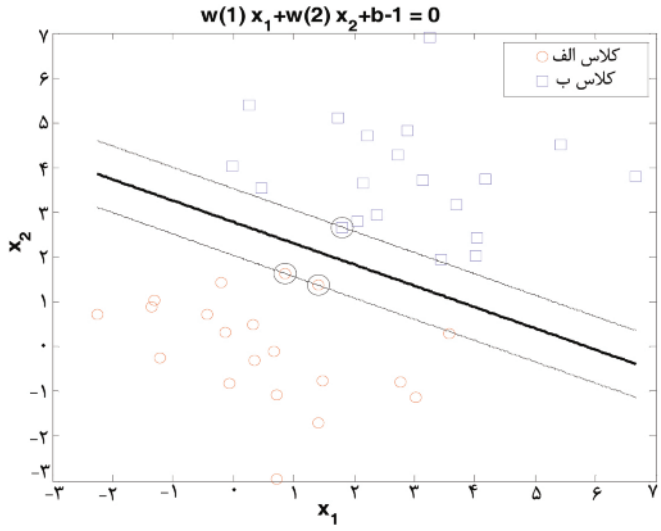
$$Se = \frac{Tp}{Tp + Fn} \qquad (۲)$$

$$Sp = \frac{Tn}{Tn + Fp} \qquad (۳)$$

1. Receiver Operating Characteristics

هـ- ماشین بردار پشتیبان

SVM یک الگوریتم یادگیری با سرپرستی است که می‌تواند در کاربردهایی نظیر جداسازی و طبقه‌بندی داده مورد استفاده قرار گیرد [۲۱]. این الگوریتم می‌تواند داده‌های آموزش را به دو دسته تقسیم کند. اصولاً داده‌ها توسط یک خط قابل جداسازی هستند، و در غیر این صوت الگوریتم با نگاشت ویژگی‌ها به وسیله تابع φ به ابعاد بالاتر سعی می‌کند توسط یک فرا صفحه با دقت بالاتری داده‌ها را تفکیک نماید (شکل ۱).



شکل ۱- جداسازی داده‌ها در ۲ کلاس توسط الگوریتم SVM

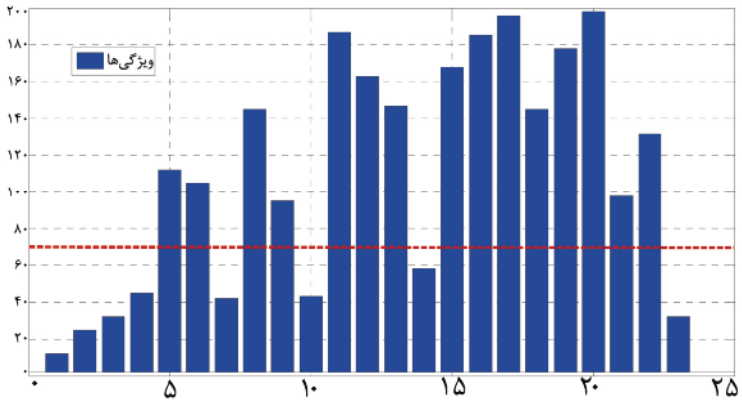
معادله (٤) رابطه فراصفحه طبقه‌بندی کننده داده‌ها است. معادلات (٥ و ٦) روابط فراصفحه‌های موازی براساس شرایط حداکثر حاشیه می‌باشند. در صورت ترسیم این توابع ساده فاصله بین دو صفحه حاشیه برابر می‌باشد.

$$w \times x - b = 0 \qquad (٤)$$

$$w \times x - b = 1 \qquad (٥)$$

$$w \times x - b = -1 \qquad (٦)$$

در این روابط x متغییر ورودی، w بردار نرمال خط جدا کننده و b عرض از مبدا خط جدا کننده است. SVM در اینجا نقش کلاس‌بندی نمونه‌های خروجی از الگوریتم ژنتیک را دارد، هدف مشخص کردن ویژگی‌هایی است که بیشترین تاثیر در تمایز بین ویژگی‌ها به منظور پیش‌بینی را داشته باشند. به منظور تسریع و کاهش محاسبات در این کار از تابع خطی SVM به منظور جداسازی ویژگی‌های موثر از غیر موثر استفاده شده است. تا این مرحله بهترین نتایج یعنی موثرترین ویژگی‌هایی که در امر پیش‌بینی نقش دارند مشخص می‌شوند، یعنی ویژگی‌های شماره (٥،٦،٨،١١،١٢،١٥،١٦،١٧،١٨،١٩،٢٠ و ٢٢) یا همان ویژگی‌هایی که در (تصویر ۲) مشخص شده است.



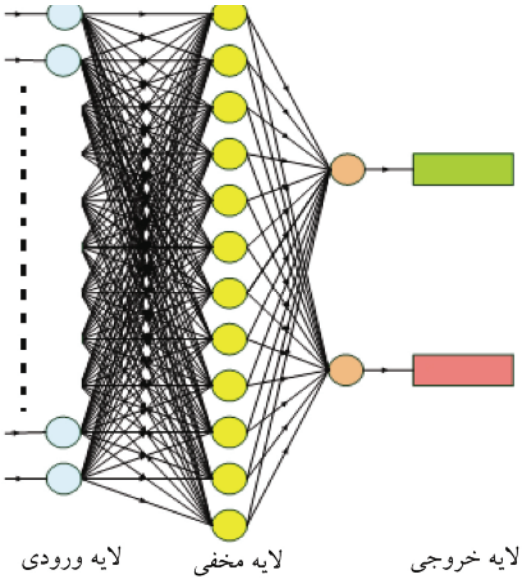
شکل ۲- بهترین ویژگی‌های انتخاب شده توسط الگوریتم ژنتیک و طبقه‌بندی کننده SVM

در اثر فقط ۲۰۰ تکرار، تمایزی فاحش بین ویژگی‌های مهم و سرنوشت‌ساز در تعیین میزان دُز دارو برای بیماران مشخص شد. در مرحله بعدی باید تنها ویژگی‌های مشخص شده از افراد را براساس اصل سری‌های زمانی (در

ذیل توضیح داده می‌شود)، به منظور انطباق، مدل‌سازی و تصمیم گیری با توجه به مدل و داده‌های ورودی به سیستم یا شبکه عصبی (مشخصات هر فرد)، تحویل شود تا تصمیمی صحیح در مورد میزان داروی تجویزی برای هر فرد بدست آید. دلیل استفاده از سری زمانی بیماران این است که از هر بیمار در ۳ نوبت با فواصل یک هفته نمونه گیری شد. در این تحقیق بدست آوردن بهترین عملکرد و یا خلاصه‌ایی از گذشته سیستم فاکتور مهمی است که توسط طبقه‌بندی کننده الگوریتم ژنتیک و الگوریتم SVM بهترین کارایی یا ویژگی سیستم در گذشته بدست می‌آید تا بوسیله شبکه‌های عصبی مصنوعی، رفتار آینده سیستم پیش بینی شود (با اقتباس از [۲۲]).

و: شبکه‌های عصبی مصنوعی

در این پروژه از یک شبکه عصبی ۳ لایه پیش خور با تابع فعالیت سیگموئید در لایه میانی و لایه خروجی و تابع خطی برای لایه ورودی استفاده گردید. تعداد نرون‌های لایه ورودی ۱۲ نرون در نظر گرفته شد. لایه میانی این شبکه برای انتخاب بهترین ساختار، از ۳ تا ۱۷ نرون یعنی به صورت (۳،۵،۷،۹،۱۱،۱۳،۱۵ و ۱۷) تغییر داده شد، تا با توجه به بهترین نتایج بدست آمده، تعداد نرون‌های لایه میانی انتخاب شود که در نهایت بهترین نتایج با ۹ نرون در لایه میانی بدست آمد. انتخاب تعداد نرون‌های لایه میانی بسیار مهم است چرا که اگر تعداد آنها کم باشد، شبکه برای حل مسائل غیر خطی و پیچیده با کمبود منابع یادگیری مواجه می‌شود و اگر زیاد باشد، خود باعث ایجاد دو مشکل می‌شود. اول آنکه زمان آموزش شبکه افزایش می‌یابد و دوم آنکه ممکن است شبکه نظام بی‌اهمیت داده‌های آموزش را یاد بگیرد و در حل مسائل ضعیف عمل نماید. در این پروژه، برای آموزش شبکه از الگوریتم پس انتشار خطا استفاده شده است. در این روش ماتریس‌های وزنی شبکه به گونه‌ایی تغییر می‌یابند، تا مقدار متوسط مربعات خطای شبکه (MSE[1] یا میانگین مربعات خطا) از مقدار خاصی که در برنامه مشخص می‌شود، (۰/۰۲) کمتر شود. خروجی مطلوب شبکه در واقع همان نتایجی است که توسط پزشک تجویز شده است. در این روش خطا به روش خاص، لایه به لایه، به عقب برگشت داده می‌شود و در هر لایه اصلاحات لازم روی وزن‌ها انجام می‌گیرد. این روند آنقدر ادامه می‌یابد تا خطای خروجی به سمت مقدار کمینه‌ای به نام Error Goal (خطای هدف) همگرا شود و با تعداد دفعات تکرار حداکثر از قبل تعیین شده برابر شود. این مقادیر باید طوری تعیین شوند که از یادگیری بیش از اندازه شبکه جلوگیری شود. تعداد دفعات تکرار در این تحقیق ۱۰۰۰ بار در نظر گرفته شده است و مقدار Error Goal نیز برابر ۰/۰۲ تعیین شده است. مدل کامل یک نوع شبکه عصبی در شکل ۳ آمده است.



شکل ۳- شمای نمونه از مدل کامل شبکه عصبی

برای بهینه سازی ضریب یادگیری، از تعداد کل پارامترهای ورودی به عنوان تعداد لایه ورودی، از ۵ نرون به عنوان تعداد نرون‌های لایه میانی و یک نرون نیز به عنوان لایه خروجی استفاده شد. با اقتباس از [۲۳] ما این تعداد نرون‌ها را به همراه، پارامترهای دیگر شبکه از قبیل دفعات تکرار و Error Goal (خطای هدف) را ثابت نگه داشتیم و ضریب یادگیری را از ۰/۰۰۱ تا ۱ تغییر دادیم و هر بار شبکه را آموزش دادیم و مقدار MSE (میانگین مربعات خطا) را در هر مرحله مشخص نمودیم. به این ترتیب نرخ یادگیری که منجر به حداقل MSE (میانگین مربعات خطا)

1. Mean Square Error

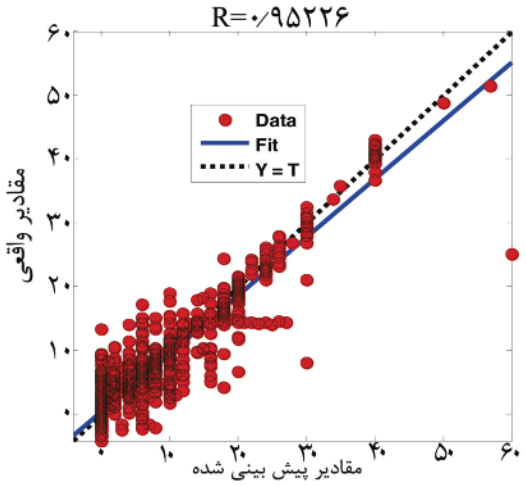
می‌شود، حدود ۰/۲ بدست آمد. مشخصات شبکه عصبی به کار رفته در این پروژه در جدل ۳ آمده است.

جدول ۳- پارامترهای شبکه عصبی پیشنهادی

مشخصات و پارامترهای شبکه عصبی	
تعداد نرون‌های لایه ورودی	۱۲ عدد
تعداد لایه‌های پنهان	۱ عدد
تعداد نرون‌های لایه پنهان	۵ عدد
تعداد نرون‌های لایه خروجی	۱ عدد
نوع تابه کار شبکه	سیگموئید
نوع آموزش شبکه	پس انتشار خطا
میانگین مربعات خطا	۰/۰۲
نرخ یادگیری	۰/۰۲
تعداد تکرار	۱۰۰۰

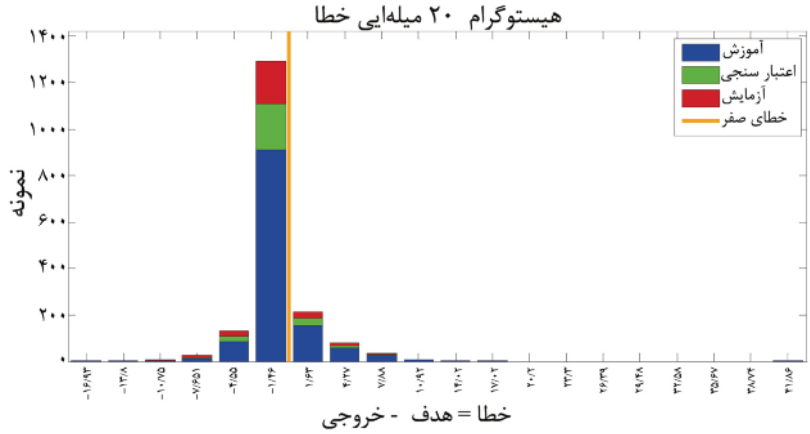
ز: صحت آزمایش و ارزیابی نهایی

در حالت کلی می‌توان تمام فرآیند انتخاب ویژگی که منجر به انتخاب ویژگی به منظور پیش‌بینی دُز صحیح دارو شد را به این صورت بیان کرد، ابتدا بانک داده تشکیل شد، سپس توسط الگوریتم باینری ژنتیک با توجه به تعداد ویژگی‌های بانک داده، بردارهایی به ازای هر نمونه سرشار از "۰" و"۱" تولید گردید، رشته در مشخصات هر نمونه (گزارش) ضرب شده و با توجه به" ۰" یا " ۱"شدن ویژگی‌های هر نمونه، آرایه حاصل وارد تابع هزینه (در تابع برازندگی) شدند، سپس خروجی تشخیصی هر نمونه بدست آمد و متناسب با خروجی‌های تشخیصی، به منظور طبقه‌بندی صحیح، نتیجه وارد الگوریتم SVM گردید و مجدداً به الگوریتم GA باز گردید. این فرآیند در تابع هزینه الگوریتم GA و تکرار الگوریتم GA منجر به انتخاب بهترین ویژگی‌های ممکن به منظور پیش‌بینی شد. براساس بهترین ویژگی‌ها، بانک داده‌ایی شامل فقط ویژگی‌های مهم وارد شبکه عصبی می‌شود. شبکه عصبی براساس مدل مشخص شده که متناسب با نظرات پزشک مبتنی بر بانک داده است، تصمیم‌گیری بر طبق پیش‌بینی را انجام می‌دهد. برای اثبات قابلیت‌های سیستم پیشنهادی، فرایند اعتبار سنجی عرضی ۱۰- دسته‌ایی (توسط K-fold) بر روی پایگاه داده (مجموعاً ۳٦۸ نمونه حاوی ویژگی‌های نمونه) پیاده شد که به ۱۰ دسته تقسیم شدند. به طوری که هر قسمت دارای ترکیبی از تمام نمونه‌ها می‌باشد. ۸ دسته ۳٦ داده‌ایی و ۲ دسته ۳۷ داده‌ایی. در این نوع اعتبار سنجی، ۹ قسمت از داده‌ها برای آموزش و قسمت باقیمانده برای آزمایش به کار رفتند. این فرایند آموزش و آزمایش ۱۰ بار به صورت متوالی چرخشی انجام شد به طوری که هر بار یک قسمت متفاوت برای آزمایش کنار گذاشته شد، تا بهترین ساختار شبکه مشخص شود، پس از مشخص شدن بهترین ساختار شبکه، می‌توان تمام آزمایشات را با این ساختار از شبکه انجام داد. بهترین نتیجه بدست آمده براساس نتایج واقعی در حالت رگرسیونی در شکل ٤ ترسیم شد، که مقدار انحراف و دقت بدست آمده قابل قبول می‌باشد (در حد بیش از ۰/۹۵ درصد).



شکل ۴- نمایش رگرسیونی مقادیر پیش‌بینی شده بر حسب نتایج واقعی

نمودار رگرسیونی در شکل ٤ نشان می‌دهد که نتایج بدست آمده در مقایسه با مقادیر واقعی در حالت کلی از دقت بالایی برخوردار است. به طوری که می‌توان گفت تقریبا مقادیر پیش‌بینی شده، بیش از ۰/۹۵ مقادیر واقعی هست. در عملکرد نهایی سیستم پیشنهادی، میزان خطایی در مقایسه بین مقادیر واقعی و مقادیر پیش‌بینی شده بوجود آمد؛ که برای نمایش این خطا یا اصطلاحاً، نمودار خطای هیستوگرام شبکه از شکل ۵ استفاده گردید.



شکل ۵- نمایش عملکرد خطای شبکه در محاسبه مقادیر واقعی و مقادیر پیش‌بینی شده

خطای سیستم که توسط تابعی در ساختار انطباقی شبکه عصبی تعریف شده است، (اختلاف مقادیر تابع هدف و تابع پیش‌بینی شده) توسط شکل ۵ بیان گردید. البته هر چقدر نمودار شکل ۵ به حالت گوسی نزدیک‌تر باشد مطمئناً شبکه عملکرد بهتری داشته است. در این شکل نمودار داری ٤ قسمت است، در محمور عمودی مقدار یا تعداد نمونه‌ها، مشخص شده است، در محور افقی، مقدارهای عددی خطا (حول نقطه صفر یا خط عمودی نارنجی رنگ)، مشخص شده است، نمودارهای میله‌ای سه رنگ متفاوت دارند، اولین رنگ قرمز است، که بیان کننده خطا در نمونه‌های تست است، رنگ سبز بیانگر مقدار خطا در نمونه‌های، اعتبار سنجی است و همین طور قسمت آبی رنگ که بیانگر مقدار خطا در نمونه‌های آموزشی هستند. تمام مقادیر خطا که در نمودار بالا ترسیم شده است، براساس نتایج نهایی و کلی است که در مقدار خطای "۱/٤٦۱-" شبکه به اشباع رسید.

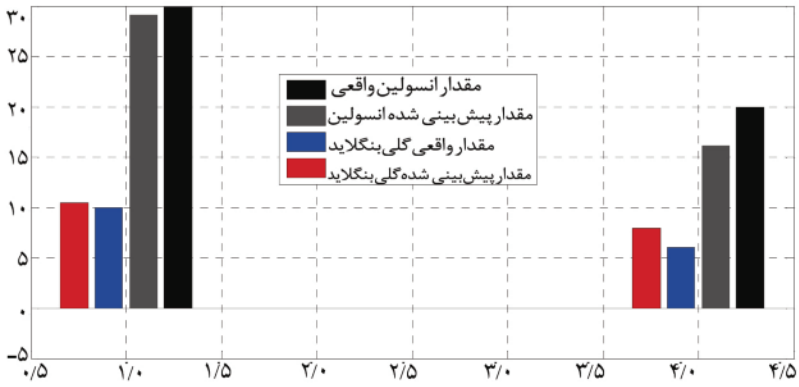
یافته‌ها

نتایج شبیه‌سازی‌ها و ارزیابی‌های انجام شده

به منظور شبیه‌سازی، از نرم‌افزار MALAB® و همین طور برای ترسیم نمودارها از نرم‌افزار ORIGIN® استفاده شده است. نتایج شبیه‌سازی‌ها نشان می‌دهد که انتخاب صحیح ورودی‌ها، به شبکه‌های عصبی بسیار مهم است، چرا که ورودی‌ها می‌توانند دقت و کارایی شبکه را افزایش و زمان همگرایی را به طور چشمگیری کاهش دهد. ویژگی‌های اصلی، سرنوشت‌ساز و تاثیرگذار در تعیین میزان دُز انسولین و دارو در بیماران تعیین شدند. سپس شبیه‌سازی نهایی توسط شبکه عصبی مصنوعی مطابق شکل ٦ برای ۲٤ ساعت آینده بیمار تهیه شد. البته مقادیر پس از این که توسط شبکه عصبی، پیش‌بینی شدند، باید به صورت قبلی بازگردند، به همین خاطر آنها طبق رابطه (۱۲) دنرمالایز گردیدند، سپس مقادیر واقعی مشخص شدند.

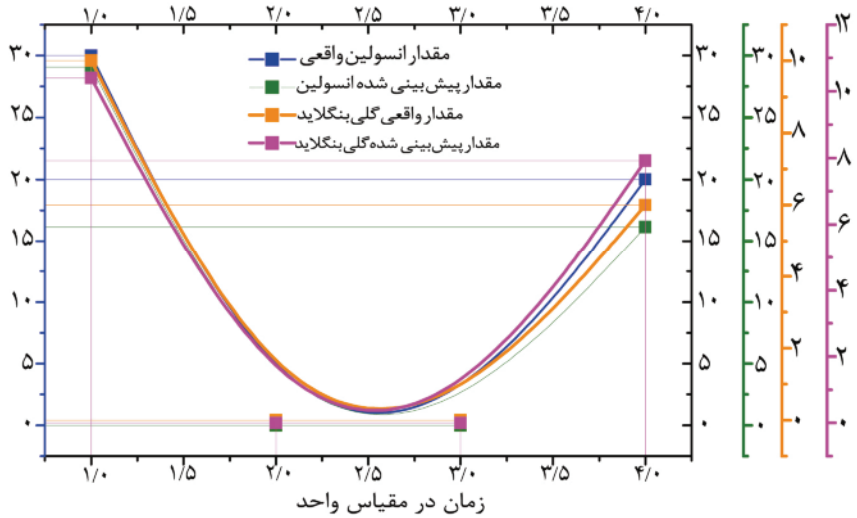
$$x_i = y_i \times \sigma_i + m \quad (۱۲)$$

در این رابطه xi مقدار واقعیِ پیش‌بینی شده است، yi مقدار پیش‌بینی شده توسط شبکه است، m مقدار میانگین داده‌ها و مقدار انحراف از معیار داده‌ها (منظور در هر ویژگی) است. در شکل ٦ میزان داروی، پیش‌بینی شده برای یک فرد که به صورت تصادفی انتخاب شده است (فرد شماره ۹)، مشخص شده است.



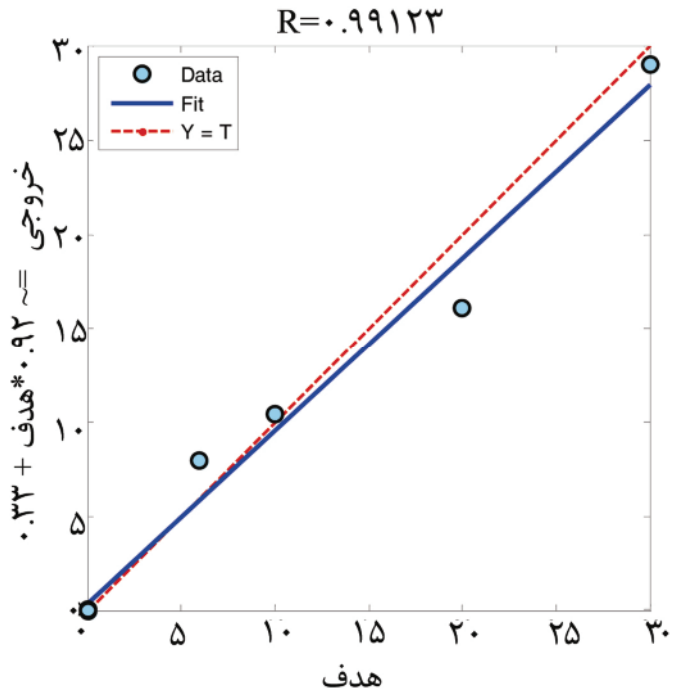
شکل ۶- مقادیر پیش‌بینی شده و واقعی برای انسولین و گلیبن‌کلامید

در شکل ٦ میزان تجویز دارو برای بیمار شماره ۹ مشخص شد. محور افقی این نمودار بیان کننده زمان واحد است، هر واحد ٦ ساعت آینده را بیان می‌کند که در مجموع ٢٤ ساعت آینده را بیان می‌شود، شماره ۱ بیان کننده ساعات (۰ بامداد تا ٦ صبح)، شماره ۲ بیان کننده ساعات (٦ صبح تا ۱۲ ظهر)، شماره ۳ بیان کننده ساعات (۱۲ظهر تا ۱۸ عصر) و شماره ٤ بیان کننده ساعات (۱۸ عصر تا ٢٤ نیمه شب) می‌باشد، برای این بیمار مقادیر واقعی که مقرر شده بودند به این صورت است که، ۳۰ واحد انسولین، ۱۰ واحد گلیبن‌کلامید در ساعات اولیه (۰ بامداد تا ٦ صبح)، ۲۰ واحد انسولین و ٦ واحد گلیبن‌کلامید در ساعت (۱۸ عصر تا ٢٤ نیمه شب)، در صورتی که مقادیر پیش‌بینی شده برای این بیمار ۲۹ واحد انسولین و ۱۰ واحد گلیبن‌کلامید در ساعات اولیه (۰ بامداد تا ٦ صبح) و مقادیر ۱٦ واحد انسولین و ۷ واحد گلیبن‌کلامید در ساعات (۱۸ عصر تا ٢٤ نیمه شب) پیش‌بینی شده است. در شکل ۷ مقادیر پیش‌بینی شده بر حسب مقدار و زمان (بر حسب واحد) برای بیمار شماره ۹ ترسیم شده است.



شکل ۷- مقادیر عددی پیش‌بینی شده و واقعی انسولین و گلیبن‌کلامید

اگر مقادیر پیش‌بینی شده را بر حسب مقادیر واقعی به وسیله نمایش رگرسیون بیان شود، شکل ۸ حاصل می‌شود که انحراف و ضریب R آن (بیش ۰/۹۹) بیان کننده مناسب و دقیق بودن پیش‌بینی صورت گرفته است.



شکل ۸- نمایش مقادیر پیش‌بینی شده و مقادیر واقعی شده انسولین و گلیبن‌کلامید برای بیمار شماره ۹

شکل ۸ بیان کننده این اصل است که برای فرد بیمار شماره ۹، صحت مقادیر پیش‌بینی شده ۰/۹۹ صحت مقادیر واقعی تجویز شده از سوی پزشک برای آن بیمار می‌باشد. تا این قسمت از مباحثی که در مورد پیش‌بینی و طراحی سیستم پیش‌بینی بحث شد، عملکرد سیستم در یک فرد تصادفی مورد بررسی قرار گرفت. که البته مسلماً با یک نفر نمی‌توان در مورد یک مجموعه صحبت، تصمیم و یا حتی اعتبار آن را بررسی نمود. به همین خاطر از نمودار ROC، پارامترهای حساسیت، ویژگی و دقت استفاده گردید.

ارزیابی مدل پیشنهادی

برای ارزیابی سیستم پیشنهادی، ۵۰ نمونه از بانک اطلاعاتی به طور تصادفی انتخاب شدند که شامل نمونه‌هایی با دُز داروهای تجویز شده در دو نوع انسولین و داروی گلیبن‌کلامید می‌باشند. معیارهای ما برای ارزیابی سیستم پیشنهادی، میزان دقت، حساسیت و ویژگی‌های تعیین شده و همچنین سطح زیر منحنی ROC[9] برای این ۵۰ نمونه می‌باشد. نحوه محاسبه این پارامترها طبق جدول ٤ می‌باشد. از آنجایی که حساسیت، ویژگی و دقت، تا حد زیادی وابسته به جمعیت بیماران غالب بوده و از این رو ممکن است نتوان به درستی و به صورت دقیق کارایی شبکه و تصمیمات پزشک را با یکدیگر مقایسه کرد، لذا از تحلیل منحنی‌های ROC برای ارزیابی کارایی مدل در مقایسه با آنچه که از پزشک (متخصص غدد) بدست آمده است، استفاده شد.

جدول ۴- پارمترهای مورد نیاز در مبحث مقایسه

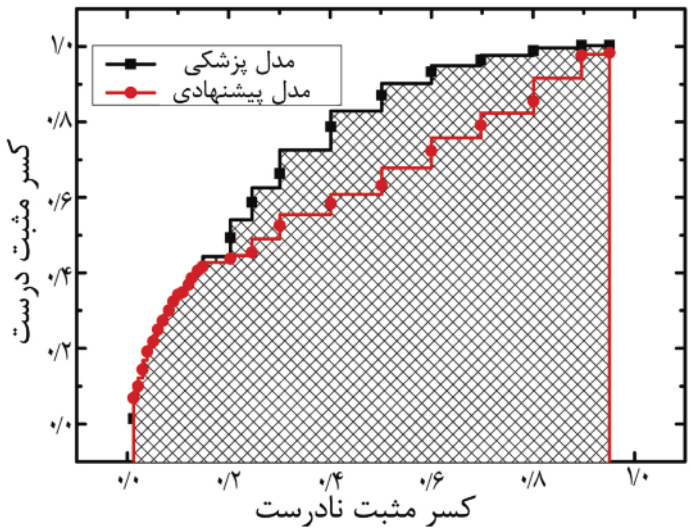
مقادیر پیش‌بینی شده	مقادیر واقعی
Fp (مثبت نادرست)	Tn (منفی درست)
Tp (مثبت درست)	Fn (منفی نادرست)
$Spe=\frac{Tn}{Tn+Fp}$ کسر ویژگی	$Sen=\frac{Tp}{Tp+Fn}$ کسر حساسیت
$Accuracy=\frac{Tp+Tn}{Tp+Fn+Tn+Fp}$ کسر دقت پیش‌بینی شده منفی	$Positive\ Predictive\ Value=TFP=\frac{Tp}{Tp+Fp}$ کسر مقادیر پیش‌بینی شده مثبت
	$Negetive\ \ Predictive\ Value\ =FPF=\frac{Tn}{Tn+Fn}$ کسر مقادیر پیش‌بینی شده منفی

براساس رابطه‌های جدول ٤، مقادیر دقت، حساسیت و ویژگی را برای داده‌های نهایی یعنی مقادیر پیش‌بینی شده بیماران را در ۲ قسمت دُز تجویز شده انسولین و دُز گلیبن‌کلامید، در جدول ۵ بررسی و مشخص شدند.

جدول ۵- مقادیر دقت نهایی، حساسیت و ویژگی پیش‌بینی شده برای ۲ مقدار انسولین و گلیبن‌کلامید

دقت الگوریتم پیشنهادی براساس درصد	گلیبن‌کلامید پیش‌بینی شده	دقت الگوریتم پیش‌بینی شده براساس درصد	انسولین پیش‌بینی شده
۸۳	حساسیت	۹۲/۷	حساسیت
٦١	ویژگی	۸٤/۱	ویژگی
۹۲/٤٥	دقت	۹٥/٥۲	دقت

پس از تحلیل کامل و بررسی چند گانه در مورد دقت و قدرت پیشگویی سیستم پیشنهادی؛ به مناسب بودن این سیستم و اطمینان این سیستم می‌توان پی برد. در مرحله بعد، به تحلیل نمودار ROC در سیستم پیشنهادی پرداخته می‌شود. تحلیل ROC، شکلی از یک اندازه‌گیری است که عمدتاً برای مقایسه کارایی روش‌های تشخیصی بکار می‌رود. در تحلیل ROC، کارائی تشخیص و یا تصمیم‌گیری بر حسب دو شاخص ارزیابی می‌شود: ۱-کسر مثبت درست (TPF)، ۲- کسر مثبت نادرست (FPF). کسر اول به عنوان کسری از تصمیم‌گیری‌هایی است که به درستی تشخیص داده شده‌اند (مثل حساسیت) و کسر دوم به کسری از تصمیم‌گیری‌هایی است که به درستی تشخیص داده نشده‌اند، از مقادیر فوق می‌توان برای رسم منحنی‌های ROC استفاده نمود. سطح زیر منحنی (ROC) معیاری برای میزان کارآیی روش بکار رفته است. همان طور که قبلاً گفته شد، هر چقدر که سطح زیر این نمودار بیشتر باشد نشان دهنده عملکرد بهتر آن سیستم خواهد بود، البته اگر در حالت مقایسه با سیستمی مثل تصمیمات پزشک باشد، سیستم پیشنهادی زمانی در حد سیستم واقعی یا اصلی عمل می‌کند، که سطح زیر منحنی برابر سطح زیر منحنی تصمیمات پزشک باشد. در اثر تعریف پارامترها و مشخص شدن بازه‌ها، نمودار ROC برای بیماران براساس نتایج سیستم پیشنهادی و تصمیمات پزشک ارائه شده ترسیم شدند.

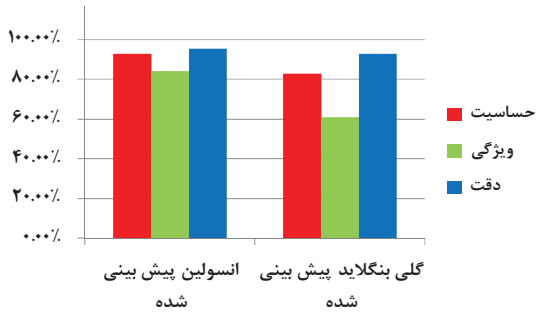


شکل ۹- نمودار ROC در سیستم پیشنهادی و سیستم بتنی بر نظریات پزشک

مساحت‌های سطح زیر ۲ نمودار در سیستم پیشنهادی و سیستمی که مبتنی بر نظریات و تصمیمات پزشک است، در شکل ۹ مشخص شده است، به راحتی می‌توان در مورد عملکرد مناسب سیستم پیشنهادی به خاطر برابر بودن نسبی مساحت‌ها و محدوده عملکرد دو سیستم با توجه به خطوط قرمز و مشکی در شکل ۹ پی برد.

بحث

دیابت از شایع‌ترین عامل‌های مرگ و میر در سراسر دنیا، در بین تقریباً تمامی اقوام جهان می‌باشد، با توجه به پیشرفت‌های بدست آمده در روش‌های تشخیص، اما هنوز هم تعیین میزان صحیح دُز داروی این بیماران خصوصاً انسولین و گلیبن‌کلامید، در مراحل اولیه کار دشوار، مهم و سرنوشت‌ساز است و در بسیاری از موارد تعیین دُز دارو، با قطعیت و صحت کامل انجام نمی‌شود. در همین راستا تحقیقات جدیدی انجام گرفته است؛ از آن جمله می‌توان به مدل‌های ریاضی و مینی‌مال براساس ویژگی‌های حساسیت بدن بیمار در پاسخ به گلوگز تزریقی، دُز انسولین مصرفی و سطح فعلی قند خون که عمدتاً به منظور تعیین حساسیت بدن بیمار در پاسخ به گلوکز تزریقی و نیز پیش‌بینی کوتاه مدت سطح قند و انسولین موجود در بدن، مورد استفاده قرار گرفته‌اند، اشاره کرد [۲۶-۲۴]. Huang به همراه همکاران، با استفاده از شبکه‌های عصبی در سال ۲۰۱۰ پیش‌بینی سطح گلوکز خون بیماران را براساس ویژگی‌ها و اطلاعات هر بیمار شامل؛ وزن، زمان اولیه تزریق، زمان انتهایی تزریق و کربوهیدارت مصرفی برای بیماران را پیش‌بینی کرده‌اند [۲۷]. Yu در سال ۲۰۱۲ توسط روش‌های محاسباتی و برنامه‌ریزی LSSVM1 پیش‌بینی دُز انسولین را انجام داده است [۲۸]. البته باید گفت که از گذشته تحقیق و تامل بر روی موضوع پیش‌بینی دُز انسولین برای بیماران رایج بوده است. Canete و همکاران، یک سیستم حلقه بسته عملی آزمایشگاهی به صورت آنلاین برای کنترل گلوکز خون براساس میانگین وزن؛ گلوکز پلاسما و انسولین پیشنهادی تهیه کرده است [۲۹]. Pappada و همکاران در سال ۲۰۰۸ تنها برای ۵۰ -۱۸۰ دقیقه توانستند براساس ویژگی‌های؛ نحوه زندگی، وضعیت روحی، دز انسولین، وعده غذایی و کربوهیدارت مصرفی پیش‌بینی سطح انسولین برای بیماران را انجام دهند [۳۰]. Stavroula در سال ۲۰۰۶ با استفاده از ٤ پارامتر عملکرد کوتاه، متوسط، سریع، بلند مدت و کربوهیدرات مصرفی و غلظت قند خون که تابعی از زمان بَعد از تزریق انسولین در بدن هستند (بیماران در اینجا کودک می‌باشند) توسط شبکه‌های عصبی مصنوعی میزان انسولین را برای بیماران مشخص کرده‌اند [۳۱]. Wahab و همکاران با استفاده از کنترل کننده تطبیقی، طراحی یک کنترل کننده PID2 براساس ۲ معادله دیفرانسل مرتبه اول برای انسولین و تنظیم گلوگز در ۲ زیر سیستم به منظور تعیین دُز صحیح انسولین را انجام داده‌اند [۳۲]. لزوم استفاده از روشی صحیح و منطقی مبتنی بر واقعیت در این فرآیند نسبتاً پیچیده و حیاتی لازم است، سیستم‌های هوشمند مصنوعی که به نوعی تلفیقی از دانش فرد خبره یا بشری با مدلی از سیستم مورد بررسی است، می‌توانند به عنوان ابزاری مناسب در جهت تعیین این فرایند مهم استفاده شود. در این تحقیق سعی کردیم مدلی ارائه شود تا با استفاده از سیستم‌های هوشمند مصنوعی و تصمیمات پزشکان از یک طرف و رفتار بیماران از طرف دیگر در قالب بانک داده، فرآیند تعیین صحیح دُز انسولین و دُز گلیبن‌کلامید در بیماران را انجام دهیم. سیستم پیشنهادی همان طور که در شکل ۱۰ می‌بینیم، به دقت (و نتایج) بیش از ۹۵٪ دست یافت که در بررسی‌های به عمل آمده از نتایج حاصل، می‌توان به دقت و صحت این سیستم پی برد.



شکل ۱۰- نتایج نهایی سیستم پیشنهادی

به طوری که نتایج موجود و به دست آمده (در مقالات و تحقیقات ارائه شده قبلی سیستمی مشابه و یا همانند نبود ؛ تا آنجا که مولفین اطلاع دارند) در مقایسه کلی نتایج، دقت و قدرت پیش‌بینی برای بیماران، به میزان قابل توجهی بهبود یافت. همچنین مقایسه نتایج به دست آمده در این تحقیق با نتایج به دست آمده از تحقیقات مشابه قبلی نشان دهنده افزایش قابل ملاحظه میزان صحت پیش‌بینی سری زمانی در سطح غلظت قند خون است، چرا که سیستم پیشنهادی موفق شد با در دست داشتن چند ویژگی از

بیماران مبتلا به دیابت، سطح قند خون تا ۲٤ ساعت آینده بیمار را پیش‌بینی و کنترل نماید. از شبکه‌های مصنوعی عصبی به منظور انطباق، یادگیری، الگوریتم بردار ماشین پشتیبان به منظور دسته‌بندی و الگوریتم ژنتیک به منظور انتخاب بهترین ویژگی‌ها استفاده شد. بسیاری از محققین تمایل زیادی در استفاده از این ابزار را دارند، اما چالش آموزش شبکه‌های عصبی مهمترین دغدغه آنها می‌باشد. ترکیب الگوریتم‌های تکاملی و سیستم‌های داده کاوی می‌تواند یک ایده مناسب برای افزایش کارایی این شبکه‌ها در جهت پیش‌بینی صحیح باشند. البته هدف اصلی در این گونه تحقیقات بدست آوردن صحت و دقت ۱۰۰ درصدی می‌باشد، تا بتوان به سیستم‌های هوشمند مصنوعی اطمینان داشت و بعد عملی به آنها بخشید. میل و رسیدن به چنین هدفی تنها در صورت دانش بالا، فهم درست و تسلط کافی از یک طرف به سیستم‌های هوشمند و از طرفی دیگر به بیماری و داده‌کاوی امکان‌پذیر است، که در آینده‌ای نه چندان دور انشااله نایل خواهد شد.

سپاسگزاری

در انتها از تمام همکاری‌ها و حمایت‌های بی‌شائبه و صمیمانه ریاست مرکز تحقیقات دیابت شهرستان سبزوار و همین طور کارکنان دلسوز و مهربان آن مرکز که ما را در تهیه این تحقیق یاری رساندند بسیار ممنون و متشکریم.

مأخذ

1. World Health Organization, Diabetes Center, Fact Sheet N 312, www.who.int/mediacentre/factsheets/fs312/en, (08/29/2012).
2. Centers for Disease Control and Prevention, National Diabetes, Fact Sheet 2011 www.cdc.gov/diabetes, (08/29/2012).

۳. نعیم آبادی، محمد رضا؛ امیر احمدی چماچار، نوشاز؛ تهامی، احسان؛ ربانی، حسین. تشخیص دیابت با استفاده از SVM. *چکیده‌نامه چهاردهمین کنفرانس دانشجویی مهندسی برق ایران*؛ کرمانشاه؛ ایران؛ کرمانشاه، دانشگاه صنعتی کرمانشاه؛ ص ۲۵۱–۲۵۸.

4. American Diabetes Association, Diabetes Basics www.diabetes.org/diabetes-basics, (08/29/2012).

۵. فیوضی، محمد؛ قره‌خانی، اعظم؛ حدادنیا، جواد. تشخیص بیماری دیابت براساس ترکیب و تعامل سیستم‌های فازی، درخت تصمیم‌گیری و سیستم‌های فازی عصبی. *چکیده‌نامه چهارمین کنفرانس ملی فناوری اطلاعات و دانش*؛ اردیبهشت ۱۳۹۱؛ بابل؛ ایران؛ بابل: دانشگاه صنعتی بابل؛ ص ۱۵۰–۱۵٦.

٦. فیوضی، محمد؛ حدادنیا، جواد؛ ملانیا، نسرین. تشخیص بیماری دیابت براساس روش ترکیبی هوشمند. *مجله علمی و ترویجی هوش مصنوعی و ابزار دقیق ایران*؛ تابستان ۱۳۹۱، دوره اول (شماره ۳۲): ۸–۱.

۷. قوچانی، سعید راحتی؛ تهامی، سیداحسان. پیش‌بینی نوسانات سطح قند خون در بیماران مبتلا به دیابت نوع ۱ با استفاده از شبکه‌های عصبی خود بازگشتی المن. *مجله علمی و پژوهشی فنی و مهندسی دانشگاه آزاد اسلامی مشهد*؛ پاییز ۱۳۸۷، دوره دوم، (شماره۱): ۹٤–۸۱.

8. Jensen R. Ph.D. Thesis: Combining rough and fuzzy sets for feature selection, [School of informatics], University of Edinburgh; 2005 .

۹. علیپور، محمد، حدادنیا جواد. معرفی یک سیستم دقیق تشخیص سرطان پستان. *فصلنامه بیماری‌های پستان ایران*، تابستان ۱۳۸۸، شماره دوم (دوره ٤)، ص ۱۸–۲۹.

10. Xiang T, and Gong S. Spectral Clustering with eigenvector selection. *Journal of Pattern Recognition*, 2008; 41(3): p 1012-1029.

۱۱. اشک زری طوسی سهیلا، صدوقی یزدی هادی.خوشه بندی طیفی با انتخاب ویژگی Kernel PCA. *چکیده‌نامه شانزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران*؛ ایران؛ دانشگاه صنعتی شریف، تهران، ایران؛ اسفند ۱۳۸۹، ص ۱۱۸.

12. Kanan H, Faez K, Hosseinzadeh M. Face recognition system using ant colony optimization-based selected features. *Proceeding of IEEE Symp and Computational Intelligence in Security and Defense Applications*, 1-5 Aprill 2007; p 57-62.
13. Siedlecki W, Sklansky J. A note on genetic Algorithms for large-scale feature selection. *Journal of Pattern Recognition Letters*, 1989. 10(5): p. 335-47.
14. Bon A, Ogier J, Razali A, Yasin A. Feature selection in beltline Moulding process using genetic algorithm. *Journal of Application Science and Research*, 2008; 4(7):p 783-92.
15. Dorigo M, Maniezzo V, and Colona A .The ant system optimization by a colony of cooperating agents. *IEEE Transaction on Systems, Man, and Cybernetics-part B*, 1996: 26 (1): p 1-13.

16. Liu B, Abbass H, McKay B. Classification rule discovery with ant colony optimization. *Proceeding of IEEE Computational Intelligence Bulletin,* 25-29 April 1986; p 152 .
17. Meng Y. A swarm intelligence based algorithm for proteomic pattern detection of ovarian cancer. *Proceeding of IEEE Symp. Computational Intelligence and Bioinformatics and Computational Biology (CIBCB)*, 18-20 April 2006: p 7.
18. Lawrence D. *Handbook of genetic algorithms. Chapman and Hall Press*, New York, 1991, pp. 7-12 .
19. Michalewicz Z. Genetic Algorithms + Data Structures = Evolution Programs. Springer Press, New York, Third Edition, 1992; pp.3-15.
20. Zhang P, Verma B, Kumar K. Statistical classifier in conjunction with genetic algorithm based feature selection. *Journal of Pattern Recognition Letters 2005*. 26(2): p. 909-918.
۲۱. نعیم آبادی، محمد رضا؛ امیراحمدی چماچار، نوشاز، تهامی، احسان و ربانی، حسین. تشخیص دیابت با استفاده از SVM و MLP، چکیده نامه چهاردهمین کنفرانس *دانشجویی مهندسی برق ایران*؛ کرمانشاه؛ ایران؛ کرمانشاه، دانشگاه صنعتی کرمانشاه؛ ص ۲۵۱-۲۵۸.
۲۲. اصغری اسکوئی، محمد رضا . کاربرد شبکه‌های عصبی در پیش‌بینی سری‌های زمانی. *فصلنامه پژوهش‌های اقتصادی ایران*، دوره ٤، شماره ۱۲، پاییز ۱۳۸۱: ۵۹-٦٨.
۲۳. نیروئی مهیار، عبدالمالکی پرویز، گیتی معصومه. شبیه‌سازی یک مدل ترکیبی به کمک الگوریتم ژنتیکی و شبکه عصبی مصنوعی برای تفکیک الگوهای خوش‌خیم و بدخیم سرطان سینه در ماموگرافی. *مجله علمی و پژوهشی فیزیک پزشکی ایران*، دوره ۳ (شماره ۱۳)، زمستان ۱۳۸۵: ۲۵-۳٦.
24. Li J, Kuang Y, Mason C. Modelling the glucose-insulin regulatory system and ultradian insulin secretary oscillations with two time delays. *The journal of Mathematics and computer Science* 2012; 8(4), 25-33.
25. Leaning M.S. and Boroujerdi M.A. A system for compartmental modelling and simulation. Computer. *Methods programs biomed Journal* 1991; 35(4), 24-36.
26. Tresp V, Moody J and Delong W.R. Neural modelling of physiological processes. *Journal of Computational Learning Theory and Natural Learning Systems* 1994; 2(5): 254-268.
27. Huang H.P, 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. *Proceeding of 9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010)*; Leuven; Belgium; 2010: 5-7 July 2010, p 105 .
28. 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.
29. Fernandez de Canete J. Gonzalez-Perez S. Ramos-Diaz J.C. Artificial neural networks for closed loop control of in silico and ad hoc type 1 diabetes. *Journal of Computer methods and programs in biomedicine* 2010; 4(2): 128-135.
30. Scott M. Pappada B.S. Brent D. Cameron, Paul M. Development of a Neural Network for Prediction of Glucose Concentration in Type 1 Diabetes Patients. *Journal of Diabetes Science and Technology* 2008; 2(5): 105-118.
31. Stavroula G. Mougiakakou , Aikaterini P, Dimitra I, Konstantina S, Nikita A. Neural Network based Glucose – Insulin Metabolism Models for Children with Type 1 Diabetes. *Journal of biosim* 2011; 25(8): 15-25.
32. Wahab A, Kong Y. K, 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 (CBMS'06); India, New Delhi; India; 2006; p 158*.