

بررسی تشخیص بیماری دیابت بر اساس اطلاعات مستخرج از سیگنال ECG با استفاده از شبکه‌های عصبی مصنوعی

مرضیه نظری^۱، بهزاد زمانی دهکردی^{۳*}، فرشاد کیومرثی دهکردی^۳

^۱دانشجو، گروه کامپیوتر، دانشگاه آزاد اسلامی، واحد شهرکرد، شهرکرد، ایران؛ ^۳دانشگاه علوم پزشکی شهرکرد، شهرکرد، ایران؛ گروه

کامپیوتر، دانشگاه آزاد اسلامی، واحد شهرکرد، شهرکرد، ایران.

تاریخ دریافت: ۹۵/۶/۲۹ تاریخ پذیرش: ۹۵/۹/۲۹

چکیده:

زمینه و هدف: بیماری دیابت یکی از شایع‌ترین بیماری‌های دنیا شناخته شده است. یکی از مشکلات اساسی مربوط به این بیماری عدم تشخیص به موقع و صحیح آن می‌باشد. هدف این پژوهش ارائه روش جدیدی برای تشخیص بیماری دیابت است و قصد دارد برای اولین بار ارتباط تصاویر ECG با تشخیص بیماری دیابت به کمک شبکه عصبی مصنوعی و الگوریتم‌های داده کاوی را بررسی کند.

روش بررسی: در این مطالعه ۸ بیمار دیابتی و ۶۴ فرد سالم حضور داشتند. الکتروکاردیوگرافی برای تمام افراد انجام گرفت. اطلاعات مورد نیاز از تصاویر ECG شامل: نام بیمار، سن، HR، p، t، RR، PP، P، PR، qtcb،qt استخراج و در پایگاه داده جمع‌آوری شد. برای طبقه‌بندی بیماران از شبکه‌های عصبی احتمالی و الگوریتم‌های استاندارد داده کاوی استفاده شده است. داده‌ها از طریق الگوریتم‌های داده کاوی و روش‌های متفاوت کلاس‌بندی مورد بررسی و ارزیابی قرار گرفتند و نتایج هر یک با توجه به نرخ صحیح مقایسه شدند. از نرم‌افزار weka برای رده‌بندی‌ها استفاده شده است.

یافته‌ها: دقت شناسایی الگوریتم‌های مبتنی بر قوانین و شبکه عصبی، نسبت به الگوریتم‌های درخت تصمیم و الگوریتم‌های مبتنی بر فاصله بالاتر و نتایج بهتری در تشخیص بیماری دیابت نشان دادند. بهترین نرخ شایستگی در الگوریتم Consistency Subset Eval با میزان ۰/۸۹ بود و موج QRS به عنوان بهترین انتخاب در همه الگوریتم‌ها گزارش می‌شود. ارزیابی داده‌های افراد دیابتی و غیر دیابتی با استفاده از الگوریتم شبکه‌های عصبی احتمالی نرخ صحیحی ۹۵٪ را نشان داد. همچنین الگوریتم KNN کمترین پیچیدگی زمانی را نشان داد.

نتیجه‌گیری: مدل مبتنی بر قوانین دقت بالاتری نسبت به کلیه الگوریتم‌های طبقه‌بندی داده کاوی مورد استفاده در پژوهش نشان داد.

واژه‌های کلیدی: بیماری دیابت، هوش مصنوعی، شبکه عصبی، الکتروکاردیوگرافی، مدل‌های پیش‌بینی.

مقدمه:

و تعداد افرادی که با سن زیر ۲۰ سال به دیابت مبتلا شده بودند، حدود ۰/۲۵٪ از جمعیت آمریکا تخمین زده شد (۴).

دیابت از مهم‌ترین عوامل خطر بروز بیماری‌های قلبی و از مهم‌ترین علل قطع اندام، نابینایی و نارسایی مزمن کلیوی است (۵). گسترده‌گی انتشار و شیوع بالای دیابت و تحمیل هزینه‌های بالای تشخیص، کنترل و

بیماری دیابت شیوع بالایی در جوامع دارد. اثر دیابت بر کاهش امید به زندگی در کشورهای درحال توسعه بیشتر است (۲،۱). در سال ۲۰۱۳ حدود ۲۰۰ میلیون نفر در جهان و بیش از دو میلیون نفر در ایران به دیابت مبتلا بودند و سالانه شیوع دیابت در جهان حدود ۶٪ افزایش می‌یابد (۳). در سال ۲۰۱۲، شیوع دیابت در ایالات متحده ۱/۷ میلیون نفر اعلام شد

*نویسنده مسئول: شهرکرد- دانشگاه آزاد اسلامی- واحد شهرکرد- گروه کامپیوتر- تلفن: ۰۲۸-۳۳۳۵۰۰۸۱، E-mail: bzamani@gmail.com

و روابط بین آن‌ها و کشف دانش در داده‌ها که انجام این امر به صورت دستی امکان‌پذیر نمی‌باشد (۱۳). استفاده از داده کاوی جهت پیش‌بینی و تشخیص سریع و کم‌هزینه بیماری‌ها می‌تواند موثر باشد (۱۴).

امروزه ANN‌ها به‌طور گسترده‌ای در تحقیقات پزشکی توانسته‌اند، در تشخیص بیماری و کاهش اشتباهات به پزشکان کمک کنند و در کشف سرطان و مشکلات قلبی و تشخیص دیابت کاربرد داشته‌اند (۷). Uguz در مطالعه‌ای از شبکه‌های عصبی مصنوعی و تحلیل مولفه‌های اصلی در تشخیص بیماری‌های دریچه قلب بهره گرفت و از صداهای قلب به‌عنوان ورودی شبکه عصبی استفاده کرد. او با استفاده از تحلیل مولفه‌های اصلی بیان کرد کاهش ابعاد سیگنال‌های صداهای قلب اثر مثبتی بر طبقه‌بندی صداهای قلبی دارد (۱۵). در مطالعه دیگر ANN در مقایسه با رگرسیون لجستیک و درخت تصمیم، رهیافت بهتری برای پیش‌بینی عوارض پس از جراحی بیماران مبتلا به سرطان معده بود (۱۶). همچنین دقت پیش‌بینی و غربالگری ANN در تشخیص بیماری انفارکتوس میوکارد بررسی شده است (۱۷).

در پژوهش دیگر مشخص شد به کارگیری ANN نسبت به مدل رگرسیون لجستیک دو متغیره برای تشخیص همزمان بیماری دیابت و فشارخون دقت بالاتری دارد (۱۸). همچنین نفیسی و هاشمی گلپایگانی سیستم تزریق هوشمند انسولین در بیماران دیابتی با استفاده از شبکه عصبی و الگوریتم فازی را ارائه دادند (۱۹).

امروزه الکتروکاردیوگرام نقش مهمی در تشخیص اولیه، پیش‌بینی و پیشگویی بقای بیماران قلبی ایفا می‌کند. الکتروکاردیوگرافی ECG = Electrocardiography حاوی مقادیر زیادی اطلاعات مهم است که می‌تواند در موارد مختلفی استفاده شود. سیگنال ECG به تجزیه و تحلیل جنبه‌های آناتومیک و فیزیولوژیک عضله قلب کمک می‌کند (۲۰).

داده‌ها برای افراد مهم‌ترین منبع تصمیم‌گیری می‌باشد. داده‌های آزمایشگاهی، ابزاری برای جمع‌آوری

درمان بیماری باعث شده است. این بیماری همواره دغدغه سیستم‌های درمانی و بهداشتی جوامع باشد و پیشگیری از ابتلا به آن در اولویت باشد. تشخیص زودهنگام به‌منظور کاهش چشمگیری عوارض دیابت و بهبود کیفیت زندگی بیماران، نیز اتخاذ سیاست‌های اصلاحی در تغییر فرهنگ و سبک زندگی مردم ضروری می‌باشد (۶).

مشکل عدم تشخیص به‌موقع دیابت همواره مطرح بوده است؛ لذا ارائه روش دقیق تشخیصی جهت پیشگیری و کنترل این بیماری به‌خصوص در مراحل ابتدایی ارزشمند است. تاکنون برخی روش‌های هوشمند به‌منظور تشخیص بیماری‌ها مطرح شده‌اند. از آن جمله می‌توان به استفاده از روش‌های تکاملی و استفاده از روش الگوریتم‌های فازی تشخیص الگو در استخراج ویژگی و همچنین استفاده از روش شبکه‌های عصبی مصنوعی اشاره نمود (۷).

روشی برای پردازش اطلاعات، شبکه عصبی مصنوعی (ANN = Artificial Neural Network) است و بر اساس متغیرهای ورودی نسبت به طبقه‌بندی افراد به بیمار و سالم اقدام می‌کند و به پیشگویی وضعیت بیمار بر اساس عوامل خطر کمک می‌کند. در ANN از سیستم‌های عصبی زیستی جهت پردازش اطلاعات الهام گرفته شده است. ANN در واقع بخشی از عملکرد مغز را شبیه‌سازی می‌کند (۸-۱۰).

ANN در حل مسائلی که راه حل الگوریتمی بسیار پیچیده‌ای دارند و مسائلی که برای انسان‌ها آسان اما برای کامپیوترهای رایج دشوار هستند مانند تشخیص تصاویر و پیش‌بینی‌ها بر پایه دانش گذشته به‌خوبی عمل کرده‌اند (۱۱). از مهم‌ترین شبکه‌های عصبی که جهت شناخت و طبقه‌بندی الگوهای استفاده می‌شوند، می‌توان به شبکه‌های خودسازمان‌ده، شبکه‌های عصبی Bayesian، شبکه‌های عصبی شعاعی محور، شبکه‌های طبقه‌بندی کننده‌های احتمالی اشاره کرد (۱۲).

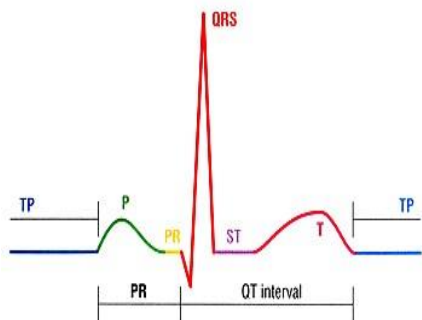
داده کاوی (Data Mining) روشی است معتبر برای تحلیل اتوماتیک داده‌ها و شناسایی الگوهای پنهان

هاجر (س) شهرکرد و ۸۰ فرد سالم در سطح استان چهارمحال و بختیاری می‌باشند، تشخیص ابتلا یا عدم ابتلا به بیماری دیابت زیر نظر مستقیم متخصص قلب انجام شد. سپس با اخذ رضایت‌نامه از تمامی افراد، تصاویر ECG با رعایت استانداردهای لازم اخذ گردید، اطلاعات بالینی بیماران نیز زیر نظر متخصص قلب اخذ و در پرسشنامه‌های مربوطه درج گردید.

معیارهای خروج شامل عدم تمایل افراد، نقص در تجهیزات الکتروکاردیوگرام و اشکال در روش الکتروکاردیوگرافی بود که با در نظر گرفتن این معیارها در نهایت ۷۲ نمونه (۸ نفر بیمار دیابتی و ۶۴ غیر دیابتی) در مطالعه حضور داشتند.

دستگاه ECG چند کاناله، ۱۰ الکتروده دارای ۱۲ کابل هادی ECG به نام BTL ساخت کشور انگلیس استفاده شد. مدت زمان گرفتن تصاویر حدوداً ۲۵ دقیقه بود. پس از تحلیل و تفسیر نوار قلب، ویژگی‌های و پارامترهای مورد نیاز جمع‌آوری و تصاویر ECG در کاغذ شطرنجی الکتروکاردیوگرام چاپ و نگهداری شد.

جریان‌های الکتریکی ماهیچه قلب را منقبض و منبسط می‌کنند. امواج الکتریکی در ECG ثبت می‌شوند. در ECG یک قطعه ST و ۴ فاصله اصلی شامل Q.R، R-R، P.R، QRS دیده می‌شود که تغییرات هر کدام نمایانگر بیماری خاصی است. (۲۰، ۲۱) (تصویر شماره ۱). مهم‌ترین پارامترهای تصاویر ECG بررسی شده در جدول شماره ۱ نشان داده شده است.



تصویر شماره ۱: نام‌گذاری پارامترها در یک

الکتروکاردیوگراف طبیعی (۲۲)

اطلاعات در تحقیقات می‌باشد که از طریق این ابزار اطلاعات تفصیلی (کمی و کیفی) مناسبی در اختیار پژوهشگر قرار می‌گیرد. داده‌کاوی ابزار و فناوری مهمی برای استفاده سودمند از این داده‌ها به شمار می‌رود. در این راستا پژوهش حاضر، از الگوریتم‌های کلاس‌بندی و ساخت درختان تصمیم‌گیری استفاده کرده است. با توجه به نتایج حاصل از مطالعات صورت پذیرفته و نقش بسیار متفاوت الگوریتم‌های داده‌کاوی و ANN در تشخیص بیماری‌های گوناگون و همچنین ارتباط مثبت آن در تشخیص بیماری دیابت و از آنجا که تاکنون هیچ مطالعه‌ای به منظور تعیین ارتباط تصاویر ECG با بیماری دیابت به کمک الگوریتم‌های داده‌کاوی و ANN انجام نشده است و با توجه به تعداد زیاد و رو به افزایش مبتلایان به دیابت در استان چهارمحال و بختیاری، مطالعه حاضر طراحی شد. در واقع در این مطالعه تلاش شده با به کارگیری ویژگی‌های مبتنی بر تصاویر ECG برای طبقه‌بندی بیماران، از شبکه‌های عصبی احتمالی (PNN= Probabilistic Neural Network) و الگوریتم‌های استاندارد داده‌کاوی بر روی پایگاه داده استاندارد دیابت؛ الگو و روش جدیدی را در پیش‌بینی بیماری دیابت بررسی کند. همچنین نوع تصاویر ECG، فراوانی و سطح پارامترهای استخراجی از تصاویر ECG در افراد بیمار و سالم تعیین و با هم مقایسه می‌گردد. فرض بر این بین تصاویر ECG و بیماری دیابت در این استان ارتباط وجود دارد و انتخاب یا اضافه کردن ویژگی‌های جدید و موثر می‌توان کارآیی شبکه عصبی مصنوعی را در تشخیص بیماری دیابت بهبود داد.

روش بررسی:

پژوهش کاربردی حاضر از نوع توصیفی-تحلیلی در حوزه کامپیوتر و علوم پایه پزشکی است که در سال ۱۳۹۴ صورت پذیرفت. در این مطالعه نمونه‌گیری به روش آسان انجام شد. نمونه بررسی شده ۲۰ بیمار مبتلا به دیابت مراجعه کرده به مرکز آموزشی درمانی

جدول شماره ۱: متغیرهای ECG برای مدل‌سازی پیش‌بینی دیابت

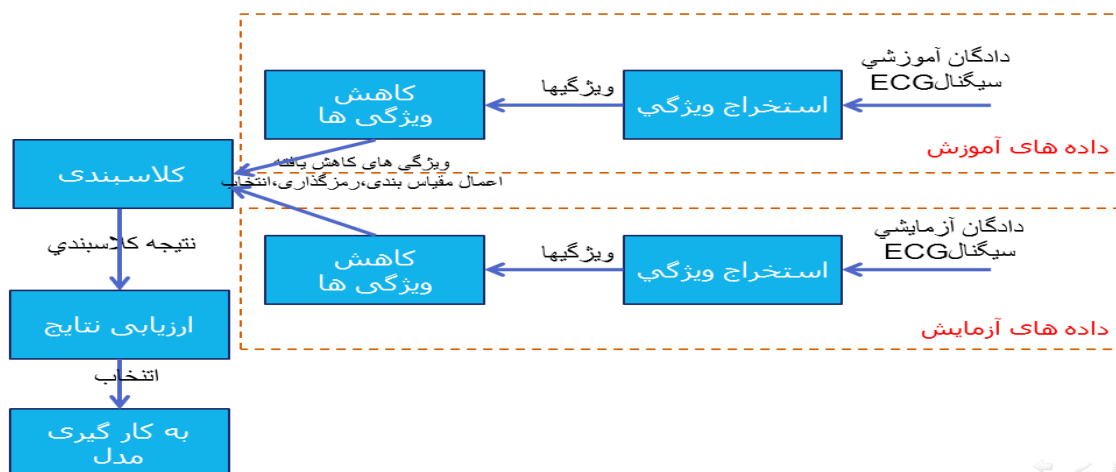
نام پارامتر	توضیح	مشخصات
موج P	عبور جریان الکتریکی دهلیزها (نشان‌دهنده دپولاریزاسیون دهلیزهاست)	اولین موج ECG - گرد و صاف و قرینه
موج T	مراحل انتهایی رپولاریزاسیون بطن‌ها	گرد و مثبت - بعد از QRS
موج U	بعد از T ظاهر می‌شود، همیشه دیده نمی‌شود	گرد و کوچک
فاصله PR	زمان سپری شده برای رسیدن موج دپولاریزاسیون از دهلیز به بطن	ابتدای P تا شروع QRS
فاصله QT	زمان لازم برای مجموع فعالیت بطن‌ها در یک چرخه قلبی	ابتدای QRS تا انتهای T
قطعه ST	مراحل ابتدایی رپولاریزاسیون بطن‌ها	انتهای QRS تا ابتدای T
کمپلکس QRS	نشان دپولاریزاسیون بطن‌ها	شامل ۳ موج

الگوریتم‌های داده‌کاوی و روش‌های متفاوت کلاس‌بندی مورد بررسی و ارزیابی قرار گرفتند و نتایج هر یک با توجه به نرخ صحیح مقایسه شدند.

برای شناسایی متغیرهای تأثیرگذار در روند پیاده‌سازی، از مصاحبه آزاد با پزشک قلب استفاده شد. نوع داده‌های این پژوهش به صورت عددی (پیوسته) و اسمی (گسسته) است. به منظور کاهش زمان و مدیریت داده برای طبقه‌بندی بیماران از شبکه‌های عصبی احتمالی و الگوریتم‌های استاندارد داده‌کاوی استفاده شده است. این پروژه داده‌کاوی، شامل ۶ مرحله است: داده‌های سیگنال ECG، استخراج ویژگی‌ها، کاهش ویژگی‌ها، ویژگی‌های کاهش‌یافته، کلاس‌بندی، ارزیابی نتایج و به کارگیری مدل (تصویر شماره ۲).

برای تفسیر ECG و قضاوت در مورد بیماری توجه به تمام اجزا، امواج، قطعات، فواصل و اطلاعات موجود بر روی نوار قلب ضروری است. مهم‌ترین آن‌ها عبارت‌اند از: محاسبه سرعت ضربان قلب (HR)؛ تعیین نظم با توجه به فواصل R-R؛ بررسی امواج P؛ تعیین فواصل و نظم PR و نهایتاً اندازه‌گیری عرض و نظم کمپلکس‌های QRS (۲۳).

در پژوهش حاضر نیز اطلاعات مورد نیاز از تصاویر ECG شامل: نام بیمار، سن، HR، t ، qrs ، p ، RR ، PP ، P ، PR ، qt ، $qtcb$ ، QRS استخراج و در پایگاه داده جمع‌آوری شد. در واقع روش حاضر، مجموعه‌ای از الگوریتم‌های یادگیری ماشینی و ابزارهای پیش‌پردازش و داده‌کاوی می‌باشد که داده‌ها را وارد نموده و از طریق



تصویر شماره ۲: ساختار بلوک دیاگرام کلی در روش پیشنهادی

است. در این پژوهش از روش داده کاوی پیش‌گویانه استفاده شد؛ چون هدف پژوهش پیش‌بینی و تشخیص بیماری دیابت است. این روش مدلی از سیستم را ارائه می‌دهد که شامل به کارگیری متغیرها و فیلدها در Data Warehouse ها جهت پیشگویی مقادیر ناشناخته می‌باشد (۲۴).

از آنجا که داده‌ها در این مطالعه به صورت نامتعادل بودند، از روش دسته‌بندی و پیش‌بینی استفاده شد (۲۵). دسته بندی افراد با توجه به ویژگی‌های تصاویر ECG انجام شد. تعداد محدود و از پیش تعیین شده‌ای از دسته‌ها وجود داشت و انتظار داریم بتوانیم هر اطلاعاتی را به یک یا دو مورد از آن‌ها تخصیص دهیم. فن‌های درخت تصمیم و نزدیک‌ترین همسایه از جمله فن‌های دسته بندی می‌باشند (۲۶).

از داده‌های پیشین برای تهیه ی یک مدل که بیانگر رفتار مشاهده شده ی کنونی است استفاده می‌شود. وقتی این مدل برای ورودی‌های کنونی به کار رفت؛ نتیجه ی کار، پیش‌بینی رفتار آینده خواهد بود. پیش‌بینی اینکه چند درصد از افراد با توجه به ویژگی‌های تصاویر ECG در آینده در معرض ابتلا به دیابت هستند.

در این مطالعه از نرم‌افزار weka برای رده‌بندی‌ها استفاده شد. این ابزار الگوریتم‌های طبقه‌بندی و پیش‌بینی بسیار متنوعی را پیاده‌سازی می‌کند. رده‌بندی‌های مورد استفاده در این بخش در ۳ گروه طبقه‌بندی می‌شوند: گروه رده‌بندی‌های مبتنی بر فاصله مانند IBK و KStar، گروه رده‌بندی‌های آماری مانند Naive Bayes و Bayes Net و گروه رده‌بندی‌های مبتنی بر درخت تصمیم‌گیری مانند J48، Lmt و RandomTree.

پایگاه داده نهایی برای داده کاوی، طبقه‌بندی و کلاس‌بندی، با ۷۲ رکورد و ۱۱ فیلد از خصیصه‌های تصاویر ECG به دست آمد. برای فراخوانی کلاس‌بندی، باید فیلدها را به‌طور ورودی معرفی کرد.

با بررسی تصاویر ECG، افراد در دو طبقه بیمار دیابتی و افراد عاری از دیابت تقسیم و دانش مربوط به پیش‌بینی بیماری بر اساس الگوریتم‌های داده کاوی استخراج شد. در مجموع ساماندهی تصاویر ECG به یک شکل استاندارد یا جدول رابطه‌ای انجام شد تا برای پردازش به‌وسیله تکنیک‌های داده کاوی استفاده شود.

مرحله استخراج ویژگی‌ها شامل انتخاب جداول، رکوردها و خصیصه‌ها و همچنین انتقال و پاک‌سازی داده برای داده کاوی انجام شد. در این مرحله، داده‌های مرتبط با تصاویر ECG که در مرحله قبل شناسایی شدند، استخراج و ثبت شدند و در یک پایگاه داده جامع و یکپارچه، به‌منزله یک پایگاه داده رابطه‌ای قرار گرفتند.

در مرحله کاهش ویژگی‌ها، همه ویژگی‌های استخراج شده از مرحله قبلی پالایش و ارزیابی می‌شود و تبدیلات خاصی روی داده‌ها انجام می‌گیرد و برخی از ویژگی‌هایی که در عمل داده کاوی از اهمیت برخوردار نیستند و مبهم می‌باشند مانند سن افراد، نام و نام و خانوادگی را از مجموعه دادگان کاهش یافتند. تبدیلات شامل آشکارسازی (حذف) داده‌های غیرعادی و تعیین ویژگی‌های مقیاس بندی، رمزگذاری و انتخاب جهت تحلیل و بررسی انجام شد.

در مرحله کلاس‌بندی، انواع روش‌های داده کاوی انتخاب و به کار گرفته می‌شوند. در پژوهش حاضر، پیاده‌سازی داده‌ها، به‌طور جداگانه روی هر الگوریتم طبقه‌بندی انجام و الگوریتمی که بالاترین صحت را داشت، مبنای مدل‌سازی و استخراج دانش از آن قرار گرفت. هدف از این کار، استخراج نتیجه موردنظر برای پیش‌بینی و تشخیص بیماری، با توجه به داده‌های آموزشی از الگوریتم انتخابی است. طبقه‌بندی روش‌های داده کاوی به دو نوع داده کاوی توصیفی و داده کاوی پیش‌گویانه

شدند. تعداد داده‌ها محاسبه می‌گردد که تعداد سطرها برابر تعداد بیماران یا همان داده‌های ورودی شبکه عصبی هستند، باید داده‌ها به صورت تصادفی انتخاب شوند که در قسمت تست و آموزش از هر دو نوع جمعیت وجود داشته باشد. سپس در قسمت Test PNN Network شبکه برای داده‌های آموزش، داده‌های تست و همه داده‌ها تست شد و خطای هر قسمت به صورت جداگانه محاسبه شد.

یافته‌ها:

در این مطالعه تمام شرکت کنندگان از جنس مونث بودند. در تکنیک داده کاوی مطالعه حاضر از گروه پیش‌بینی کننده‌ها و از روش دسته‌بندی استفاده شد، رتبه‌بندی متغیرهای ورودی توسط الگوریتم‌های کلاس‌بندی برای پیش‌بینی انجام شد. پس از ورود داده‌ها، روش‌های کلاس‌بندی یا طبقه‌بندی الگوریتم‌های متفاوت انجام و نتایج هر یک با توجه به نرخ صحیح بررسی و مقایسه شد.

یافته‌های مربوط به ارزیابی مدل پیشنهادی روش‌های کلاس‌بندی Bayesian Logistic Regression و Voted Perceptron از انواع الگوریتم‌های مبتنی بر قوانین و مدل با نرخ صحیح ۸۸/۸۹٪ بهترین نرخ صحیح را دارند. روش Naive Bayes از انواع الگوریتم‌های مبتنی بر قوانین و مدل با عدد ۰/۸۷ بهترین دقت و صحت را دارد. Bayes Net از انواع الگوریتم‌های مبتنی بر قوانین و مدل و SMO از انواع الگوریتم‌های مبتنی بر قوانین با میزان ۰/۸۹٪ بهترین نرخ فراخوانی را دارند. Naive.Bayes Updateable از انواع الگوریتم‌های مبتنی بر قوانین و مدل با ۰/۸۸٪ بهترین نرخ و دقت اندازه‌گیری را داشته و در اغلب موارد نسبت به دیگر تبدیلات عملکرد بهتری داشته است (جدول شماره ۲).

برای سنجش اعتبار و صحت مدل، از تقسیم نمونه به دو مجموعه داده آموزشی و آزمایشی استفاده شده است. میزان اعتبار با نتایج داده‌های جدید آزمون می‌شود و داده‌های آزمایشی به منزله ناظر به الگوریتم وارد شده و نتایج میزان صحت آن را ارزیابی می‌کند. داده‌های آموزش بدون هیچ خطایی خروجی هدف را دنبال می‌کردند؛ ولی داده‌های تست دارای تفاوت‌هایی بودند. معیار اعتبار و صحت مدل، بسته به صحت کلاس‌بندی یا تفکیک داده‌های آزمایشی است. پس از اینکه فرایند دوباره تکرار شد، دقت نهایی مدل ارائه می‌شود.

شبکه‌های عصبی احتمالی دارای ۳ ورودی می‌باشد که به ترتیب مربوط به الگوهای ورودی، الگوهای خروجی و فضایی است که برای همسایگی داده‌ها مشخص می‌شود که اگر این مقدار نزدیک به صفر باشد، شبکه عمل طبقه‌بندی را نسبت به نزدیک‌ترین همسایه انجام می‌دهد و هرچه دارای مقدار بزرگ‌تری باشد فاصله را نسبت به چندین همسایه (بردار) نزدیک محاسبه می‌کند و عمل طبقه‌بندی را انجام می‌دهد. سیستم پیشنهادی با استفاده از ترکیب روش‌های مذکور موفق شد با تکیه بر ویژگی‌های پایگاه داده در قالب ترکیب و تعامل به نرخ صحیح و دقت شناسایی دست یابد.

مقادیر در نرم‌افزار اکسل به صورت یک ماتریس 12×72 می‌باشند که ۱۱ ستون اول ورودی‌ها و ستون ۱۲ خروجی سیستم است (داده‌های مربوط به سن بیماران لحاظ نشده است؛ زیرا تأثیر زیادی در بیماری دیابت و تشخیص آن با شبکه عصبی ندارد). پس با توجه به این موضوع داده‌های ستون اول تا یازدهم در Inputs و داده‌های ستون دوازدهم در Targets ریخته شد. بعد از آن داده‌های آموزش و تست از هم جدا شدند. در این پژوهش ۷۰٪ داده‌ها از نوع داده‌های آموزش و ۳۰٪ داده‌های تست انتخاب

جدول شماره ۲: میزان نرخ صحیح، دقت و صحت، فراخوانی و دقت اندازه‌گیری روش‌های مختلف کلاس بندی

نوع الگوریتم	روش کلاس‌بندی	نرخ صحیح	دقت و صحت	فراخوانی	دقت اندازه‌گیری	
مبتنی بر فاصله	Knn (k=3)	٪۸۶/۱۱	۰/۷۹	۰/۸۶	۰/۸۲	
	Knn (k=5)	٪۸۶/۱۱	۰/۷۹	۰/۸۶	۰/۸۲	
	Kstar	٪۸۴/۷۲	۰/۸۴	۰/۸۵	۰/۸۴	
	IBk	٪۸۰/۵۶	۰/۸۲	۰/۸۰	۰/۸۱	
	LWL	٪۸۳/۳۳	۰/۸۳	۰/۸۳	۰/۸۳	
مبتنی بر قوانین و مدل	Bayes Net	٪۸۸/۸۸	۰/۷۹	۰/۸۹*	۰/۸۳	
	Naïve Bayes	٪۸۷/۵	۰/۸۷*	۰/۸۸	۰/۸۷	
	Bayesian Logistic Regression	٪۸۸/۸۹*	۰/۷۹	۰/۸۸	۰/۸۳	
	Naive Bayes Updateable	٪۸۷/۵	۰/۸۶	۰/۸۸	۰/۸۸*	
مبتنی بر قوانین	SMO	٪۸۸/۸۸	۰/۷۹	۰/۸۹*	۰/۸۴	
	RBF network	٪۸۶/۱	۰/۸۲	۰/۸۶	۰/۸۴	
	VFI	٪۶۳/۸۸	۰/۷۹	۰/۶۳	۰/۷۰	
	SPegasos	٪۷۹/۱۷	۰/۸۰۲	۰/۷۹	۰/۷۸	
	Voted Perceptron	٪۸۸/۸۹*	۰/۷۹	۰/۸۹	۰/۸۳	
	Logistic	٪۸۰/۵۶	۰/۷۹	۰/۸۰	۰/۸۰	
	Multilayer Perceptron	٪۸۳/۳۳	۰/۸۳	۰/۸۳	۰/۸۳	
	مبتنی بر درخت تصمیم	Decision Stump	٪۸۶/۱۱	۰/۷۹	۰/۸۶	۰/۸۲
		OneR	٪۸۳/۳۳	۰/۷۸	۰/۸۳	۰/۸۰
		J48	٪۷۹/۱۶	۰/۷۸	۰/۷۹	۰/۷۹
LMT		٪۸۱/۹۴	۰/۷۸	۰/۸۱	۰/۸۰	
BF TREE		٪۸۷/۵	۰/۷۹	۰/۸۸	۰/۸۳	
Random Tree		٪۸۱/۹۴	۰/۸۰	۰/۸۱	۰/۸۱	
ADTree		٪۸۴/۷۲	۰/۸۲	۰/۸۵	۰/۸۳	

*: بهترین مقادیر مربوط به هر آیتیم.

یکسانی در همه الگوریتم‌ها برخوردار است؛ ولی بهترین نرخ شایستگی در الگوریتم Consistency Subset Eval بالاترین میزان و برابر ۰/۸۹ بود.

یافته‌های مربوط به الگوریتم‌های انتخاب ویژگی طبق جدول شماره ۳ برای انتخاب بهترین جستجو، از الگوریتم Rank Search استفاده شد که لیست داده‌ها، از ترتیب

جدول شماره ۳: رتبه‌بندی داده‌های ECG به روش Rank Search

Cost Sensitive Subset Eval	Wrapper Subset Eval	Consistency Subset Eval	Classifier Subset Eval	Cfs Subset Eval	Attribute Evaluator Selection	داده‌های ECG
۱۱	۱۱	۱۱	۱۱	۱۱		HR
۳	۳	۳	۳	۳		qtcb
۲	۲	۲	۲	۲		qt
۵	۵	۵	۵	۵		P zar
۴	۴	۴	۴	۴		PR
۶	۶	۶	۶	۶		RR
۱۰	۱۰	۱۰	۱۰	۱۰		qrt
۹	۹	۹	۹	۹		t
۸	۸	۸	۸	۸		P
۷	۷	۷	۷	۷		PP
۱	۱	۱	۱	۱		QRS
۰	۰/۱۱	۰/۸۹	۰/۱۱	۰		نرخ بهترین شایستگی

الگوریتم‌ها برخوردار است و QRS به‌عنوان بهترین انتخاب گزارش می‌شود (جدول شماره ۴).

برای داشتن بهترین انتخاب، از الگوریتم Best First استفاده شد که لیست داده‌ها از ترتیب یکسانی، در همه

جدول شماره ۴: رتبه‌بندی داده‌های ECG به کمک الگوریتم Best First

Cost Sensitive Subset Eval	Wrapper Subset Eval	Consistency Subset Eval	Classifier Subset Eval	Cfs Subset Eval	Attribute Evaluator Selection	داده‌های ECG
HR	HR	HR	HR	HR		HR
qtcb	qtcb	qtcb	qtcb	qtcb		qtcb
qt	qt	qt	qt	qt		qt
P zar	P zar	P zar	P zar	P zar		P zar
PR	PR	PR	PR	PR		PR
RR	RR	RR	RR	RR		RR
qrt	qrt	qrt	qrt	qrt		qrt
t	t	t	t	t		t
P	P	P	P	P		P
PP	PP	PP	PP	PP		PP
QRS	QRS	QRS	QRS	QRS		QRS

بعد از ارزیابی داده‌های افراد دیابتی و غیر دیابتی با استفاده از الگوریتم شبکه‌های عصبی احتمالی نرخ صحیحی ۹۵٪ به دست آمد. ارزیابی پیچیدگی زمانی الگوریتم‌های با نرخ صحیح بالا نشان داد. الگوریتم KNN پیچیدگی زمانی کمتری نسبت به بقیه الگوریتم‌ها دارد (جدول شماره ۵).

برای انتخاب جستجو ژنتیکی، از روش Genetic Search استفاده شد که با الگوریتم‌های متفاوت، از نتایج یکسان برخوردار است. در الگوریتم‌های متفاوت انتخاب پارامتر qtcb، احتمال تقاطع ۰/۶، احتمال جهش ۰/۰۳۳، فرکانس گزارش ۲۰، کاوش تصادفی ۱، تعداد نسل ۲۰ به دست آمد.

جدول شماره ۵: ارزیابی پیچیدگی زمانی الگوریتم‌ها

نوع الگوریتم	پیچیدگی زمانی	توضیحات
Best-first Search	$O(V + E)$	جستجوی سطح به سطح
KNN	$o(n)$	طبقه‌بندی نزدیک‌ترین مجاور
τ -PNN	$O(\tau N^2)$	N تعداد موضوعات آموزشی
Ad tree	$O(N^2)$	N تعداد موضوعات آموزشی
Naive Bayes	$O(D Lave + C V)$	زمان آزمون (طول متوسط یک سند آزمون)

بحث:

ویژگی مشترک انتخاب شد. این داده‌ها بیولوژیکی بوده و مربوط به بیان تصویر قلب انسان است؛ حجم دیتاست‌های مورد استفاده چندان زیاد نیست.

پژوهش حاضر یک مطالعه تشخیصی بود که بر اساس متغیرهای ورودی، به تشخیص بیماری دیابت پرداخت. مدل مبتنی بر قوانین دقت بالاتری نسبت به کلیه الگوریتم‌های طبقه‌بندی داده‌های مورد استفاده در پژوهش دارد که مهم‌ترین شاخص‌ها در این الگوریتم، بررسی و دقت و صحت هر یک ارزیابی شدند. بنابراین دانش استخراج شده از این درخت، مورد اعتقادترین دانش حاصل از داده‌های مورد بررسی به شمار رفته و می‌تواند مبنای استخراج قوانین داده‌های مورد بررسی باشد.

با توجه به مطالعه سایر الگوریتم‌ها در زمینه تشخیص دیابت، نتایج حاکی از آن است که هرگز نمی‌توان الگوریتمی را به‌عنوان الگوریتم بهینه معرفی

با توجه به اهمیت و حساسیت شبکه عصبی مصنوعی و داده‌کاوی در علم پزشکی و همچنین نیاز مبرم این صنعت به تغییر مسیر از پزشکی سنتی به سمت پزشکی مبتنی بر شواهد، لذا در پژوهش حاضر تشخیص بیماری دیابت به کمک تصاویر ECG با الگوریتم‌های داده‌کاوی و کلاس‌بندی و شبکه عصبی بررسی شد. ویژگی‌های استخراج شده از تصاویر ECG مورد تفسیر قرار گرفت و روش پیشنهادی ارتباط تصاویر ECG با بیماری دیابت بیان و یک دیتاست از مجموعه داده‌های ECG گرفته شد که ویژگی‌های کلی این دیتاست‌ها به شرح ذیل بیان شده است:

این داده‌ها از طریق دستگاه‌های ECG استخراج شده‌اند؛ فرض بر این بود که داده‌های مورد استفاده داده‌های گم شده ندارند؛ داده‌های منتخب تک‌بعدی هستند و در صورت داشتن ویژگی‌های مختلف تنها یک

بیشترین تأثیر را داشته‌اند. به کمک درخت تصمیم ایجاد شده، قوانینی استخراج شده که می‌تواند به‌عنوان الگویی برای پیش‌بینی وضعیت بیماران و احتمال بروز عوارض در آن‌ها استفاده شود. صحت مدل ایجاد شده بر روی داده‌های مورد استفاده در درخت تصمیم ۸۹/۷۴٪ و در شبکه عصبی مصنوعی ۵۱/۲۸٪ گزارش شد (۳).

Cho و همکاران به کمک دسته‌بندی (SVM= Support Vector Machine) و با روش انتخاب ویژگی و تجسم سازی، وجود نروپاتی در بیماران دیابتی را پیش‌بینی کردند (۳۰). در مطالعه دیگر که به‌منظور مقایسه شبکه عصبی پرسپترون و رگرسیون لجستیک در بررسی متابولیسم قند بیماران دیابتی، مدل شبکه عصبی پرسپترون از قدرت بیشتری در پیش‌بینی مدل برخوردار بود (۳۱).

برفنه ئی و همکاران جهت پیش‌بینی ابتلا به دیابت، مدل ANN ۳ لایه با معماری (۵۳:۲۰:۲) با صحت پیش‌بینی آموزش ۹۲٪ و پیش‌بینی آزمون ۹۱/۶٪ بهترین مدل دانستند. پیشنهاد شد از این مدل برای پیش‌بینی ابتلا به بیماری دیابت استفاده شود (۶).

با توجه به مطالعات انجام گرفته در این زمینه، می‌توان گفت بهره‌گیری از داده‌های پزشکی، بیشتر در حوزه طبقه‌بندی بیماری‌ها و علائم و پیش‌بینی بیماری آن‌ها بوده است. نظر به اهمیت مقوله داده‌های پزشکی، استفاده از شبکه‌های عصبی مصنوعی در علم پزشکی، لازم است در کشور ما نیز در واحدهای پژوهشی و مراکز تحقیقاتی به این حیطه از علم توجه ویژه گردد و از داده‌های فراوان موجود و داده‌های به‌دست آمده، جهت استفاده هوشمند در پیش‌بینی، تشخیص و درمان بیماری‌ها بهره‌برند.

مزایای استفاده از شبکه‌های عصبی مصنوعی این است که تحت تأثیر عوامل خستگی، شرایط کاری و موقعیت‌های عاطفی قرار نمی‌گیرند. از دیگر مزیت‌ها خودسازمان‌دهی، یادگیری تطبیقی، تحمل خطا، دسته‌بندی، تعمیم دهی در تشخیص بیماری‌ها می‌باشد (۲۶).

کرد و برای هر کاربرد باید با توجه به مجموعه داده مورد استفاده بهترین الگوریتم را معرفی نمود (۲۷). به‌رحال الگوریتم‌های مبتنی بر قوانین و شبکه عصبی احتمالی، یکی از ابزارهای قوی و متداول در داده‌کاوی می‌باشند که درک، پیاده‌سازی و کاربرد آسان داشته و از نظر محاسباتی ارزان می‌باشد. بر اساس نتایج از بین پارامترهای ECG بررسی شده در این مطالعه موج QRS بهترین گزینه جهت پیش‌گویی ابتلا به بیماری دیابت می‌باشد و پیشنهاد می‌گردد جهت استفاده از شبکه‌های عصبی در بررسی ابتلا به بیماری دیابت از این پارامتر بهره‌گرفته شود. کمپلکس QRS شامل ۳ موج می‌باشد که بعد از موج P با مکتبی کوتاه ظاهر می‌شود و خط نوار قلب حرکتی سریع به سمت پایین، بالا و دو مرتبه به سمت پایین دارد. این ترکیب در اثر انقباض بطن‌ها ایجاد می‌شود.

مطالعات در خصوص استفاده از شبکه عصبی مصنوعی بر روی تصاویر ECG محدود بوده است. Bani-Hasan و همکاران برای شناسایی سیگنال‌های الکتروکاردیوگرام ضربان غیرطبیعی قلب، روشی جدید ارائه کردند. آن‌ها از یک شبکه عصبی مصنوعی چندلایه پیش‌رو با الگوریتم پس‌انتشار برای دسته‌بندی انواع ضرب‌آهنگ‌های قلبی استفاده کردند (۲۸). در یک پژوهش از شبکه‌های عصبی مصنوعی برای پردازش‌های چندمنظوره ECG استفاده شد که نتایج نشان داد، استفاده از شبکه عصبی در کاهش نویزهای موجود در نوار قلب موثر است (۲۹). مطالعات متعددی جهت ارزیابی کارایی شبکه‌های عصبی مصنوعی در تشخیص بیماری‌ها و حتی تشخیص دیابت انجام شده است؛ ولی هیچ‌کدام از مطالعات موجود ارتباط تصاویر ECG با تشخیص بیماری دیابت به کمک ANN و الگوریتم‌های داده‌کاوی را بررسی نکرده‌اند.

در مطالعه عامری و همکاران، به‌منظور استخراج دانش از داده‌های بیماران دیابتی با استفاده از داده‌های پزشکی و روش‌های آن مشخص شده که متغیرهای فشارخون بالا، سن و سابقه خانوادگی در عوارض مشاهده شده

سلامت و بیمارستان‌ها و نیز مقایسه نتایج انواع الگوریتم‌های دیگر برای رسیدن به هدف بهتر پیشنهاد می‌گردد. پیشنهاد می‌گردد تحقیق بر روی داده‌های بیشتری (بیماران بیشتر) صورت پذیرد و قبل از انجام مراحل داده‌کاوی (Preprocessing) نیز بر روی داده‌ها صورت پذیرد تا نتایج دقیق‌تر حاصل گردد.

نتیجه‌گیری:

نظر به اهمیت تشخیص به‌موقع بیماری دیابت در این تحقیق، از شبکه عصبی مصنوعی و داده‌کاوی تکنیک‌های آن از شبکه عصبی مصنوعی و الگوریتم‌های داده‌کاوی جهت بررسی ارتباط داده‌های تصاویر ECG با تشخیص بیماری دیابت استفاده شد. در مجموع بعد از تست‌های انجام شده برتری نسبی روش‌های پیشنهادی مشخص بود. از میان انواع الگوریتم‌های کلاس‌بندی، الگوریتم‌های مبتنی بر قوانین و شبکه عصبی احتمالی، در تشخیص بیماری دیابت از دقت و عملکرد بالاتری برخوردارند و در راستای بهبود کیفیت سلامت نقش مهمی دارند.

تشکر و قدردانی:

این مقاله ناشی از پایان نامه کارشناسی ارشد دانشگاه آزاد اسلامی واحد شهرکرد با کد طرح ۲۹۲۷-۹۵-۰۱-۱۳۹۵ می‌باشد. بدین وسیله از حمایت این دانشگاه و همچنین دانشگاه علوم پزشکی شهرکرد و کلیه افرادی که در انجام این مطالعه همکاری نمودند، قدردانی می‌گردد.

استفاده همزمان و ترکیبی الگوریتم‌های تکاملی، دستگاه‌های داده‌کاوی و شبکه عصبی مصنوعی، می‌تواند منجر به افزایش کارایی شبکه‌های مصنوعی و دسته‌بندی صحیح داده‌ها شود. البته هدف اصلی در این تحقیقات به دست آوردن صحت، دقت و اطمینان کامل به سیستم‌های هوشمند مصنوعی نیست و باید با دانش بالا، فهم درست و تسلط کافی به سیستم‌های هوشمند و ANN و تسلط به بیماری مورد نظر و داده‌کاوی آن، سعی کرد تا میزان اشتباهات را به کمترین میزان رساند (۳۲). استفاده همزمان از الگوی تشخیصی هوشمند و اطلاعات پزشکی بر اساس مشاهدات بالینی می‌تواند روند تشخیص و پیش‌بینی بروز عوارض و درمان بیماران دیابتی را تسریع کرده و با عوارض غیرقابل کنترل دیابت مقابله شود (۶).

نتایج مطالعه حاضر نشان می‌دهد ANN در رابطه با داده‌های مربوط به دیابت از توانایی پیش‌بینی با دقت بالایی برخوردار می‌باشد و تأییدی بر نتایج مطالعات قبلی در زمینه کارایی شبکه‌های عصبی مصنوعی در تشخیص بیماری‌ها می‌باشد و به کمک مدل شبکه‌های مصنوعی می‌توان به مدیریت بیماری دیابت کمک کند.

با توجه به اینکه این مطالعه اولین تحقیق در این زمینه بود؛ بنابراین محدودیت‌هایی در انجام مطالعه وجود داشت. برای پژوهش‌های آتی انجام مطالعه در محیط پژوهش گسترده‌تر، استفاده از انواع دستگاه‌های الکتروکاردیوگرام، مقایسه اثربخشی مراقبت و درمان قبل از اجرای داده‌کاوی و بعد از اجرای آن در مراکز

منابع:

1. Zandkarim Ei, Afshari Safavi A. Comparison of artificial neural network predictive power with multiple logistic regressions to determine patients with and without diabetic retinopathy. Razi J Med Sci. 2014; 21(124): 79-90.
2. Roglic G, Unwin N, Bennett PH, Mathers C, Tuomilehto J, Nag S, et al. The burden of mortality attributable to diabetes. Diabetes Care. 2005; 28(9): 2130-5.
3. Ameri H, Alizadeh S, Barzegari A. Knowledge extraction of diabetics' data by decision tree method. J Ment Health Adm. 2013; 16(53): 58-72.

4. American Diabetes Association. Diabetes basics. Last Reviewed: May 18, 2015. Available from: <http://www.diabetes.org/diabetes-basics/statistics>.
5. Aram Ahmadi M, Bahrampour A. Comparison of logistic regression and discriminant analysis in predicting type 2 diabetes. *Iran J Epidemiol*. 2015; 11(3): 62-9.
6. Barfei F, Salehi M, Najafi I. Predicting diabetes using artificial neural network. *Razi J Med Sci*. 2015; 22(135): 28-37.
7. Khosravianian A, Ayat SS. Presenting an intelligent system for diagnosis of coronary heart disease by using probabilistic neural network. *Health Inf Manage*. 2015; 12(1): 3-13.
8. Dunne RA. A statistical approach to neural networks for pattern recognition: John Wiley and Sons; 2007.
9. Livingstone DJ. Artificial neural networks: Methods and applications (methods in molecular biology). New York: Humana Press; 2008.
10. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: A comparison of three data mining methods. *Artif Int Med*. 2005; 34(2): 113-27.
11. Zini G, d'Onofrio G. Neural network in hematopoietic malignancies. *Clinica Chimica Acta*. 2003; 333(2): 195-201.
12. Sun G, Dong X, Xu G. Tumor tissue identification based on gene expression data using DWT feature extraction and PNN classifier. *Neurocomputing*. 2006; 69(4): 387-402.
13. El-Sappagh S, Elmogy M, Riad A. A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis. *Artif Int Med*. 2015; 65(3): 179-208.
14. Baron-Epel O, Heymann AD, Friedman N, Kaplan G. Development of an unsupportive social interaction scale for patients with diabetes. *Patient Prefer Adherence*. 2015; 9: 1033-41.
15. Uguz H. A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases. *J Med Syst*. 2012; 36(1): 61-72.
16. Chien C-W, Lee Y-C, Ma T, Lee T-S, Lin Y-C, Wang W, et al. The application of artificial neural networks and decision tree model in predicting post-operative complication for gastric cancer patients. *Hepato-gastroenterology*. 2008; 55(84): 1140-5.
17. Andaieshgar B, Sedehi M, Kheiri S, Farahani nia M. Comparison of classical discriminant methods with artificial neural network using different algorithm to the diagnosis of myocardial infarction. *J Health Syst Res*. 2015; 11(2): 349-59.
18. Adavi M, Salehi M, Roudbari M, Asgari F, Rafei A. The comparison of the predictive precision of artificial neural networks and bivariate logistic regression in diagnosis of patients with hypertension. *Razi J Med Sci*. 2014; 21(123): 54-61.
19. Nafisi V, Hashemi Golpayegani S. Intelligent insulin injection system for the diabetic patients based on neural network and fuzzy algorithm. *Iran J Endocrinol Metabol*. 2003; 5(3): 203-10.
20. Dey N, Prasad Dash T, Dash S. ECG signal denoising by Functional Link Artificial Neural Network (FLANN). *Int J Biomed Engin Technol*. 2011; 7(4): 377-89.
21. Yanowitz FG. Introduction to ECG interpretation. LDS Hospital and Intermountain Medical Center. 2012.
22. Pongpon Sri S, Yu X-H. An adaptive filtering approach for Electrocardiogram (ECG) signal noise reduction using neural networks. *Neurocomputing*. 2013; 117: 206-13.
23. Alexakis C, Nyongesa H, Saatchi R, Harris N, Davies C, Emery C, et al., editors. Feature extraction and classification of Electrocardiogram (ECG) signals related to hypoglycaemia. *Comput Cardiol*. 2003; 4(23): 537-40.
24. Han J, Kamber M. Data mining: Concepts and techniques. Chapter 9: Graph mining, social network analysis and multirelational data mining. 2nd ed. United States: Morgan Kaufmann Pub. 2006: 535-89.

25. Mikut R, Reischl M. Data mining tools. Wiley Interdiscip Rev Data Min Knowl Discov. 2011; 1(5): 431-43.
26. Al-Shayea QK. Artificial neural networks in medical diagnosis. Int J Comput Sci Issues. 2011; 8(2):150-4.
27. Mahmoodi SA. Assessment of classification algorithms in the diagnosis of diabetes and breast cancer. Iran J Med Inform. 2013; 2(2). 19-22.
28. Bani-Hasan MA, Kadah YM, Rasmy ME, El-Hefnawi FM. Electrocardiogram signals identification for cardiac arrhythmias using prony's method and neural network. Conf Proc IEEE Eng Med Biol Soc. 2009; 2009: 1893-6.
29. Mateo J, Rieta JJ. Application of artificial neural networks for versatile preprocessing of electrocardiogram recordings. J Med Eng Technol. 2012; 36(2): 90-101.
30. Cho BH, Yu H, Kim KW, Kim TH, Kim IY, Kim SI. Application of irregular and unbalanced data to predict diabetic nephropathy using visualization and feature selection methods. Artif Intell Med. 2008; 42(1): 37-53.
31. Kazemnejad A, Batvandi Z, Faradmali J. Comparison of artificial neural network and binary logistic regression for determination of impaired glucose tolerance/diabetes. East Mediterr Health J. 2010; 16(6): 615-20.
32. Fiuzy M, Ghara-Khani A, Hadadnia J. Providing an intelligent hybrid system for detecting diabetes. 20th Iran Conf Electr Eng. Iran; Tehran University; 2012.

Diagnosis diabetes on the basis of information extracted from the ECG signal using artificial neural networks

Nazari M^{1,2}, Zamani Dehkordi B^{3*}, Kiomarsi Dehkordi F³

¹Student, Computer Dept., Shahrekord Branch, Islamic Azad University, Shahrekord, I.R. Iran;

²Shahrekord University of Medical Sciences, Shahrekord, I.R. Iran; ³Computer Dept., Shahrekord Branch, Islamic Azad University, Shahrekord, I.R. Iran.

Received: 19/Sep/2016

Accepted: 19/Dec/2016

Background and aims: Diabetes is known to be one of the most common diseases worldwide. Lack of an early and appropriate diagnosis is considered to be a main problem associated with diabetes. The aim of this study was to offer a novel approach to diagnose diabetes and, for the first time, investigates the correlation between ECG and diagnosis of diabetes using artificial neural network and data analysis algorithms.

Methods: In this study, 8 patients with diabetes and 64 healthy subjects were enrolled. ECG was conducted on all the participants. The necessary data including age, HR, p, t, RR, PP, P, PR, qt, and qtcb were drawn from ECG and collected in database. To classify the patients, tentative neural networks and standard algorithms were used. The data were analyzed using data analysis algorithms and different approaches, and the results of each investigation were compared with reference to appropriate rate. Weka software was used for ranking.

Results: The accuracy of detection of regulations-based algorithms and neural network, with better results in diabetes diagnosis, was higher than that of decision tree and interval-based algorithms. The best qualification rate (0.89) was obtained for ConsistencySubset Eval and QRS wave was reported the best choice in all algorithms. Investigating the data on people with and without diabetes using tentative neural networks showed an appropriate rate of 95%. Furthermore, KNN algorithm displayed the lowest time complexity.

Conclusion: Regulations-based model displayed the highest accuracy compared with all classification algorithms for data analysis used in the study.

Keywords: Diabetes mellitus, Artificial intelligence, Neural networks, Electrocardiography, Prediction models.

Cite this article as: Nazari M, Zamani Dehkordi B, Kiomarsi Dehkordi F. Diagnosis diabetes on the basis of information extracted from the ECG signal using artificial neural networks. J Shahrekord Univ Med Sci. 2017; 19(4): 64-77.

***Corresponding author:**

Computer Dept., Shahrekord Branch, Islamic Azad University, Shahrekord, I.R. Iran.
Tel: 00983833350081, E-mail: bzamani@gmail.com