



یافتن الگوریتم بهینه جهت پیش بینی و تشخیص بیماری دیابت با استفاده از روش های جمعی

علی اخلاقی^۱، هاله همایونی^۲، دکتر منصور امینی لاری^۳

^۱ دانشجوی کارشناسی ارشد مهندسی کامپیوتر موسسه آموزش عالی آپادانا شیراز، ایران ، iakhlaghi65@gmail.com

^۲ عضو هیات علمی ، موسسه غیرانتفاعی آپادانا شیراز، ایران ، Haleh.homayouni@gmail.com

^۳ استاد یار و عضو هیئت علمی دانشگاه آزاد اسلامی واحد مرودشت ، Aminilari@gmail.com

چکیده

این نوع لوزالمعده شخص مبتلا به دیابت قادر به ترشح انسولین نمی باشد و یا نوع غیر وابسته به انسولین (NIDDM) که در آن لوزالمعده شخص مبتلا به دیابت قادر به تولید و ترشح انسولین می باشد اما میزان جذب آن در بدن بسیار اندک است. مشکل عمده ای که در رابطه با بیماری دیابت وجود دارد عدم تشخیص به موقع و یا به طور کلی ضعف در تشخیص این بیماری است که این ضعف نیز به دلیل عدم انتخاب الگوی مناسب توسط پزشک و یا عدم استفاده مناسب از الگوهای استاندارد است. بنابراین پیاده سازی روشی که بتواند هر فرد را در تشخیص صحیح ابتلا یا عدم ابتلا به این بیماری یاری رساند می تواند گام مهمی در جهت پیشگیری و کنترل این بیماری به خصوص در مراحل ابتدایی آن باشد. تاکنون روشهای هوشمند گوناگونی به منظور حل این مشکل بنیادی در سراسر جهان ارائه گردیده که از آن جمله میتوان به استفاده از روشهای تکاملی و نیز استفاده از روش الگوریتم های فازی تشخیص الگو در استخراج ویژگی و همچنین استفاده از روش شبکه های عصبی اشاره نمود. در این مقاله، مسئله پیش بینی و تشخیص بیماری دیابت با استفاده از روش پیشنهادی (روش های جمعی) مورد توجه قرار گرفته است.

الگوریتم هایی پر کاربرد یادگیری سنتی

الگوریتم ماشین بردار پشتیبان

این الگوریتم از روشهای یادگیری با ناظر است که از آن برای طبقه بندی و رگرسیون استفاده می کنند. این روش از جمله روشهای نسبتاً جدیدی است که در سالهای اخیر کارایی خوبی نسبت به روشهای قدیمی تر برای طبقه بندی از جمله شبکه های عصبی پرسپترون نشان داده است. مبنای کاری دسته بندی کننده SVM دسته بندی خطی داده هاست و در تقسیم خطی داده ها سعی بر آن است خطی انتخاب شود که حاشیه اطمینان بیشتری داشته باشد. این الگوریتم دارای مبانی نظری قوی و بی نقصی می باشد، فقط به یک دوجین نمونه احتیاج دارد و به تعداد ابعاد مسأله حساس نمی باشد. در یک فرایند یادگیری که شامل دو کلاس می باشد، هدف SVM پیدا کردن بهترین تابع برای طبقه بندی می باشد به نحوی که بتوان اعضای دو کلاس را در مجموعه داده ها از هم تشخیص داد.

یکی از مشکلات SVM پیچیدگی محاسباتی آن است. با اینحال این مشکل به طور قابل قبولی حل شده است. یک راه حل این است

امروزه در دانش پزشکی جمع آوری داده های فراوان در مورد بیماری های مختلف از اهمیت فراوانی برخوردار است. از سوی دیگر حجم مفید اطلاعات پزشکی که حتی با یک محدوده کوچک تشخیصی در ارتباط است، به قدری زیاد می باشد که تصمیم گیری سریع و دقیق را برای پزشک دشوار خواهد ساخت و همچنین فناوریهای مدرن روز به روز بر حجم این اطلاعات می افزایند و مساله را دشوارتر از پیش خواهند ساخت؛ به صورتیکه پزشکان تاکنون با چنین حجم وسیعی از اطلاعات مواجه نشده بودند. مراکز پزشکی با مقاصد گوناگونی به جمع آوری این داده ها می پردازند. تحقیق روی این داده ها و به دست آوردن نتایج و الگوهای مفید در رابطه با بیماری ها، یکی از اهداف استفاده از این داده ها است. حجم زیاد این داده ها و سردرگمی حاصل از آن مشکلی است که مانع رسیدن به نتایج قابل توجه می شود. این متد پیشنهادی می تواند در پیش بینی و تشخیص بیماری دیابت کارساز باشد. از آنجایی که در کارهای صورت گرفته پیشین عموماً از روشهای شبکه عصبی مصنوعی، درخت تصمیم (درخت رگرسیون)، SVM، و ... استفاده شده است لذا در این مقاله داده های مربوط به بیماری دیابت توسط نگارنده جمع آوری شده و روشهای پیشنهادی که بیشتر در قالب متدهای Ensemble هستند بر روی این داده ها (دیتاست) با استفاده از نرم افزار رپیدماینر به مرحله اجرا، خروجی و در نهایت مقایسه بین روشها خواهیم پرداخت.

واژه های کلیدی

الگوریتم بهینه، طبقه بندی، پیش بینی و تشخیص بیماری، الگوریتم یادگیری جمعی، دیابت

مقدمه

بیماری دیابت یکی از شایعترین بیماریهای حاضر دنیا شناخته شده است که علیرغم گستردگی شیوع این بیماری هنوز روشی به منظور ریشه کن کردن و از بین بردن آن در دنیا شناخته نشده است هرچند که روشهای مختلفی جهت تشخیص و کنترل آن در حال حاضر مورد استفاده قرار میگیرد. از جمله عوارضی که به دنبال مبتلا شدن افراد به این بیماری گریبانگیر آنها خواهد شد میتوان به گرفتگی عروق قلبی و در نوع پیشرفته آن به نابینایی، قطع اعضای بدن، اختلالات فکری و غیره اشاره نمود. بیماری دیابت را از نظر تقسیم بندی میتوان به دو نوع وابسته به انسولین (IDDM) که در