Archive of SID

Comparison of Artificial Neural Network and Decision Tree to Identify and Predict Factors Associated with Type 2 Diabetes

Mirzakhani F[1], Kazemi A[2], Rasoulian Kasrineh M[3], Javad Moosavi S.Y[4], Amirabadizade A.R[5]

Abstract

Purpose: One of the goals of medical research is to determine the factors association of diseases in prognosis. One of the most common metabolic diseases in Iran is diabetes. The aim of this study was to identify the related factors that predict diabetes by using artificial neural network and decision tree algorithms. In this study we will compare the performance of these models.

Methods: In this study, 901 cases of people referred to health centers in Mashhad were used. Initially, data were analyzed using descriptive and analytical statistics. Then, 70% of the data were randomly selected for constructing artificial neural network and decision tree models and the remaining 30% were used to compare the performance of the models. Finally, the performance of the models was compared using the ROC curve.

Results: Development of two predictive models was performed by using13 input (independent) variables and 1 output (dependent) variable. The two models were evaluated in terms of area under the ROC curve, sensitivity, specificity and accuracy. Area under ROC curve, sensitivity, specificity and accuracy for artificial neural network model were 69.1, 74.2, 56.03 and 61.3. For CART algorithm of decision tree the under ROC curve, sensitivity, specificity and accuracy were obtained as 68.9, 64.77, 63.47 and 65.3 respectively. In all modes, family history of diabetes, triglycerides, body mass index, low density lipoprotein, and systolic and diastolic blood pressure were the most important factors associated with type 2 diabetes.

Conclusion: The results showed that the perceptron multi-layer neural network model had a better result than the CART decision tree in term of area under the ROC curve for prediction of diabetes type 2. Also, low density lipoprotein was identified as the most important related factor of type 2 diabetes. The study suggests that modern data mining techniques such as artificial neural network and decision trees can be used to identify associated disease factors.

Keywords: Decision tree, Artificial neural network, Data mining, Diabetes

Received: 2017.09.12; Accepted: 2018.04.21

مقایسه مدل شبکه عصبی مصنوعی و درخت تصمیم برای شناسایی و پیش بینی عوامل مرتبط با دیابت نوع۲

فرزاد میرزاخانی[۱]، آذر کاظمی[۲]، مرجان رسولیان کسرینه[۳]، سید یوسف جواد موسوی[۴]، علیرضا امیرآبادی‌زاده[۵]

هدف: یکی از اهداف تحقیقات پزشکی تعیین عوامل مرتبط در پیش بینی بیماری می باشد. یکی از شایع ترین بیماری های متابولیک در ایران، دیابت می‌باشد. هدف از این مطالعه شناسایی عوامل موثر در پیش بینی دیابت با استفاده از مدل های شبکه عصبی مصنوعی و درخت تصمیم می‌باشد.

روش بررسی: برای انجام مطالعه، پرونده ۹۰۱ تن از افرادی که در سال های ۹۱ و ۹۲ به مراکز بهداشتی شهر مشهد مراجعه کرده بودند، استفاده گردید. در ابتدا با استفاده از روش های آمار توصیفی و تحلیلی، داده ها آنالیز شدند. سپس ۷۰٪ داده ها به طور تصادفی برای ساخت مدل های شبکه عصبی مصنوعی و درخت تصمیم انتخاب شدند. ۳۰٪ باقیمانده برای مقایسه عملکرد مدل ها استفاده شد. در نهایت عملکرد مدل ها با استفاده از سطح زیر منحنی راک (ROC) مورد مقایسه قرار گرفت.

یافته ها: توسعه دو مدل پیش بینی با استفاده از ۱۴ متغیر انجام شد. دو مدل از نظر سطح زیر منحنی راک، حساسیت، ویژگی و صحت مورد ارزیابی قرار گرفتند. برای مدل شبکه عصبی، سطح زیر منحنی راک و حساسیت به ترتیب ۶۹/۱ و ۷۴/۲

www.SID.ir

Archive of SID

بدست آمد. برای مدل درخت تصمیم نیز سطح زیر منحنی راک و حساسیت به ترتیب ۶۸/۹ و ۶۴/۷۷ بدست آمد. در هر دو مدل متغیرهای سابقه خانوادگی دیابت، تری گلیسرید، شاخص توده بدنی، لیپوپروتئین با چگالی کم و فشار خون سیستولیک و دیاستولیک مهم ترین عوامل مرتبط در شناسایی دیابت نوع ۲ بودند.

نتیجه گیری: نتایج نشان داد که مدل شبکه عصبی چند لایه سطح زیر منحنی راک بهتری نسبت به درخت تصمیم CART در پیش بینی دیابت نوع ۲ دارد. همچنین لیپوپروتئین با چگالی کم مهم ترین عوامل مرتبط در شناسایی دیابت نوع ۲ می باشد. مطالعه حاکی از آنست که روش های داده‌کاوی نوین از جمله شبکه عصبی مصنوعی و درخت تصمیم می توانند برای شناسایی عوامل مرتبط با بیماری ها مورد استفاده قرار گیرند.

کلمات کلیدی: درخت تصمیم، شبکه عصبی مصنوعی، داده‌کاوی، دیابت

نویسنده مسئول: علیرضا امیرآبادی زاده، amirabadiza921@gmail.com، ORCID: 0000-0002-2495-5042

آدرس: مرکز تحقیقات مسمومیت ها و سو مصرف مواد، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران

۱- دانشجوی دکترای تخصصی انفورماتیک پزشکی، گروه مدیریت اطلاعات سلامت، دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی ایران، تهران، ایران

۲- دانشجوی دکترای تخصصی انفورماتیک پزشکی،گروه انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۳- دانشجوی کارشناسی ارشد گروه انفورماتیک پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۴- دانشجوی پزشکی عمومی، کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران

۵- کارشناس مرکز تحقیقات مسمومیت ها و سو مصرف مواد، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران

مقدمه

دیابت یک اختلال درون ریز است و توسط هایپر گلیسمی مزمن که نتیجه نقص در تولید انسولین یا مقاومت به آن است مشخص و شناخته می شود. دیابت شایع ترین بیماری متابولیک در جهان است و در ایران نیز از شیوع بالایی برخوردار می باشد (۱). دیابت یک بیماری مزمن است که هزینه‌های زیادی برای خدمات پزشکی در بردارد (۲). دیابت چهارمین یا پنجمین علت مرگ و میر در اغلب کشورهای با درآمد سرانه بالا بوده و شواهد زیادی از همه-گیری دیابت در کشورهای صنعتی یا در حال صنعتی شدن نیز وجود دارد. میزان شیوع کلی این بیماری به ویژه در کشورهای درحال توسعه رو به افزایش است (۱). به علت شیوع قابل توجه چاقی و به موازات آن شیوع دیابت قندی در بیشتر کشورهای دنیا، «چاقی» و «دیابت» اپیدمی دو قلوی قرن ۲۱ نام گرفته اند (۳). مطمئناً یکی از مشکلات چالش برانگیز در قرن ۲۱ گستردگی زیاد این بیماری در کشورهای در حال توسعه است. تقریباً حدود ۳۹۰ میلیون نفر درسرتاسر جهان تا سال ۲۰۱۵ دارای دیابت شیرین هستند و پیش بینی می شود که تعداد مبتلایان به این بیماری به حدود ۵۹۲ میلیون نفر تا سال ۲۰۳۵ برسد (۴). براساس نقطه نظرات کارشناسان سازمان جهانی بهداشت، شیوع دیابت نوع ۲ در ایران تا سال ۲۰۲۵ حدود۵۲۱۵۰۰۰ نفر پیش بینی شده است و این عدد به ۶/۴ میلیون نفر در سال ۲۰۳۰ می رسد (۵). براساس برآورد وزارت بهداشت، درمان و آموزش پزشکی حدود ۱/۱ میلیارد دلار از کل بودجه ۱۳ میلیاردی این وزارتخانه صرف درمان عوارض ۲۲۰۵۰۰۰ نفر بیمار دیابتی شناخته شده در کشور می شود (۶). از آنجا که بیماری های قلبی، سکته مغزی، نابینایی، بیماری های کلیوی و قطع عضو از جدی‌ترین عوارض دیابت می باشند، بنابراین شناسایی و تشخیص بیماران دیابت نوع ۲ برای هر دو سیاست بهداشت عمومی و بالینی بسیار مهم است (۷).

با توجه به شیوع دیابت نوع ۲ در سراسرجهان، روش - های جدید در تحقیقات زیست-پزشکی بسیار مورد توجه قرار گرفته است، که از آن جمله می‌توان به تکنیک های داده‌کاوی اشاره نمود (۸، ۶). داده‌کاوی یکی از روش هایی است که می تواند ارتباطات و وابستگی های جدید و بدیعی را کشف کند که برای پزشکان مفید هستند (۶). استخراج اطلاعات از بین حجم انبوه اطلاعات با استفاده از فرآیند داده کاوی ما را در شناسایی و پیش بینی بیماری کمک می کند (۸). داده کاوی در حوزه سلامت

www.SID.ir

Archive of SID

کاربردهای فراوانی دارد که از جمله آن ها می‌توان به موارد ذیل اشاره کرد: تشخیص بیماری ها، دسته بندی بیماران در مدیریت بیماری، پیدا کردن الگوهایی برای تشخیص سریعتر بیماران و جلوگیری از بروز عوارض در آن ها (۹).

در داده‌کاوی از روش های مختلف آماری، یادگیری ماشین، هوش مصنوعی و غیره برای کشف روابط پنهان بین داده ها استفاده می شود. یکی از روش‌هایی که به طور گسترده در داده‌کاوی مورد استفاده قرار می گیرد، شبکه عصبی مصنوعی می باشد که با استفاده از آن به بررسی ارتباطات غیرخطی و اثرات متقابل پیچیده بین عوامل مختلف پرداخته می‌شود (۱۱، ۱۰). این مدل، شناسایی ارتباط بین عوامل مختلف، که برخی از تکنیک-های آماری مانند رگرسیون چندگانه[1] ممکن است قادر به تشخیص آن ها نباشد، را ممکن می سازد (۱۲). مدل شبکه عصبی مصنوعی یکی از مدل های مورد استفاده در حوزه هوش مصنوعی، یادگیری ماشین، داده‌کاوی[2] و... است که با الهام گرفتن از شبکه نورون‌ها در مغز انسان ایجاد شده است (۱۴، ۱۳). یکی از شبکه های عصبی پرکاربرد شبکه عصبی چند لایه پرسپترون با روش یادگیری پس انتشار می باشد که قادر است هر سیستم غیرخطی را مدل سازی نماید (۱۵). یکی دیگر از روش-های رایج مورد استفاده در داده‌کاوی، درخت تصمیم است (۱۷، ۱۶، ۶). درخت تصمیم یکی از روش های ناپارامتری رده بندی کردن است (۱۷). این روش با به کارگیری تکنیک های بسیار ساده یک الگوی رده بندی برای مشاهدات موجود معرفی می‌نماید (۱۷). الگوی معرفی شده توسط این روش از ساختار بسیار ساده و قابل درک برای تصمیم گیری برخوردار می باشد. با اینکه این روش از تکنیک های بسیار ساده استفاده می نماید ولی در زمینه تشخیص و پیشگویی می تواند به خوبی روش-های پیچیده نظیر شبکه عصبی مصنوعی عمل نماید (۱۷، ۱۶). بنابراین هدف ما در این مطالعه شناسایی و پیش بینی عوامل مرتبط با دیابت نوع ۲ با استفاده از مدل درخت تصمیم و شبکه عصبی و مقایسه این دو مدل می باشد.

[1] Multiple regression
[2] Data mining

روش بررسی

داده های مورد استفاده در این مطالعه توصیفی- تحلیلی، از پرونده افرادی که در سال های ۹۱ و ۹۲ هجری شمسی جهت کنترل و پیشگیری از دیابت به پنج مراکز بهداشت شهر مشهد مراجعه کرده بودند، جمع آوری گردید. متغیرهای لازم از روی پرونده افراد و فرم‌های ارجاع آنها به مراکز قلب، چشم، کلیه و... از ۹۰۱ پرونده استخراج گردید. متغیرهای مستقل در این مطالعه براساس اطلاعات موجود در پرونده افراد شامل سن، جنسیت، مصرف سیگار، وضعیت تاهل، وضعیت شغل، سابقه خانوادگی دیابت، شاخص توده بدنی، فشار خون سیستولیک (SBP)[3] و دیاستولیک (DBP)[4]، تری گلیسرید[5] (TG)، کلسترول (TC)[6]، لیپوپروتئین با چگالی کم (LDL)[7] و بالا (HDL)[8] بودند (۱۶، ۶). متغیر قند خون با استفاده از نقطه برش ۱۲۶ میلی‌گرم بر دسی-لیتر به دو گروه تقسیم بندی شد. افراد با قند خون بیشتر مساوی ۱۲۶ میلی‌گرم بر دسی‌لیتر به عنوان افراد دیابتی و افراد با قند خون کمتر از ۱۲۶ میلی‌گرم بر دسی‌لیتر به عنوان افراد غیردیابتی به عنوان متغیر وابسته در نظر گرفته شدند (۱۸).

ابتدا برای ایجاد دو مدل، داده ها به صورت تصادفی به دو قسمت داده های آموزش و داده های تست تقسیم-بندی شد. ۷۰ درصد داده ها (۶۴۳ نفر) به عنوان داده-های آموزش در نظر گرفته شد و از آن ها برای ساخت مدل درخت تصمیم و شبکه عصبی استفاده گردید و ۳۰ درصد باقیمانده (۲۵۸ نفر) نیز به عنوان داده های تست و برای ارزیابی نتایج حاصل از مدل سازی به کار گرفته شد. برای طراحی مدل شبکه عصبی و پیدا کردن مدل بهینه، از روش آزمون و خطا استفاده گردید. برای انجام این کار در ابتدا نرخ یادگیری، ۰/۰۵و تعداد نورون‌ها، ۲ در نظر گرفته شد و از تابع فعالیت سیگموئید[9] در لایه پنهان و تابع فعالیت خطی در لایه خروجی استفاده گردید. به تدریج نرخ یادگیری تا ۰/۵ و تعداد نورون‌ها تا ۱۸ عدد

[3] Systolic blood pressure
[4] Diastolic blood pressure
[5] Triglyceride
[6] Cholesterol
[7] LDL
[8] HDL
[9] Sigmoid

www.SID.ir

Archive of SID

افزایش یافت و ترکیب های مختلف نرخ یادگیری و تعداد نورون‌ها با یکدیگر در نظر شد و سپس براساس معیار خطای میانگین مربعات خطا(MSE)[10]، ترکیب بهینه جهت مدل سازی نهایی انتخاب گردید. برای طراحی مدل درخت تصمیم از الگوریتم CART[11] که از مشهورترین و پرکاربردترین الگوریتم های درخت تصمیم است، استفاده و از شاخص جینی به عنوان معیار تجزیه استفاده گردید. لازم به ذکر است که الگوریتم CART، متغیرهای ورودی را برای یافتن بهترین تجزیه می‌آزماید تا شاخص ناخالصی حاصل از تجزیه، حداقل مقدار ممکن را اختیار کند.

برای ارزیابی کیفیت مدل های شبکه عصبی و درخت تصمیم از معیارهای حساسیت، ویژگی، صحت و سطح زیر منحنی راک[12] استفاده گردید. منحنی راک یک تکنیک برای تجسم، سازماندهی و انتخاب طبقه بندی براساس عملکرد آن ها می باشد. علت این که به آن نمودار مشخصه عملکرد نسبی می گویند این است که مقایسه بین دو مشخصه عملکرد، یعنی مثبت حقیقی و مثبت کاذب در کران های مختلف تصمیم گیری انجام می دهد. در فضای راک، حساسیت محور عمودی و یک منهای ویژگی محور افقی را تشکیل می دهد (۱۹). حساسیت[13] و ویژگی[14]در آمار دو شاخص برای ارزیابی نتیجه یک آزمایش دسته بندی هستند. حساسیت به معنی نسبتی از افراد بیمار است که آزمایش آن ها به درستی بیمار تشخیص داده است. بنابراین هرچه این مقدار بزرگتر باشد، نشان دهنده ی دقت در تشخیص افراد بیمار است. همچنین ویژگی به معنی نسبتی از افراد سالم است که مدل به درستی آن ها را سالم تشخیص داده است. منظور از معیار صحت یا نرخ دسته بندی نسبت تعداد افراد سالم و بیمار است که به درستی تشخیص داده می شوند (۶).

برای متغیرهای کیفی با استفاده از آزمون کای دو به مقایسه بین دو گروه پرداخته شد و برای مقایسه متغیرهای کمی نیز، پس از بررسی نرمال بودن با استفاده از آزمون کلوموگروف اسمیرنوف توزیع[15] داده ها از آزمون تی مستقل برای مقایسه دو گروه استفاده گردید. متغیرهای که در آزمون تک متغیره سطح معناداری آن ها کمتر از ۰/۲ بود وارد مدل رگرسیون لجستیک[16] چند گانه شدند. تمامی آنالیزها و برازش مدل ها با استفاده از نرم افزارهای۱۹ SPSS و R انجام شد.

یافته ها

از ۹۰۱ فرد مورد بررسی ۵۷/۹ درصد زن بودند. تعداد ۱۶۰ نفر دیابتی (قند خون بیشتر مساوی ۱۲۶ میلی‌گرم بر دسی‌لیتر) و ۷۴۱ نفر نیز غیر دیابتی (قند خون کمتر از ۱۲۶ میلی‌گرم بر دسی‌لیتر) بودند. همچنین در افراد دیابتی میانگین شاخص توده بدنی، فشار خون سیستولیک[17] و دیاستولیک[18] ، تری گلیسرید[19]، کلسترول[20] و لیپوپروتئین با چگالی کم[21] به طور معنا داری بالاتر از افراد غیر دیابتی بود (جدول۱).

نتایج آزمون رگرسیون لجستیک چند گانه نشان داد که متغیر های سن، سابقه خانوادگی دیابت، شاخص توده بدنی، فشار خون سیستولیک و دیاستولیک، کلسترول و لیپوپروتئین با چگالی کم معنادار[22] بودند ($p<۰/۰۵$). نتایج نشان داد که به ازای یک سال افزایش سن شانس ابتلا به دیابت ۱/۰۴ برابر افزایش پیدا می کند. همچنین در افرادی که سابقه خانوادگی دیابت دارند شانس ابتلا به دیابت نوع۲، ۲/۵۵ برابر افزایش پیدا می کند (پیوست۱). پس از در نظر گرفتن ترکیب های مختلف نرخ های یادگیری و تعداد نورون ها برای ساخت شبکه عصبی مصنوعی، در نهایت یک شبکه عصبی پرسپترون سه لایه با ۲۱ گره در لایه ورودی، ۱۸ گره در لایه پنهان و ۲ گره در لایه خروجی، با نرخ یادگیری ۰/۲ استفاده شد که در بین مدل‌های دیگر، کمترین میانگین مربعات خطا را داشت (جدول۲).

نتایج آنالیز حساسیت و اهمیت نسبی متغیرها در مدل شبکه عصبی نشان داد که متغیرهای سن، فشار خون سیستولیک، لیپوپروتئین با چگالی کم، شاخص توده بدنی

[10] Mean squared error
[11] Classification And Regression Tree
[12] ROC Curve
[13] Sensitivity
[14] Specificity
[15] Distribution
[16] Logistic regression
[17] Systolic blood pressure
[18] Diastolic blood pressure
[19] Triglyceride
[20] Cholesterol
[21] LDL
[22] Significant

www.SID.ir

Archive of SID

جدول۱: مقایسه متغیرهای دموگرافیک و آزمایشگاهی بین گروه های دیابتی و غیر دیابتی

متغیر		کل	دیابتی	غیر دیابتی	آماره آزمون	p- مقدار
سن (سال)		۴۸/۳۵±۸/۲۱	۵۱/۴±۷/۸۱	۴۷/۵±۸/۱۲	t=۱۷/۷۹	<۰/۰۰۱
جنسیت	مرد	۳۷۹(٪۴۲)	۶۱(٪۳۸/۱)	†۳۱۹(٪۴۳)	χ^2=۰/۸۷	۰/۳۵
	زن	۵۲۱(٪۵۸)	۹۹(٪۶۱/۹)	۴۲۲(٪۵۷)		
وضعیت تاهل	مجرد	۴(٪۰/۴)	۲(٪۱/۳)	۲(٪۱/۳)	χ^2=۴۶/۶۵	<۰/۰۰۱
	متاهل	۸۵۴(٪۹۴/۹)	۱۴۸(٪۹۲/۵)	۷۰۷(٪۹۵/۴)		
	مطلقه	۱۰(٪۱/۱)	۲(٪۱/۳)	۸(٪۱/۱)		
	بیوه	۳۲(٪۳/۶)	۸(٪۵)	۲۴(٪۳/۲)		
وضعیت شغلی	دانشجو	۷(٪۰/۸)	۳(٪۱/۹)	۴(٪۰/۵)	χ^2=۴۷/۹۲	<۰/۰۰۱
	کارمند	۴۶۲(٪۵۱/۳)	۷۶(٪۴۷/۵)	۳۸۷(٪۵۲/۲)		
	بیکار	۲۹۷(٪۳۳)	۵۳(٪۳۳/۱)	۲۴۴(٪۳۳)		
	بازنشسته	۱۳۴(٪۱۴/۹)	۲۸(٪۱۷/۵)	۱۰۶(٪۱۴/۳)		
مصرف سیگار	دارد	۱۵۲(٪۱۶/۹)	۲۰(٪۱۲/۵)	۱۳۲(٪۱۷/۸)	χ^2=۲/۸۴	۰/۰۴
	ندارد	۷۴۸(۸۳/۱٪)	۱۴۰(۸۷/۵٪)	۶۰۹(۸۲/۲٪)		
سابقه خانوادگی دیابت	بله	۲۸۰(٪۳۱/۱)	۷۳(٪۴۵/۶)	۲۰۷(٪۲۸)	χ^2=۲۶۶/۲۴	<۰/۰۰۱
	خیر	۶۲۰(٪۶۸/۹)	۸۷(٪۵۴/۴)	۵۳۳(٪۷۲)		
شاخص توده بدنی(کیلوگرم بر متر مربع)		#۲۸/۰۱±۴/۷۲	۲۹/۰۹±۴/۷	۲۷/۷±۴/۶۷	¥t=۱۰/۵۲	<۰/۰۰۱
فشارخون سیستولیک(میلی متر جیوه)		۱۲۲/۳۶±۱۹/۵۱	۱۲۸/۸۵±۲۹/۰۴	۱۲۰/۵۷±۱۸/۹۸	t=۱۴/۸۷	<۰/۰۰۱
فشارخون دیاستولیک(میلی متر جیوه)		۷۹/۳±۱۱/۹۹	۸۱/۷۵±۱۲/۸۴	۷۸/۶۳±۱۱/۶۶	t=۹/۳۳	<۰/۰۰۱
تری گلیسرید(میلی گرم بر دسی لیتر)		۱۴۶/۵±۹۶/۵	۱۸۰/۴±۱۱۸/۳	۱۳۷/۱±۸۷/۳	t=۱۳/۷۳	<۰/۰۰۱
کلسترول(میلی گرم بر دسی لیتر)		۱۹۲/۵۴±۳۹/۶	۲۰۲/۲±۴۵/۳	۱۸۹/۸۸±۳۷/۵	t=۱۰/۰۴	<۰/۰۰۱
لیپوپروتئین با چگالی کم(میلی گرم بر دسی لیتر)		۱۱۷/۰۶±۳۵/۶	۱۲۰/۱۲±۳۹/۷	۱۱۶/۲±۳۴/۴	t=۳/۶	<۰/۰۰۱
لیپوپروتئین با چگالی بالا(میلی گرم بر دسی لیتر)		۴۲/۷۷±۹/۹	۴۲/۴±۱۰/۳	۴۲/۸±۹/۸	t=۱/۴۷	۰/۱۴

#: میانگین ± انحراف معیار ، †: فراوانی (درصد)، ¥:آماره آزمون تی مستقل

جدول۲: میانگین مربعات خطا در مدل شبکه عصبی با توجه به تعداد گره‌های لایه مخفی و نرخ یادگیری

نرخ یادگیری	تعداد گره‌های لایه مخفی					
	۲	۷	۱۰	۱۲	۱۶	۱۸
۰/۰۵	۲۸/۵۱	۲۷/۷۴	۲۷/۵۶	۲۷/۵۲	۲۷/۱۲	۲۷/۰۴
۰/۰۱	۲۸/۴۶	۲۷/۸۰	۲۷/۴۴	۲۷/۵۶	۲۶/۹۴	۲۶/۸۴
۰/۲	۲۸/۳۸	۲۷/۸۷	۲۷/۵۴	۲۷/۳۹	۲۷/۲۳	۲۶/۸
۰/۵	۲۸/۳۷	۲۷/۸۶	۲۷/۵۱	۲۷/۳۸	۲۷/۳۷	۲۶/۹۴

www.SID.ir

Archive of SID

، سابقه خانوادگی دیابت، تری گلیسرید و فشار خون سیستولیک مهم‌ترین متغیرها در پیش‌بینی و شناسایی عوامل مرتبط دیابت نوع ۲ می باشند (شکل۱). برای طراحی مدل درخت تصمیم، از الگوریتم CART با معیار تجزیه جینی مورد استفاده قرار گرفت. نتایج آنالیز حساسیت و اهمیت نسبی متغیرها در مدل درخت تصمیم نشان داد که متغیرهای تری گلیسرید، فشار خون سیستولیک، لیپوپروتئین با چگالی کم، سابقه خانوادگی دیابت، فشار خون دیاستولیک و شاخص توده بدنی مهم-ترین متغیرها در پیش‌بینی و شناسایی عوامل مرتبط دیابت نوع ۲ می باشند (شکل۲). درخت تصمیم نهایی دارای ۲۷ گره، ۷ لایه و ۱۴برگ می باشد (شکل۳). متغیر اول (ریشه) در درخت، فشار خون سیستولیک می باشد که مهم ترین عامل می‌باشد.نتایج قوانین شرطی "اگر-آنگاه"بدست آمده از مدل درخت تصمیم در پیوست ۲ آورده شده است. مسیر ترسیم قوانین بدست آمده از گره ریشه به سمت گره برگ می‌باشد. نتایج نشان داد اگر فشار خون سیستولیک بیشتر مساوی۱۴۰ میلی‌متر جیوه، تری- گلیسرید بیشتر مساوی ۱۲۸ میلی‌گرم بر دسی‌لیتر، و فشار خون دیاستولیک بیشتر مساوی ۹۲ میلی‌متر جیوه باشد، آنگاه با احتمال ۸۴ درصد فرد دیابتی خواهد بود. در زیر گروه دیگری از درخت دیگر نشان داده شد که اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی‌متر جیوه، لیپوپروتئین با چگالی کم کمتر از ۶۱ میلی‌گرم بر دسی-لیتر و تری- گلیسرید کمتر از ۱۷۲ میلی‌گرم بر دسی‌لیتر باشد آنگاه با احتمال ۷۸ درصد فرد غیر دیابتی خواهد بود. با توجه به نتایج مقایسه عملکرد معیارهای دو مدل شبکه عصبی و درخت تصمیم معیارهای حساسیت و سطح زیر منحنی راک در مدل شبکه عصبی بیشتر از درخت تصمیم می باشد (جدول۳). در این مطالعه، دو مدل شبکه عصبی و درخت تصمیم برای پیش بینی عوامل مرتبط دیابت نوع ۲ استفاده گردید که حساسیت و ویژگی برای این دو به ترتیب (۷۴/۲ و۷ /۶۴) و (۵۶/۰۳ و ۶۳/۴۷) به دست آمد (شکل۴).

بحث و نتیجه گیری

دیابت از شایع ترین بیماری های متابولیک جهان و یکی از مشکلات چالش برانگیز قرن حاضر است(۴، ۱). دیابت نتیجه فعل و انفعالات بین استعداد ژنتیکی و رفتاری و عوامل زیست محیطی می باشد (۲۰). با توجه به شیوع دیابت نوع ۲ در سراسر جهان شناسایی ارتباطات و کشف قوانین جدید برای پزشکان مفید است (۸). در این مطالعه با استفاده از مدل شبکه عصبی مصنوعی و درخت تصمیم به پیش بینی عوامل مرتبط مرتبط با دیابت نوع ۲ پرداخته شد.

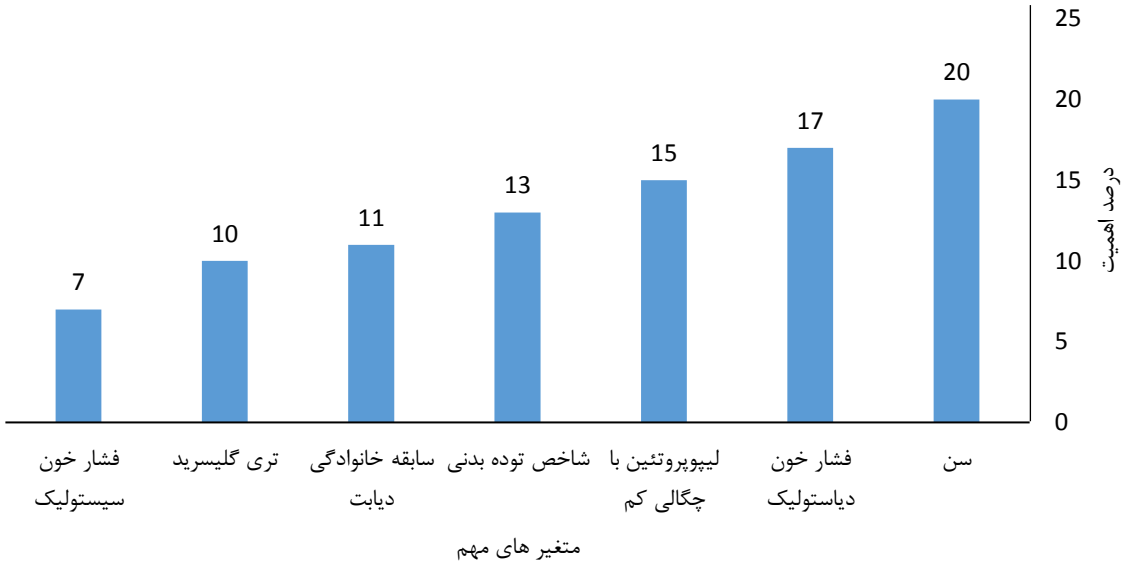
مرتبط ترین عوامل در دو مدل شبکه عصبی و درخت تصمیم سن بالا، تری گلیسرید، فشار خون سیستولیک و لیپوپروتئین با چگالی کم، شاخص توده بدنی و سابقه خانوادگی دیابت بود. نتایج مطالعات قبلی، عوامل مرتبط دیابت را شبیه نتایج حاصل از فاکتورهای مهم در مدل درخت تصمیم و شبکه عصبی مطالعه حاضرگزارش کرده اند (۲۳-۲۱، ۱۷). مطالعه Meng و همکاران نشان داد که در مدل شبکه عصبی و درخت تصمیم سن، سطح تحصیلات، سابقه خانوادگی دیابت، شاخص توده بدنی، جنسیت، فعالیت فیزیکی و وضعیت تاهل مهم‌ترین عوامل مرتبط با دیابت نوع۲ هستند (۱۶). دلیل تفاوت برخی متغیرها با نتایج مطالعه ما عدم استفاده از متغیرهای بیوشیمی در مطالعه Meng و همکاران است. نتایج مطالعات قبلی ریسک فاکتورهایی مانند سن، جنسیت، فشار خون سیستولیک و شاخص توده بدنی را برای دیابت گزارش شده اند که با نتیجه مطالعه حاضر همخوانی دارد (۲۵، ۲۴).

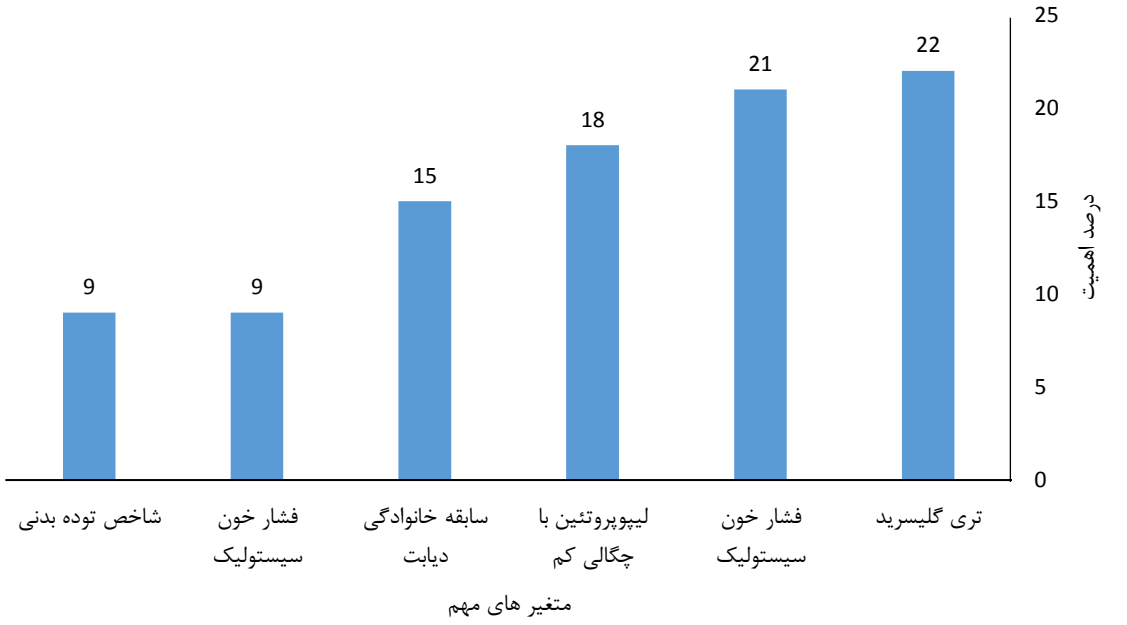
برای مقایسه مدل‌های پیش بینی معمولاً از دو معیار حساسیت و ویژگی استفاده می شود که این دو معیار تعیین می کنند که هر مدل تا چه حدی می تواند نتایج مثبت و منفی را با دقت تشخیص دهد و از این جنبه، این دو معیار برای مدل های پیش بینی اهمیت بسیاری دارند (۲۶)؛ و اعتبار مدلی بالا می باشد که هر دو مقدار حساسیت و ویژگی آن بالا باشد (۲۸، ۲۷). نتایج مطالعه حاضر از نظر میزان حساسیت و ویژگی مدل ها شبیه مطالعات قبلی بود (۳۰، ۲۹) اما با یافته های مطالعه کاظم نژاد و همکاران در تناقض بود (۳۱) که این تفاوت را می توان به تفاوت های جغرافیایی و رژیم غذایی مرتبط دانست.

در این مطالعه، علاوه بر محاسبه ویژگی و حساسیت، از نمودار راک نیز برای مقایسه عملکرد دو مدل استفاده شد. صحت و سطح زیر منحنی راک به ترتیب (۶۱/۳ و ۶۵/۳) و (۶۹/۱ و ۶۸/۹) در مدل شبکه عصبی و درخت

www.SID.ir

Archive of SID

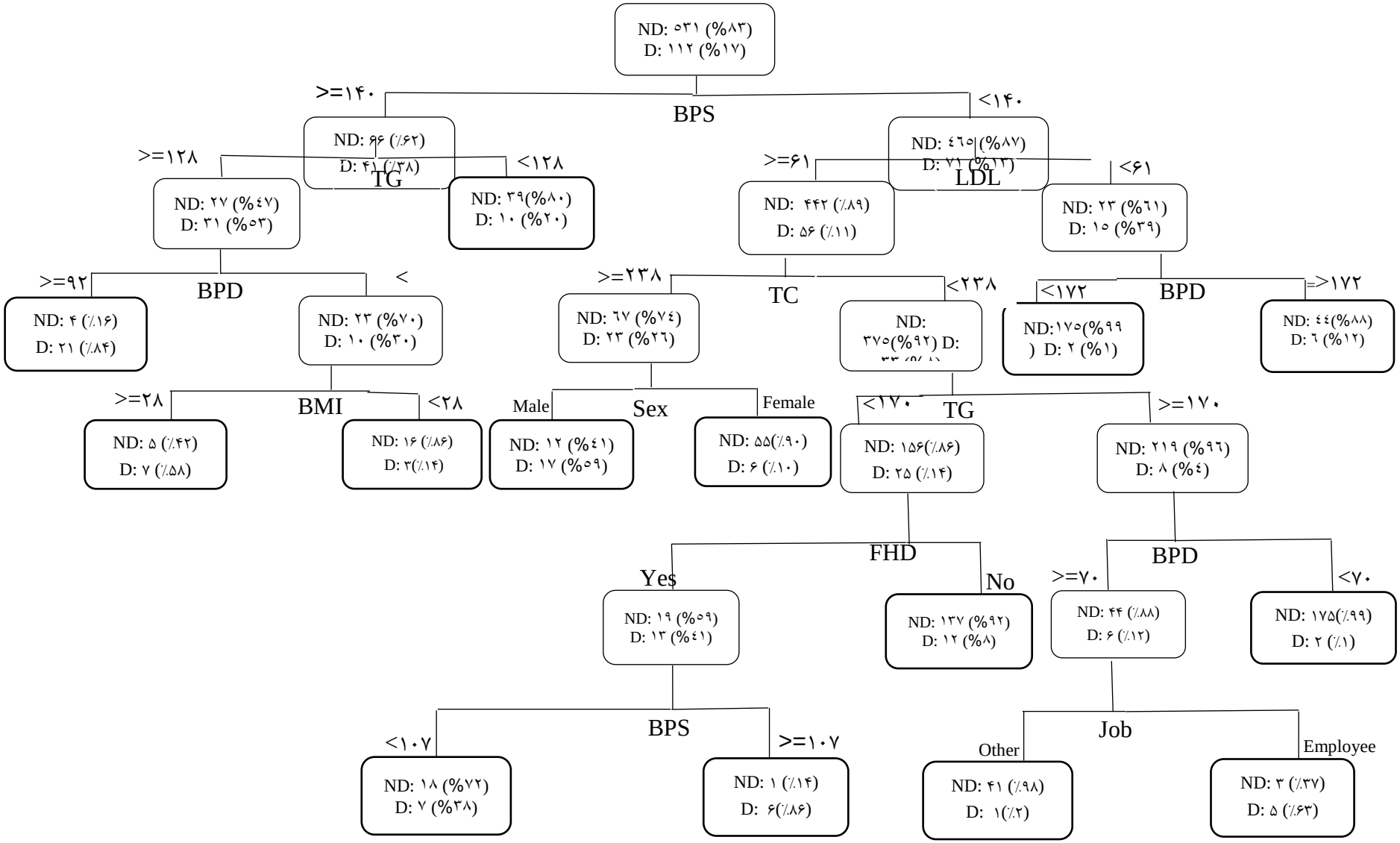




شکل۲: نتایج آنالیز حساسیت و اهمیت نسبی متغیرها در مدل درخت تصمیم جهت شناسایی عوامل مرتبط با دیابت

www.SID.ir

Archive of SID



شکل۳: مدل درخت تصمیم برای پیش بینی دیابت نوع ۲

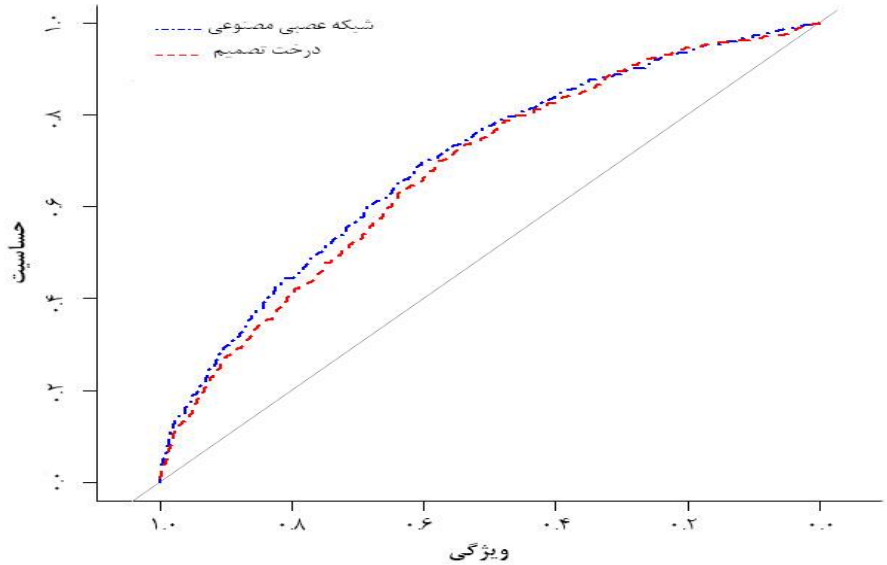
(D: Diabetic; ND: Non-Diabetic; BMI: Body Mass Index; BPS: Systolic Blood Pressure; BPD: Diastolic Blood Pressure; TG: Triglyceride; LDL: Low Density Lipoprotein; TC: Total cholesterol; FHD: family history of diabetes)

www.SID.ir

Archive of SID

جدول۳: عملکرد شبکه عصبی و درخت تصمیم برای پیش بینی دیابت نوع ۲

مدل	حساسیت(فاصله اطمینان۹۵٪)	ویژگی(فاصله اطمینان۹۵٪)	صحت (فاصله اطمینان۹۵٪)	سطح زیر منحنی راک (فاصله اطمینان۹۵٪)
شبکه عصبی	۷۴/۲ (۶۱/۲-۸۱/۳)	۵۶/۰۳ (۴۹/۵-۶۸/۵)	۶۱/۳ (۵۵/۸-۶۷/۷)	۶۹/۱ (۶۶/۵-۷۱/۸)
درخت تصمیم	۶۴/۷۷(۵۰/۵-۸۵/۹)	۶۳/۴۷(۵۲/۳-۷۷/۷)	۶۵/۳ (۵۱/۲-۷۲/۵)	۶۸/۹ (۶۶/۲-۷۱/۶)



شکل ۴: سطح زیر منحنی راک برای مدل شبکه عصبی و درخت تصمیم در داده های آزمایشی

تصمیم بود. این نتایج نشان می‌دهد که اختلاف کمی بین عملکرد دو مدل شبکه عصبی و درخت تصمیم برای پیش بینی عوامل مرتبط با دیابت وجود دارد. همچنین سطح زیر منحنی راک (۰/۲ درصد) و حساسیت (۹/۴۳ درصد) در مدل شبکه عصبی بیشتر از درخت تصمیم می باشد، در حالی که مقادیر صحت (۴ درصد) و ویژگی (۷/۴۴ درصد) در مدل درخت تصمیم بالاتر بود. بنابراین در مجموعه داده‌های تست مدل شبکه عصبی به دلیل بالا بودن سطح زیر منحنی راک، عملکرد و صحت بالاتری نسبت به مدل درخت تصمیم دارد و مدل شبکه عصبی نشان داد که عملکرد بهتری نسبت به مدل درخت تصمیم دارد و این نتیجه از مطالعه ما، مشابه با نتایج حاصل از مطالعات قبلی (۳۲-۳۴) بود. در مطالعه Meng و همکاران که به مقایسه روش های شبکه عصبی، درخت تصمیم و رگرسیون لجستیک در پیش بینی دیابت پرداخته شده است، صحت و ویژگی در مدل درخت تصمیم از دیگر مدل ها بیشتر بود. همچنین حساسیت در مدل شبکه عصبی بالاتر از سایر مدل‌ها بود (۱۶). در مطالعه Kurt و همکاران نیز که به مقایسه عملکرد شبکه عصبی و درخت تصمیم پرداختند سطح زیر منحنی راک در مدل شبکه عصبی بالاتر بود (۳۵) که با نتایج مطالعه ما هم خوانی دارد.

در این مطالعه به مقایسه و شناسایی عوامل خطر و پیش‌بینی بیماری دیابت نوع ۲ با استفاده از دو مدل شبکه عصبی و درخت تصمیم با ۲۱ متغیر پرداخته شد. یافته ها نشان داد که مدل شبکه عصبی ابزار مناسب تری برای پیش بینی عوامل مرتبط با دیابت می باشد. از طرف دیگر شواهد قوی وجود دارد که متغیرهای تری گلیسرید، شاخص توده بدنی، فشار خون و سن با دیابت در ارتباط می باشند. پس با اطلاع رسانی و آگاهی افراد جامعه و تشویق به تحرک و ورزش بیشتر می توان از ابتلا به دیابت نوع۲ جلوگیری کرد. با این حال پیگیری و جمع آوری اطلاعات کامل تر برای شناسایی عوامل مفید می‌باشد.

www.SID.ir

Archive of SID

از جمله محدودیت های مطالعه حاضر می‌توان به عدم جمع آوری اطلاعات بیماران دیابتی در مراکز بهداشت اشاره نمود که باعث کم شدن حجم نمونه مطالعه گردید. همچنین متغیر A1C که امروزه به عنوان معیاری برای شناسایی افراد دیابتی مورد استفاده قرار می گیرد در پرونده بیماران ثبت نشده بود.در مطالعات آینده می توان از روش ها و الگوریتم های جدید داده کاوی در شناسایی عوامل خطر بیماری ها استفاده کرد.

سپاسگزاری

بر خود لازم می دانیم که از همکاری معاونت و مراکز بهداشتی دانشگاه علوم پزشکی مشهد و پرسنل این مجموعه که همکاری لازم در زمینه جمع آوری اطلاعات داشتند، تشکر نماییم.

منابع

1. Hossain P, Kawar B, El Nahas M. Obesity and diabetes in the developing world--a growing challenge. The New England journal of medicine 2007; 356(3): 213-5.

2. Booth GL, Kapral MK, Fung K, Tu JV. Relation between age and cardiovascular disease in men andwomen with diabetes compared with non-diabetic people: a population-based retrospective cohort study. Lancet (London, England) 2006; 368(9529): 29-36.

3. Yach D, Hawkes C, Gould CL, Hofman KJ. The global burden of chronic diseases: overcoming impedimentsto prevention and control. Jama 2004; 291(21): 2616-22.

4. Guariguata L, Whiting DR, Hambleton I, Beagley J, et.al. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes research and clinical practice 2014; 103(2): 137- 49.

5. Sedehi M, Mehrabi Y, Kazemnejad A, F. H. Comparison of Artificial Neural Network, Logistic Regression and Discriminant Analysis Methods in Prediction of Metabolic Syndrome. IJEM. 2009; 11(6): 638-46.

6. Tapak L, Mahjub H, Hamidi O, PoorolajalJ. Real-data comparison of data mining methods in prediction of diabetes in iran. Healthcare informatics research 2013; 19(3): 177-85.

7. Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011and 2030. Diabetes research and clinical practice 2011; 94(3): 311-21.

8. Jayalakshmi T, Santhakumaran A, et al. A Novel Classification Method for Diagnosis of Diabetes Mellitus Using Artificial Neural Networks. 2010 International Conference on Data Storage and Data Engineering; 2010 9-10 Feb. 2010.

9. Choi SB, Kim WJ, Yoo TK, Park JS, et al. Screening for Prediabetes Using Machine Learning Models. Computational and Mathematical Methods in Medicine 2014; 2014:8.

10. Franco L, Jerez JM ,Alba E, et al. Artificial neural networks and prognosis in medicine. Survival

www.SID.ir

Archive of SID

analysis in breast cancer patients. 13th European Symposium on Artificial Neural Networks; 2005; Bruges, Belgium.

11. Jerez-Aragones JM, Gomez-Ruiz JA, Ramos-Jimenez G, Munoz-Perez J, et al. A combined neural network and decision trees model for prognosis of breast cancer relapse. Artificial intelligence in medicine 2003; 27(1): 45-63.

12. Bourd, Bonnevay S, et al. Comparison of Artificial NeuralNetwork with Logistic Regression as Classification Models for Variable Selection for Prediction of Breast Cancer Patient Outcomes. Advances in Artificial Neural Systems 2010; 2010: 11.

13. Chou S-M, Lee T-S, Shao YE, Chen IF. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications 2004; 27(1): 133-42.

14. Endo A, Shibata T, Tanaka H. Comparison of Seven Algorithms to Predict Breast Cancer Survival(<Special Issue>Contribution to 21 Century Intelligent Technologies and Bioinformatics). International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association 2008; 13(2): 11-6.

15. Haykin SS. Neural Networks and Learning Machines: Prentice Hall; 2009.

16. Meng X-H, Huang Y-X, Rao D-P, Zhang Q, Liu Q. Comparison of three data mining models for predicting diabetes or prediabetes by risk factors. The Kaohsiung Journal of Medical Sciences 2013; 29(2): 93-9.

17. Ramezankhani A, Pournik O, Shahrabi J, Khalili D, et al. Applying decision tree for identification of a low risk population for type 2 diabetes. Tehran Lipid and Glucose Study. Diabetes research and clinical practice 2014; 105(3): 391-8.

18. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. Diabetes care 2008; 31(5): 1040-5.

19. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters 2006; 8: 861-874.

20. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. The New England journal of medicine 2001; 344(18): 1343-50.

21. Ameri H, Alizadeh S, Barzegari A. Knowledge Extraction of Diabetics' Data by Decision Tree Method. Journal of Health Administration 2013; 16(53): 58-72.

22. Ramezankhani A, Hadavandi E, Pournik O, Shahrabi J, et al .Decision tree-based modelling for identification of potential interactions between type 2 diabetes risk factors: a decade follow-up in a Middle East prospective cohort study. BMJ open 2016; 6(12): e013336.

23. Wang C, Li L, Wang L, Ping Z, et al. Evaluating the risk of type 2 diabetes mellitus using artificial neural network: an effective classification approach. Diabetes research and clinical practice 2013; 100(1): 111-8.

24. Glumer C, Carstensen B, Sandbaek A, Lauritzen T, et al. A Danish diabetes risk score for targeted screening: the Inter99 study. Diabetes care 2004; 27(3): 727-33.

25. Hu D, Sun L, Fu P, Xie J, et al. Prevalence and risk factors for type 2 diabetes mellitus in the Chinese adult population: the InterASIA Study. Diabetes research and clinical practice 2009; 84(3): 288-95.

26. Ho WH, Lee KT, Chen HY, Ho TW, et al. Disease-free survival after hepatic resection in hepatocellular carcinoma patients: a prediction approach using artificial neural network. PloS one 2012; 7(1): e29179.

27. Walker HK, Hall WD, Hurst JW, et al. Clinical Methods: The History, Physical, and Laboratory Examinations. Boston: Butterworths. Butterworth Publishers, a division of Reed Publishing; 1990.

28. Wang CJ, Li YQ, Wang L, Li LL, et al. Development and evaluation of a simple and effective prediction approach for identifying those

www.SID.ir

Archive of SID

at high risk of dyslipidemia in rural adult residents. PloS one 2012; 7(8): e43834.

29. Barakat NH, Bradley AP, Barakat MN. Intelligible support vector machines for diagnosis of diabetes mellitus. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society 2010; 14(4): 1114-20.

30. UbeyliED. Modified mixture of experts for diabetes diagnosis. J Med Syst 2009; 33(4): 299-305.

31. Kazemnejad A, Batvandi Z, Faradmal J. Comparison of artificial neural network and binary logistic regression for determination of impaired glucose tolerance/diabetes. Eastern Mediterranean health journal = La revue de sante de la Mediterranee orientale = al-Majallah al-sihhiyah li-sharq al-mutawassit 2010; 16(6): 615-20.

32. Kang JO, Chung S-H, Suh Y-M. Prediction of Hospital Charges for the Cancer Patients with Data Mining Techniques. J Korean Soc Med Inform 2009; 15(1): 13-23.

33. Lee S-M, Kang J-O, Suh Y-M. Comparison of Hospital Charge Prediction Models for Colorectal Cancer Patients: Neural Network vs. Decision Tree Models. Journal of Korean Medical Science 2006; 19(5): 677-81.

34. Wang J, Li M, Hu YT, Zhu Y. Comparison of hospital charge prediction models for gastric cancer patients: neural network vs. decision tree models. BMC health services research 2009; 9: 161.

35. Kurt I, Ture M, Kurum AT. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Systems with Applications 2008; 34(1): 366-74.

www.SID.ir

Archive of SID

پیوست۱

نتایج حاصل از مدل رگرسیون لجستیک در افراد دیابتی و غیر دیابتی

متغیر		نسبت شانس (فاصله اطمینان ۹۵٪)	p- مقدار
سن (سال)		۱/۰۴(۱/۰۱-۱/۰۷)	۰/۰۱
وضعیت تاهل	مجرد	۶/۷۲(۰/۵-۱۷/۰۲)	۰/۱۴
	متاهل	۱/۲۸(۰/۴۷-۳/۵۰)	۰/۶۲
	مطلقه	۳/۲۳(۰/۴۸-۸/۶۲)	۰/۲۳
	بیوه	**سطح رفرنس**	
وضعیت شغلی	دانشجو	۴/۸(۰/۹-۹۳/۸۲)	۰/۰۶
	کارمند	۱/۱۴(۰/۵۹-۲/۱۶)	۰/۶۹
	بیکار	۱/۳۲(۰/۶۸-۲/۵۵)	۰/۴۱
	بازنشسته	**سطح رفرنس**	
مصرف سیگار	دارد	۱/۱۷(۰/۱-۹۹/۵۱)	۰/۰۹
	ندارد	**سطح رفرنس**	
سابقه خانوادگی دیابت	بله	۲/۵۵(۱/۶۷-۳/۹۰)	<۰/۰۰۱
	خیر	**سطح رفرنس**	
شاخص توده بدنی(کیلوگرم بر متر مربع)		۱/۰۴(۱/۰۱-۱/۰۹)	۰/۰۴
فشارخون سیستولیک(میلی متر جیوه)		۱/۰۵(۱/۰۳-۱/۰۶)	<۰/۰۰۱
فشارخون دیاستولیک(میلی متر جیوه)		۰/۹۳(۰/۸۹-۰/۹۶)	<۰/۰۰۱
تری گلیسرید(میلی گرم بر دسی لیتر)		۰/۹۸(۰/۹۹-۱/۰)	۰/۰۶
کلسترول(میلی گرم بر دسی لیتر)		۱/۰۴(۱/۰۳-۱/۰۵)	<۰/۰۰۱
لیپوپروتئین با چگالی کم(میلی گرم بر دسی لیتر)		۰/۹۶(۰/۹۵-۰/۹۷)	<۰/۰۰۱
لیپوپروتئین با چگالی بالا(میلی گرم بر دسی لیتر)		۰/۹۸(۰/۱-۹۶/۰۱)	۰/۲۹

www.SID.ir

Archive of SID

پیوست ۲

مجموعه قوانین بدست آمده از مدل درخت تصمیم CART برای پیش بینی دیابت نوع۲

۱- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم کمتر از ۶۱ میلی گرم بر دسی لیتر و تری گلیسرید کمتر از ۱۷۲ میلی گرم بر دسی لیتر باشد آنگاه با احتمال ۷۸ درصد فرد غیر دیابتی خواهد بود.

۲- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم کمتر از ۶۱ میلی گرم بر دسی لیتر و تری گلیسرید بیشتر مساوی ۱۷۲ میلی گرم بر دسی لیتر باشد آنگاه با احتمال ۸۲ درصد فرد دیابتی خواهد بود.

۳- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول بیشتر مساوی ۲۳۸ میلی گرم بر دسی لیترو مرد باشد آنگاه با احتمال ۵۹ درصد دیابتی خواهد بود.

۴- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول بیشتر مساوی ۲۳۸ میلی گرم بر دسی لیتر و زن باشد آنگاه با احتمال ۹۰ درصد غیر دیابتی خواهد بود.

۵- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید بیشتر مساوی ۱۷۰ میلی گرم بر دسی لیتر و فشار خون دیاستولیک کمتر از ۷۰ میلی متر جیوه باشد آنگاه با احتمال ۹۹ درصد غیر دیابتی خواهد بود.

۶- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید بیشتر مساوی ۱۷۰ میلی گرم بر دسی لیتر و فشار خون دیاستولیک بیشتر مساوی ۷۰ میلی متر جیوه و کارمند باشد آنگاه با احتمال ۶۳ درصد دیابتی خواهد بود.

۷- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید بیشتر مساوی ۱۷۰ میلی گرم بر دسی لیتر و فشار خون دیاستولیک و شغل بیکار یا دانشجو باشد آنگاه با احتمال ۹۸ درصد غیر دیابتی خواهد بود.

۸- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید کمتر از ۱۷۰ میلی متر جیوه و سابقه خانوادگی دیابت نداشته باشد آنگاه با احتمال ۹۲ درصد غیر دیابتی خواهد بود.

۹- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید کمتر از ۱۷۰ میلی متر جیوه و سابقه خانوادگی دیابت داشته و فشار خون سیستولیک بیشتر مساوی ۱۰۷ میلی متر جیوه باشد آنگاه با احتمال ۸۶ درصد دیابتی خواهد بود.

۱۰- اگر فشار خون سیستولیک کمتر از ۱۴۰ میلی متر جیوه، لیپوپروتئین با چگالی کم بیشتر مساوی ۶۱ میلی گرم بر دسی لیتر، کلسترول کمتر از ۲۳۸ میلی متر جیوه، تری گلیسرید کمتر از ۱۷۰ میلی متر جیوه و سابقه خانوادگی دیابت داشته و فشار خون سیستولیک کمتر از ۱۰۷ میلی متر جیوه باشد آنگاه با احتمال ۷۲ درصد غیر دیابتی خواهد بود.

۱۱- اگر فشار خون سیستولیک بیشتر مساوی ۱۴۰ میلی متر جیوه و تری گلیسرید کمتر از ۱۲۸ میلی گرم بر دسی لیتر باشد آنگاه با احتمال ۸۰ درصد غیر دیابتی خواهد بود.

۱۲- اگر فشار خون سیستولیک بیشتر مساوی ۱۴۰ میلی متر جیوه و تری گلیسرید بیشتر مساوی ۱۲۸ میلی گرم بر دسی لیتر و فشار خون دیاستولیک بیشتر مساوی ۹۲ میلی متر جیوه باشد آنگاه با احتمال ۸۴ درصد فرد دیابتی خواهد بود.

۱۳- اگر فشار خون سیستولیک بیشتر مساوی ۱۴۰ میلی متر جیوه و تری گلیسرید بیشتر مساوی ۱۲۸ میلی گرم بر دسی لیتر و فشار خون دیاستولیک کمتر از ۹۲ میلی متر جیوه و شاخص توده بدنی کمتر از ۲۸ باشد آنگاه فرد با احتمال ۸۵ درصد غیر دیابتی خواهد بود.

۱۴- اگر فشار خون سیستولیک بیشتر مساوی ۱۴۰ میلی متر جیوه و تری گلیسرید بیشتر مساوی ۱۲۸ میلی گرم بر دسی لیتر و فشار خون دیاستولیک کمتر از ۹۲ میلی متر جیوه و شاخص توده بدنی بیشتر مساوی ۲۸ باشد آنگاه با احتمال ۵۸ درصد فرد دیابتی خواهد بود.

www.SID.ir