Archive of SID

ORIGINAL ARTICLE

Investigating the Risk Factors for Low Birth Weight Using Data Mining: A Case Study of Imam Ali Hospital, Zahedan, Iran

Farzad Firouzi Jahantigh[1],
Rahele Nazarnejad[2],
Mahrokh Firouzi Jahantigh[3]

[1] Assistant Professor, Department of Industrial Engineering, Faculty of Industrial Engineering, University of Sistan and Baluchestan, Zahedan, Iran
[2] MSc in Industrial Engineering, Faculty of Engineering, Payam Noor University, Tehran, Iran
[3] Pediatricians, Zahedan University of Medical Sciences, Zahedan, Iran

(Received August 8, 2014 Accepted November 25, 2015)

Abstract

Background and purpose: High rate of mothers and infants' death and preterm birth are amongst major public health problems worldwide. The aim of this study was to identify, evaluate and rank the factors responsible for low birth weight using data mining techniques and also investigating the impact of predictor variables and developing a decision support system which could help physicians to make better treatment decisions at the birth of low weight infants.

Materials and methods: Relevant information was collected from Imam Ali Hospital affiliated with Zahedan University of Medical Sciences in 2013 including smoking, the age of mothers, etc. Different data mining algorithms were applied for modeling. Data analysis was performed in Clementine software.

Results: The variables that were very influential in predicting the low weight of infants at birth were mother's weight (100%), mother's age (98%), the number of doctor visits in the first trimester of pregnancy (45.86%), and previous preterm delivery (43.11%). Other variables poorly influenced the prediction.

Conclusion: The findings revealed some relationships between the low weight of infants at birth and mother's weight, mother's age, number of doctor visits in the first trimester of pregnancy, previous preterm delivery, high blood pressure, race, uterine irritability, and smoking. The accuracy of prediction improved via data mining techniques compared to logistic regression. Classification tree could determine the low weight of infants at birth well and random forest technique had an important role in making the diagnosis.

Keywords: low birth weight, infant, health care, data mining

J Mazandaran Univ Med Sci 2016; 26(134): 171-182 (Persian).

www.SID.ir

Archive of SID

پژوهشی

مجله دانشگاه علوم پزشکی مازندران
دوره بیست و پنجم شماره ۱۳۳ بهمن سال ۱۳۹۴ (۱۸۲-۱۷۱)

شناسایی، بررسی و رتبه بندی عوامل خطر وزن کم نوزادان در زمان تولد با استفاده از تکنیک های داده کاوی: مطالعه موردی بیمارستان امام علی (ع) دانشگاه علوم پزشکی زاهدان

فرزاد فیروزی جهانتیغ [۱]
راحله نظرنژاد [۲]
ماهرخ فیروزی جهانتیغ [۳]

چکیده

سابقه و هدف: نرخ بالای مرگ و میر مادران، نوزادان، و تولد زودرس یک مسئله مهم بهداشت عمومی در سراسر جهان است. هدف مطالعه شناسایی، بررسی و رتبه‌بندی عوامل کم وزنی نوزادان با تکنیک‌های داده کاوی برای شناسایی میزان تاثیر متغیرهای پیش‌بینی کننده و توسعه یک سیستم پشتیبانی تصمیم‌گیری برای کمک به پزشکان در تصمیم‌گیری بهترهنگام تولد نوزادان کم وزن است.

مواد و روش‌ها: اطلاعات بخش اطفال بیمارستان امام علی دانشگاه علوم پزشکی زاهدان از جمله مصرف دخانیات، سن مادران و ... در سال ۱۳۹۲ مورد بررسی قرار گرفت. در بخش مدل‌سازی از الگوریتم‌های مختلف داده کاوی استفاده گردید. برای تحلیل داده‌ها از نرم‌افزار کلمنتاین استفاده شده است.

یافته‌ها: متغیرهایی که برای پیش‌بینی کم وزنی نوزادان در زمان تولد بسیار موثرند وزن مادر (۱۰۰ درصد)، سن مادر (۹۸ درصد)، تعداد مراجعه به پزشک در سه ماهه اول بارداری (۴۵/۸۶ درصد) و سابقه زایمان زودرس (۴۳/۱۱ درصد) و سایر متغیرها به طور ضعیفی پیش‌بینی را تحت تاثیر قرار دادند.

استنتاج: نتایج مطالعه بیانگر ارتباط وزن کم نوزادان در زمان تولد با وزن مادر، سن مادر، تعداد مراجعه به پزشک در سه ماهه اول بارداری، سابقه زایمان زودرس قبلی، فشارخون بالا، نژاد، تحریک‌پذیری رحم و مصرف دخانیات بود. صحت پیش‌بینی که با رگرسیون لجستیک مقایسه شده، در تکنیک‌های داده کاوی بهبود می‌یابد. درخت طبقه‌بندی به خوبی وزن کم نوزادان در زمان تولد را مشخص می‌کند و تکنیک جنگل تصادفی در تشخیص بیماری نقش مهمی دارد.

واژه های کلیدی: وزن کم نوزادان در زمان تولد، مراقبت‌های بهداشتی، داده کاوی

مقدمه

بهبود وضعیت سلامت زنان باردار و نوزادان یک مسئله مهم بهداشت عمومی در سراسر جهان است و نیاز به رسیدگی فوری دارد، زیرا منجر به اثرات کوتاه مدت و بلندمدت در افراد، خانواده، جامعه و سیستم مراقبت‌های بهداشتی می‌شود(۱). هدف اصلی از ارائه مراقبت‌های بهداشتی، تقویت سیستم مراقبت‌های بهداشتی با شناسایی و ردیابی بیماری کشورها و بیماران در معرض خطر است. برنامه کاربردی داده کاوی برای طراحی

مولف مسئول: فرزاد فیروزی جهانتیغ – سیستان و بلوچستان: دانشکده مهندسی صنایع، دانشگاه سیستان و بلوچستان E-mail: firouzi@eng.usb.ac.ir
۱. استادیار، گروه مهندسی صنایع، دانشکده مهندسی صنایع، دانشگاه سیستان و بلوچستان، زاهدان، ایران
۲. کارشناس ارشد مهندسی صنایع، صنایع، دانشکده فنی مهندسی، دانشگاه پیام نور، تهران، ایران
۳. متخصص اطفال، دانشگاه علوم پزشکی زاهدان، زاهدان، ایران
تاریخ دریافت: ۱۳۹۴/۵/۱۷ تاریخ ارجاع جهت اصلاحات: ۱۳۹۴/۶/۲۱ تاریخ تصویب ۱۳۹۴/۹/۴

www.SID.ir

Archive of SID

مداخلات مناسب و کاهش تعداد پذیرش بیمارستان می‌تواند توسعه داده شود(۲).

پرونده بیماران یک منبع اصلی برای تحقیقات پزشکی است. به منظور استفاده از داده‌های بیماران و استخراج اطلاعات پنهان و الگوها از روش‌های داده کاوی استفاده می‌گردید. داده کاوی برای تصمیم‌گیری بهتر در بسیاری از زمینه‌ها مثل بازاریابی، تشخیص تقلب، سرمایه گذاری، تولید، ارتباطات راه دور، مهندسی، پزشکی و تحقیقات بیوپزشکی به‌کار می‌رود(۵،۴). تکنولوژی داده کاوی به سرعت در حال تحول است و به صنایع مراقبت‌های بهداشتی برای تشخیص و درمان بهتر کمک می‌کند(۷). داده کاوی، فرایند انتخاب، بررسی و مدل سازی حجم زیادی از داده‌ها را برای کشف الگوهای ناشناخته تسهیل می‌کند، به همین دلیل نسبت به روش‌های طبقه‌بندی سنتی مناسب‌تر است. کشف روند و الگوهای پنهان با داده کاوی به طور قابل توجهی درک ما را از پیشرفت بیماری، مدیریت بیماری، کمک در تشخیص پزشکی، بهبود کیفیت مراقبت از بیماران و ... افزایش می‌دهد(۶). سازمان بهداشت جهانی، وزن کم تر از ۲۵۰۰ گرم در هنگام تولد را به عنوان وزن کم نوزادان در زمان تولد (LBW) تعریف نموده است، زیرا کم‌تر از این مقدار وزن مشخص، مرگ و میر نوزادان شروع و به سرعت افزایش می‌یابد(۸). بررسی‌ها نشان داد که LBW یکی از علل اصلی مرگ و میر نوزادان است به طوری که مرگ و میر نوزادان LBW، ۴۰ برابر بیش‌تر از نوزادانی است که وزن طبیعی دارند(۳). وزن کم هنگام تولد باعث آسیب رسیدن به مغز، بیماری ریوی و کبدی مزمن، کری، کوری، صرع، معلولیت‌های یادگیری، فلج مغزی، عقب ماندگی ذهنی، اختلالات عصبی، ناتوانی جسمی، بیماری‌های تنفسی، مشکلات دستگاه گوارش، فشار خون بالا و دیابت نوع II می‌شود(۹). بنابراین به دلیل اهمیت این موضوع به بررسی عوامل موثر بر وزن کم نوزادان در زمان تولد در بیمارستان امام علی (ع) دانشگاه علوم پزشکی زاهدان پرداختیم. عوامل زیادی در بروز LBW تاثیر گذارند، شامل آنتروموتری/ وضعیت تغذیه، کم خونی، مصرف دخانیات، استفاده از مواد مخدر، کار/ فعالیت فیزیکی، مراقبت زنان در دوران بارداری، واژینوز باکتریال (عفونت) و زایمان متعدد. نوزادان کم وزن بیش‌تر در معرض خطر مرگ در طول سال اول زندگی هستند و احتمال بیش‌تری دارد با مشکلات روانی، اجتماعی، مشکلاتی در مدرسه و هنگام نوجوانی روبه رو شوند و در فعالیت‌های فکری، مخصوصاً ریاضی (محاسبات) موفقیت کم‌تری کسب کنند(۹). Hosmer و Lemeshow برای پیش‌بینی LBW از رگرسیون لجستیک استفاده کردند(۱۰). صحت طبقه‌بندی رگرسیون لجستیک برای مجموعه داده‌ها با اندازه‌های کوچک و متوسط مناسب است(۱۱).

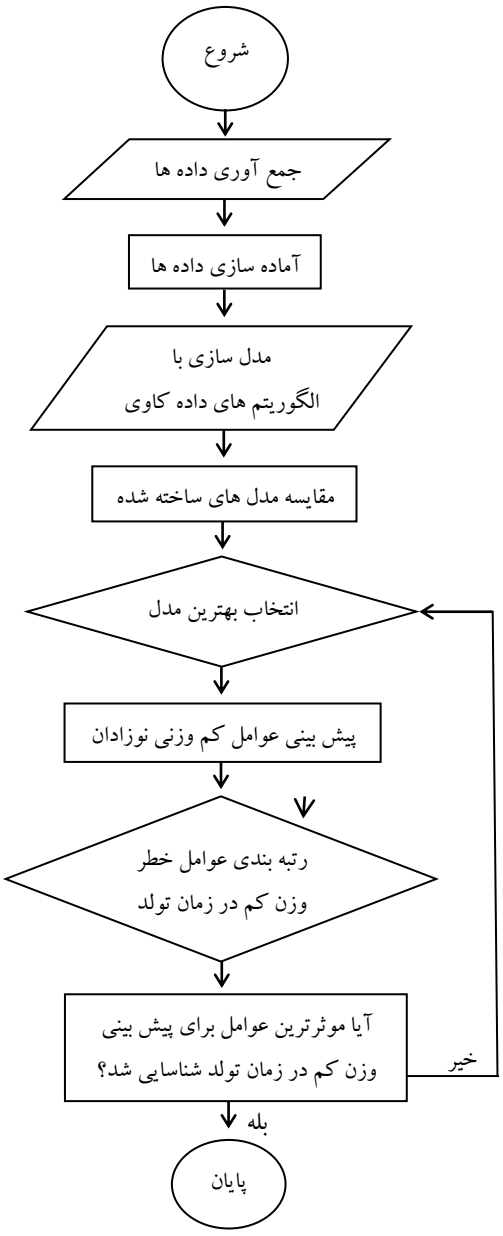
وزن هنگام تولد با تولد زودرس در ارتباط است. علت ۳۵ درصد از ۳/۱ میلیون مرگ و میر نوزادان در جهان در هر سال ناشی از تولد زودرس می باشد. نرخ مرگ و میر کودکان زیر ۵ سال، ۶۳ در ۱۰۰۰ تولد زنده در ۲۰۱۰ برآورد شده بود، در حالی که هدف توسعه هزاره، کاهش این میزان به ۳۸ است. نسبت مرگ مادران ۲۵۴ در ۱۰۰۰۰۰ تولد زنده در سال ۲۰۰۴ تا سال ۲۰۰۶ (MDGs کم‌تر از ۱۰۰) برآورده شده بود. نرخ شیوع کم‌وزنی کودکان ۴۳ درصد در سال ۲۰۰۶-۲۰۰۵ (MDGs ۲۷ درصد) برآورد شده بود(۱۲). در این مقاله تکنیک‌های داده‌کاوی برای شناسایی متغیرهای مهم پیش‌بینی به کار رفت تا یک سیستم پشتیبانی برای کمک به پزشکان برای تصمیم‌گیری بهتر در هنگام تولد نوزادان کم وزن ارائه شود.

مواد و روش ها

برای داشتن یک داده کاوی موثر علاوه بر نیاز به داده‌های مرتبط، باید از یک فرایند و روش داده کاوی مناسب نیز بهره مند شویم. مراحل انجام کار در نمودار شماره ۱ قابل مشاهده است.

www.SID.ir

Archive of SID



نمودار شماره۱: مراحل کلی انجام مطالعه

این مطالعه با اطلاعات ۱۸۹ مادر و نوزاد جمع آوری شده در بخش اطفال بیمارستان امام علی (ع) دانشگاه علوم پزشکی زاهدان انجام شد و نوزادان مرده به دنیا آمده، کنار گذاشته شدند. اطلاعات شامل مصرف دخانیات، سن مادران، سابقه زایمان زودرس قبلی، وزن در آخرین دوره قاعدگی، مراقبت از تحریک‌پذیری رحم، تعداد مراجعه به پزشک در طول سه ماه اول بارداری، فشارخون بالا و نژاد از طریق مصاحبه با خانم‌های باردار بعد از زایمان و پرونده‌های پزشکی مادران و نوزادان استخراج و در پرسشنامه‌های از پیش طراحی شده ثبت شد. مصاحبه با مادران به طور ثابت از طریق یکی از دانشجویان پزشکی انجام پذیرفت. تمامی نوزادان بعد از تولد، با استفاده از ترازوی مخصوص نوزاد با مارک سگا وزن می‌شدند. در این مطالعه هر نوزادی که وزن کم‌تر از ۲۵۰۰ گرم داشت، به عنوان نوزاد کم وزن انتخاب می‌شد. هم‌چنین زایمان زودرس، به زایمان قبل از هفته ۳۷ حاملگی اطلاق شد. داده‌های جمع‌آوری شده با الگوریتم‌های مختلف داده‌کاوی از جمله رگرسیون لجستیک، بیز ساده، جنگل تصادفی، ماشین بردار پشتیبانی، شبکه‌های عصبی و درخت طبقه‌بندی مدل‌سازی شد و با استفاده از نرم‌افزار کلمنتاین مورد تجزیه و تحلیل قرار گرفت. الگوریتم‌های داده کاوی در زیر توضیح داده شده است.

رگرسیون لجستیک

رگرسیون لجستیک یک مورد خاص از تعمیم مدل خطی است. هم‌چنین برای مدل‌سازی‌هایی که نتایج دودویی دارند، یک روش آماری است. از این مدل برای به دست آوردن نتایج بهینه پیش‌بینی استفاده می‌شود. از رگرسیون لجستیک می‌توان به عنوان تنوع دیگری از الگوریتم شبکه‌های عصبی نام برد. برخلاف درخت‌های تصمیم‌گیری یا SVM، می‌توانیم به آسانی مدل را برای داده جدید به روز رسانی کنیم. از معایب رگرسیون لجستیک این است که از SVM کندتر است و شناسایی اثرمتقابل دشوار است. ممکن است در زمینه‌های پزشکی و روانشناسی مناسب باشد، ولی یک انتخاب همیشگی نیست(۱۳).

بیز ساده

این الگوریتم بر پایه قضیه بیز برای مدل‌سازی پیش‌گویانه ارائه شده است. قضیه بیز از روشی برای دسته‌بندی پدیده‌ها بر پایه احتمال وقوع یا عدم وقوع یک پدیده استفاده می‌کند و احتمال رخ دادن یک پدیده محاسبه و دسته بندی می‌شود. طبقه بندی بیز

www.SID.ir

Archive of SID

سریع‌تر از رگرسیون لجستیک همگرا می‌شود. این روش برای کاربردهای مختلف با چندین ویژگی بسیار پرکاربرد است. عملکرد آن در شرایط پیچیده جهانی و مسائل دنیای واقعی خوب است و می‌تواند برای طبقه‌بندی دودویی و چندگانه استفاده شود و نتایج دقیق ارائه دهد. عملکرد طبقه‌بندی بیز در مقایسه با درخت تصمیم‌گیری بهتر است. به برنامه‌های تخمین پارامتر تکرار شونده پیچیده نیازی ندارد، برای ویژگی‌های گسسته، پیوسته و مجموعه داده بزرگ می‌تواند به کار برده شود و تفسیر آن آسان است. ممکن است بهترین طبقه‌بندی ممکن در یک کاربرد خاص نباشد، اما اغلب می‌توان به قوی بودن و عملکرد عالی آن اطمینان کرد. بیز ساده نرخ خطای کمی در مقایسه با طبقه‌بندی‌های دیگر دارد، اما در عمل همیشه درست نیست و در مقایسه با طبقه‌بندی‌های دیگر صحت کم‌تری دارد(۱۴).

جنگل تصادفی

جنگل تصادفی یک مجموعه طبقه‌بندی، مثل درخت تصمیم‌گیری است که برای حل مسائل طبقه‌بندی و رگرسیون می‌تواند استفاده شود. اکثر اوقات در مقایسه با درخت‌های تصمیم‌گیری به نتایج بهتری می‌رسد. جنگل تصادفی در سال ۲۰۰۱ پایه گذاری شده و درتعداد متنوعی ازبرنامه‌های کاربردی شامل تحقیقات پزشکی، پردازش تصاویر و... استفاده شده است. جنگل تصادفی برای مدل‌سازی داده با بعد زیاد مناسب است و نیاز به هرس درختان ندارد. در ساخت پایگاه داده‌های بزرگ با دقت پیش‌بینی بالا به طور موثر اجرا می‌شود. هم‌چنین از یک روش برای تشخیص اثر متقابل بین متغیرها پشتیبانی می‌کند(۱۵).

ماشین بردار پشتیبانی

ماشین بردار پشتیبانی یکی از روش‌های یادگیری با نظارت است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌کنند. این روش برای جاهایی‌که با دقت بسیار بالا نیاز به ماشینی داده‌ها داریم، به شرط این که توابع نگاشت را به درستی انتخاب کنیم، بسیار خوب عمل می‌کند. ماشین بردار پشتیبانی برای مجموعه داده غیر خطی کاربرد دارد. صحت آن نسبت به سایر طبقه‌بندی شناخته شده بالاتر است. یک طبقه‌بندی خوب است، نیازی به دانش پیشین ندارد و فضای ورودی آن بسیار زیاد است. آموزش نسبتاً ساده است و برخلاف شبکه‌های عصبی در ماکزیمم‌های محلی گیر نمی‌افتد. برای داده‌های با ابعاد بالا تقریباً خوب جواب می‌دهد. مصالحه بین پیچیدگی دسته‌بندی کننده و میزان خطا به طور واضح کنترل می‌شود. SVM ها دارای یک شالوده نظری منسجم بوده و جواب‌های تولید شده توسط آن‌ها، سراسری و یکتا می‌باشد. امروزه ماشین‌های بردار پشتیبان، به متداول‌ترین تکنیک‌های پیش‌بینی در داده‌کاوی تبدیل شده‌اند(۱۷،۱۶).

شبکه های عصبی

شبکه‌های عصبی از پرکاربردترین و عملی‌ترین روش‌های مدلسازی مسائل پیچیده و بزرگ که شامل صدها متغیر هستند، می‌باشد. شبکه‌های عصبی می‌توانند برای مسائل طبقه‌بندی (که خروجی یک کلاس است) یا مسائل رگرسیون (که خروجی یک مقدار عددی است)، استفاده شوند. هر شبکه عصبی شامل یک لایه ورودی است که هر گره در این لایه، معادل یکی از متغیرهای پیش‌بینی می‌باشد. این الگوریتم برای تجزیه و تحلیل داده‌های پیچیده‌ای که انجام آن توسط سایر الگوریتم‌ها به سادگی انجام نمی‌گیرد کاربرد دارد. شبکه‌های عصبی مصنوعی الگویی برای پردازش اطلاعات می‌باشند که با تقلید از شبکه‌های عصبی بیولوژیکی مغز انسان، الهام گرفته شده است. عضو کلیدی این ساختار جدید، سیستم پردازش اطلاعات است که تعدادی از آن‌ها به صورت مجتمع مانند نورون‌های مغز با یکدیگر کار می‌کنند تا بتوانند مسائل خاصی مانند تشخیص الگو یا طبقه‌بندی داده‌ها را از

www.SID.ir

Archive of SID

طریق فرایند یادگیری حل کنند. یادگیری در شبکه‌های عصبی رایج به شکل یادگیری تحت نظارت می‌باشد. در این روش داده‌های ورودی و خروجی هر دو به سیستم داده می‌شوند. سپس داده‌ها به دو دسته داده‌های آموزش و داده‌های آزمون تقسیم می‌شوند. سیستم بر اساس داده‌های آموزش، روابط بین ورودی‌ها و خروجی را کشف می‌کند و بعد از آن بر اساس داده‌های آزمون، اعتباریابی می‌شود. بیش‌ترین کاربرد آن در مورد مسائلی است که یک ارتباط غیرخطی میان پارامترهای معلوم و مجهول آن مسائل وجود دارد(۱۴،۱۳).

درخت طبقه بندی

درخت تصمیم یکی از قوی‌ترین و پرکاربردی‌ترین الگوریتم‌های داده‌کاوی است که برای کاوش در داده‌ها و کشف دانش کاربرد دارد. این الگوریتم داده‌ها را به مجموعه‌های مشخصی تقسیم می‌کند. هر مجموعه شامل چندین زیر مجموعه از داده‌های کم و بیش همگن که دارای ویژگی‌های قابل پیش‌بینی هستند، تقسیم می‌شود. درخت تصمیم برخلاف شبکه های عصبی، به تولید قانون می پردازد. در حالی که در شبکه‌های عصبی تنها پیش‌بینی بیان می‌شود و چگونگی آن در خود شبکه پنهان باقی می‌ماند. توضیح و تفسیر درخت تصمیم توسط کاربر غیرحرفه‌ای آسان است. درخت تصمیم قادر به بررسی هر دو ورودی عددی و غیرعددی است، نیازمند آماده‌سازی داده کم است و در پردازش با مجموعه داده‌های بزرگ کارایی دارد. به هیچ حوزه دانش یا تنظیم پارامتر نیاز ندارد و به همین دلیل برای کشف دانش اکتشافی مناسب است(۱۹،۱۸). کارایی الگوریتم‌های داده‌کاوی از نظر صحت (Accuracy)، حساسیت (Sensitivity)، تشخیص (Specificity)، سطح زیرمنحنی راک (AUC)، مقدار F، دقت (Precision) و بازخوانی (recall) بررسی شد و داده با استفاده از Leave-one-out، اعتبار سنجی متقابل (cross validation) شد، به این صورت که داده‌ها به صورت تصادفی به k قسمت مساوی و مجزا تقسیم می‌شوند. سپس k-1 قسمت برای ساخت مدل (training) و یک قسمت باقی مانده برای آزمون مدل (testing) استفاده می‌شود. آن‌گاه به طورچرخشی یک قسمت دیگر برای آزمون و بقیه برای ساخت مدل استفاده می‌شوند و تا زمانی این چرخش ادامه می‌یابد که تمام قسمت‌ها در مرحله آزمون مدل استفاده شوند. میزان صحت پیش‌بینی در هر مرحله، از داده‌های آزمون محاسبه شده و درصد صحت کلی به عنوان معیاری برای اعتبار مدل محاسبه می‌شود، نمونه‌گیری تصادفی و آزمون بر روی دنباله داده ارزیابی شد. صحت مدل بر روی مجموعه داده آزمایش شده، درصدی از مجموعه آزمایش است که به درستی طبقه‌بندی شده است. عملیات تعریف شده به صورت زیر است(۱۵-۱۲) که برای مقایسه مدل‌های ساخته شده از این شاخص‌ها استفاده کرده‌ایم: که t-pos عدد مثبت واقعی است، pos عدد مثبت است (به بیان دیگر حساسیت، حاصل تقسیم موارد مثبت واقعی به مجموع موارد مثبت واقعی و منفی کاذب است)، t-neg عدد منفی واقعی، neg عدد منفی است و f-pos عدد مثبت کاذب است. دقت به صورت زیر تعریف شده است:

$$\text{میزان حساسیت} = \frac{t-pos}{pos} \quad (۱)$$

$$\text{تشخیص} = \frac{t-neg}{neg} \quad (۲)$$

$$\text{دقت} = \frac{t-pos}{t-pos+f-pos} \quad (۳)$$

$$\text{دقت} = \text{میزان حساسیت}\frac{pos}{(pos+neg)} + \text{تشخیص}\frac{neg}{(pos+neg)} \quad (۴)$$

دقت

دقت نسبت تعداد داده های مرتبط بازیابی شده به تعداد کل داده های بی ربط و مرتبط بازیابی شده است. دقت معمولاً به صورت درصد بیان می‌شود:

$$\text{دقت} = \frac{\text{مثبت واقعی}}{\text{مثبت کاذب} + \text{مثبت واقعی}} \quad (۵)$$

www.SID.ir

Archive of SID

بازخوانی

بازخوانی نسبت تعداد رکوردهای بازیابی شده مرتبط به تعداد کل پرونده‌های مربوطه در پایگاه داده است. بازخوانی معمولاً به صورت درصد بیان می شود:

$$\text{بازخوانی} = \frac{\text{واقعیتی}}{\text{کاذبمنفی + واقعیتی}} \quad (۶)$$

F - اندازه متوسط هارمونیک وزنی از دقت و بازخوانی است.

$$F = \frac{2 * \text{دقت} * \text{بازخوانی}}{\text{دقت} + \text{بازخوانی}} \quad (۷)$$

تکنیک سطح زیر منحنی راک (AUC) یک ارزیابی جامع از دقت پیش‌بینی محدوده غربالگری از مقادیرآستانه برای تصمیم گیری فراهم می کند. این تکنیک برای مقایسه کارایی مدل‌های پیش بینی استفاده می گردد. هرچه مساحت زیر منحنی ROC بیش‌تر باشد، دقت مدل پیش‌بینی در تفکیک صحیح مقادیر پیامد، بیش‌تر خواهد بود. AUC را از منحنی راک می‌توان با معادله زیر اندازه‌گیری کرد، وقتی (ویژگی-۱) =t و ROC (t) حساسیت است:

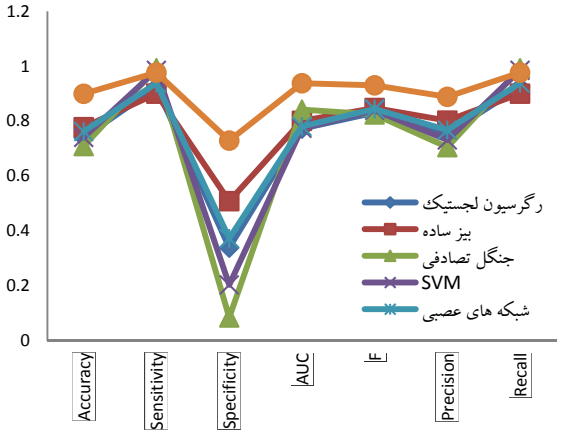
$$AUC = \int_0^1 ROC(t)\,dt \quad (۸)$$

یافته ها

در مطالعه حاضر کلیه نوزادان متولد شده در طی انجام مطالعه مورد بررسی قرار گرفتند و هیچ موردی معیار خروج از مطالعه را نداشت. در این تحقیق با استفاده از تکنیک‌های داده کاوی، نوزادان کم وزن در زمان تولد پیش‌بینی شدند. برای ساخت مدل‌های مورد نظر از تکنیک‌های رگرسیون لجستیک، بیز ساده، جنگل تصادفی، ماشین بردار پشتیبانی، شبکه‌های عصبی و درخت طبقه‌بندی استفاده کردیم. این مدل‌ها را برای داده‌های جمع آوری شده از بخش اطفال بیمارستان امام علی (ع) دانشگاه علوم پزشکی زاهدان پیاده‌سازی کرده ایم.کارایی الگوریتم‌های داده کاوی از نظر صحت، حساسیت، تشخیص، سطح زیر منحنی راک، مقدار F، دقت و بازخوانی و بررسی شدند و داده با استفاده از Leave-one-out، اعتبار سنجی متقابل، نمونه‌گیری تصادفی و آزمون بر روی دنباله داده ارزیابی شد. درستی اقدامات الگوریتم‌های داده‌کاوی مختلف با تکنیک‌های اعتبارسنجی مختلف با استفاده از نرم افزار کلمنتاین در جدول شماره ۱ نشان داده شده است و در نمودار شماره ۲ نمایش داده ایم.

جدول شماره ۱: عملکرد پیش بینی با روش های طبقه بندی مختلف

الگوریتم	صحت	حساسیت	تشخیص	AUC	F	دقت	بازخوانی
رگرسیون لجستیک	۰/۷۴۰۷	۰/۹۲۳۱	۰/۳۳۹۰	۰/۷۷۲۴	۰/۸۳۰۴	۰/۷۵۴۷	۰/۹۲۳۱
بیز ساده	۰/۷۷۷۸	۰/۹۰۰۰	۰/۵۰۸۵	۰/۸۰۰۸	۰/۸۴۷۸	۰/۸۰۱۴	۰/۹۰۰۰
جنگل تصادفی	۰/۷۰۹۰	۰/۹۹۲۳	۰/۰۸۴۷	۰/۸۴۲۰	۰/۸۲۴۳	۰/۷۰۴۹	۰/۹۹۲۳
SVM	۰/۷۴۰۷	۰/۹۸۴۶	۰/۲۰۳۴	۰/۷۷۳۸	۰/۸۳۹۳	۰/۷۳۱۴	۰/۹۸۴۶
شبکه های عصبی	۰/۷۶۱۹	۰/۹۳۸۵	۰/۳۷۲۹	۰/۷۸۰۴	۰/۸۴۴۳	۰/۷۶۷۳	۰/۹۳۸۵
درخت طبقه بندی	۰/۸۹۹۵	۰/۹۷۶۹	۰/۷۲۸۸	۰/۹۳۸۰	۰/۹۳۰۴	۰/۸۸۸۱	۰/۹۷۶۹



نمودار شماره ۲: اندازه گیری صحت روش های طبقه بندی

مشاهده شده که برای پیش‌بینی وزن کم نوزادان در زمان تولد، استفاده از مدل آزمون اعتبار سنجی بر روی دنباله داده‌های درخت طبقه‌بندی، صحت پیش‌بینی بیش‌تری دارد و به‌طور کلی صحت (۸۹/۹۵ درصد)، تشخیص (۷۲/۸۸ درصد)، AUC (۹۳/۸ درصد)، مقدار F- (۹۳/۰۴ درصد) و دقت (۸۸/۸۱ درصد) در مقایسه با روش‌های طبقه‌بندی دیگر، بیش‌تر است. سطح زیر

www.SID.ir

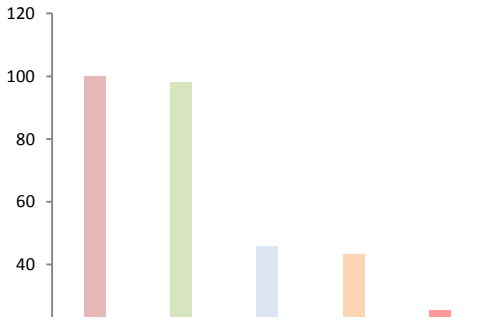
Archive of SID

منحنی راک (AUC) در درخت طبقه‌بندی از روش‌های دیگر بالاتر است، در نتیجه نشان می‌دهد که درخت طبقه‌بندی به خوبی نوزادان کم وزن در زمان تولد را مشخص می‌کند و کارایی بیش‌تری دارد. بنابراین درخت طبقه‌بندی در مقایسه با الگوریتم‌های دیگر بهتر است. حساسیت در تشخیص بیماری نقش مهمی ایفا می‌کند. جنگل تصادفی صحت پیش‌بینی پایین‌تر(۷۰/۹ درصد) و بالاترین حساسیت (۹۹/۲۳ درصد) نسبت به الگوهای دیگر را دارد. رتبه موثرترین پیش‌بینی در پیش‌بینی وزن کم نوزادان در زمان تولد در جدول شماره ۲ و نمودار شماره ۳ نمایش داده شده است. مطابق جدول شماره ۲، برای پیش‌بینی وزن کم نوزادان در زمان تولد، آخرین وزن مادر قبل ازبارداری پارامتر تاثیرگذاری است، یعنی اگر وزن مادر قبل از حاملگی کم باشد، امکان تولد نوزاد با وزن کم، زیاد است.

جدول شماره ۲: رتبه بندی عوامل خطر وزن کم در زمان تولد

خصوصیت	رتبه بندی
LWT	۱۰۰/۰۰۰۰
AGE	۹۸/۰۰۲۴
FTV	۴۵/۸۵۶۳
PTL	۴۳/۱۱۳۲
HT	۲۵/۴۳۰۳
RACE	۱۸/۲۶۳۰
UI	۱۵/۰۱۷۴
SMOKE	۷/۷۳۶۹

Mothers age in years (AGE)
the weight of the subject at her last menstrual period (LWT)
the number of physician visits during the first trimester of pregnancy (FTV)
race (RACE),
lifestyle information such as smoking (SMOKE)
a history of previous preterm delivery (PTL)
the presence of uterine irritability (UI)
hypertension (HT)



نمودار شماره ۳: رتبه بندی عوامل خطر وزن کم هنگام تولد

با توجه به نتایج به دست آمده، ۴ متغیر اول، آخرین وزن مادر قبل از باردار شدن (۱۰۰ درصد)، سن مادران (۹۸ درصد)، تعداد مراجعه به پزشک در سه ماهه اول بارداری (۴۵/۸۶ درصد) و سابقه زایمان زودرس (۴۳/۱۱ درصد) بسیار موثرهستند. فشارخون بالا، نژاد، تحریک‌پذیری رحم و مصرف دخانیات تاثیر ضعیفی در پیش‌بینی دارند.

بحث

در این تحقیق با استفاده از تکنیک‌های مختلف داده کاوی، تلاش کردیم عوامل کم وزنی نوزادان را پیش‌بینی کنیم. اطلاعات شامل مصرف دخانیات (SMOKE)، سن مادران (AGE)، سابقه زایمان زودرس قبلی (PTL)، وزن در آخرین دوره قاعدگی (LWT)، مراقبت از تحریک‌پذیری رحم (UI)، تعداد مراجعه به پزشک در طول سه ماه اول بارداری (FTV)، فشارخون بالا (HT) و نژاد (RACE) مورد بررسی قرار گرفت. Hosmer و Lemeshow از تغییرات و تکنیک‌های گام به گام انتخاب متغیر برای مدل‌های رگرسیون لجستیک استفاده کرده‌اند(۲۳،۱۰). مهم‌ترین هدف، شناسایی موثرترین عوامل برای تشخیص و پیش‌بینی بیماری است. در موثرترین پیش‌بینی، همیشه صحت پیش‌بینی مدل در حال افزایش است. تعمیم اعتبار سنجی متقابل (GCV) توسط Wahba و Craven توسعه یافته است. Friedman از فرم اصلاح شده اعتبارسنجی متقابل کلی برای شناسایی موثرترین پیش‌بینی، رتبه پیش‌بینی و حذف پیش‌بینی کم اهمیت استفاده کرد(۲۱،۲۷-۲۴). متغیرهایی که در پیش‌بینی وزن کم نوزادان در زمان تولد بسیار موثرند، آخرین وزن مادر قبل از باردار شدن، سن مادر، تعداد مراجعه به پزشک در سه ماهه اول و تعداد زایمان زودرس قبلی است(۲۹،۲۸).

در سال‌های اخیر محققان در زمینه‌های پزشکی برای پیش‌بینی زایمان زودرس و عوامل خطر آن به طور مداوم تلاش کردند. حاجیان در تحقیقی در سال ۱۳۷۷

www.SID.ir

Archive of SID

به بررسی علل شیوع کم وزنی نوزادان و برخی از عوامل خطر آن در شهر بابل پرداخته است. برای این کار از روش تجزیه و تحلیل آماری داده‌ها با استفاده از مدل رگرسیون لجستیک استفاده کرده است. نتایج این پژوهش نشان می‌دهد که چند قلو بودن، جنسیت، تعداد دفعات بارداری و سن مادر از عوامل تاثیرگذار بر کم‌وزنی نوزاد هستند(۳۰).

انصاری و همکاران ارتباط خشونت‌های دوران بارداری را با وزن کم نوزادان در زمان تولد بررسی کردند و به این نتیجه رسیدند که اعمال خشونت در دوران حاملگی، بر وزن کم نوزادان در زمان تولد موثر است. بنابراین برنامه‌ریزی گسترده برای کاهش خشونت‌ها، به ویژه خشونت فیزیکی و توجیه خانواده‌ها در مورد تبعات بعدی آن اهمیت به سزایی دارد(۳۱). حسینی و همکاران عوامل خطرساز LBW را بررسی کردند و نشان دادند که رتبه حامگی، نوع زایمان و فاصله با زایمان قبلی، به لحاظ آماری با وزن کم نوزادان در زمان تولد مرتبط بوده، اما با جنس نوزاد و سن مادر مرتبط نیست(۳۲).

هارون رشیدی و همکاران درماندگی روانشناختی مادران دارای نوزادان کم وزن و مادران دارای نوزادان با وزن مطلوب را مقایسه کردند. بین مادران دارای نوزادان کم وزن با مادران دارای نوزادان با وزن مطلوب در مقیاس افسردگی، اضطراب و استرس تفاوت معنی‌داری وجود دارد(۳۳). خلیلیان و همکاران در تحقیقی به بررسی عوامل موثر بر کم وزنی نوزادان با استفاده از آنالیز مسیری پرداختند. نتایج آنالیز مسیری نشان داد زایمان زودرس شانس کم‌وزنی نوزاد را به شدت افزایش می‌دهد. دومین متغیر اثرگذار شناسایی شده بر کم وزنی، سن حاملگی است، به صورتی که با بالا رفتن سن حاملگی، احتمال کم وزنی نوزاد کاهش می‌یابد(۳۴). بسیاری محققان دیگر در زمینه وزن کم نوزادان در زمان تولد مطالعاتی انجام دادند که در جدول شماره ۳ به برخی از آن‌ها اشاره می‌کنیم(۴۲-۳۵). با توجه به جدول شماره ۳ و مطالعات انجام شده به نظر می رسد در زمینه پیش بینی وزن کم نوزادان در زمان تولد با تکنیک‌های مختلف داده کاوی تحقیقی صورت نپذیرفته است، بنابراین ما در این مطالعه برای پیش بینی نوزادان کم وزن در زمان تولد و شناسایی متغیرهای مهم پیش‌بینی از تکنیک‌های داده‌کاوی استفاده کردیم. در پایان می‌توان نتیجه‌گیری

جدول شماره ۳: مرور نتایج برخی از محققان

نویسندگان	سال تحقیق	مسئله تحقیق	نتایج
E. Nkwabong و همکاران (۲۸)	۲۰۱۵	پیش بینی عوامل موثر بر کم وزنی نوزادان	BMI کم تر از ۲۵، قد کم تر از ۱۶۵ cm و سن کم تر از ۲۰ سال، خطر تولد نوزاد کم وزن را افزایش می دهد.
هوانگ و همکاران	۲۰۱۵	پیش بینی عوامل موثر بر کم وزنی نوزادان	فاکتورهای افزایش ناکافی یا بیش از حد وزن بارداری، فشار خون دوران بارداری، تعداد مراقبت های قبل از بارداری و تعادل غذایی بر روی کم وزنی نوزادان تاثیر دارند.
Nynke R. Vanden Broek و همکاران	۲۰۱۴	پیدا کردن ارتباط عوامل مستقل با تولد زودرس، در اولین و آخرین تولد زودرس با رگرسیون لجستیک چندگانه	تاریخ بارداری زنان و شناسایی مادر کم وزن، مالاریا و کم خونی از عوامل خطر برای تولد زودرس است، وضعیت HIV در تولد زودرس نقشی ندارد.
Yorfuji و همکاران	۲۰۱۳	بررسی کردند که آیا با استفاده از مدل رگرسیون لجستیک، موقعیت اجتماعی و اقتصادی و خصوصیات والدین بر تولد نوزادان کم وزن مرتبط است و آیا بین آلودگی هوا و عوارض جانبی تولد ارتباطی وجود دارد؟	آلودگی هوا یک عامل خطر مستقل برای عوارض جانبی تولد است.
Hsiang-yang chen و همکاران	۲۰۱۱	بررسی عوامل خطر زایمان زودرس با استفاده از داده کاوی با شبکه عصبی و درخت تصمیم گیری	تولد متعدد و خونریزی در دوران بارداری از عوامل مهم برای تولد زودرس است. این مطالعه اشاره دارد که مردان نیز در تولد زودرس نوزادان نقش دارند.
Aparajita Dasgupta و همکاران	۲۰۱۱	استفاده ازمدل های لجستیک چندگانه برای شناسایی عوامل موثر بر LBW	کم خونی در دوران بارداری به طور قابل توجهی با LBW در ارتباط است.
معصومه دل آرام	۲۰۱۰	استفاده از رگرسیون لجستیک و آزمون کای اسکوئر برای شناسایی عوامل موثر بر LBW	کم وزنی در نوزادان چند قلو حدود ۱۶/۵ برابر نوزادان تک قلو، در نوزادان با سن حاملگی کم تر از ۳۷ هفته حدود ۱۰ برابر نوزادان بیش از ۳۷ هفته، در نوزادان دختر حدود ۲/۵ برابر نوزادان پسر، در سن کم تر از ۲۰ سال ۲/۷۵ برابر سن بیش تر از ۲۰ سال و در حاملگی اول ۳/۶۹ برابر حاملگی های چندم است.
Nasreen و همکاران	۲۰۱۰	بررسی تاثیر افسردگی و علائم اضطراب مادر بر تولد نوزادان کم وزن	افسردگی و نشانه های اضطراب مادر در طی بارداری بر تولد نوزادان کم وزن تاثیر دارد.
Kleanthis C.Neokleous و همکاران	۲۰۰۸	پیش بینی خطر زایمان زودرس با شبکه عصبی و بررسی اثر طول دهانه رحم برای پیش بینی زایمان زودرس	حساسیت به طول دهانه رحم بسیار قابل توجه است.

www.SID.ir

Archive of SID

کرد که وزن کم نوزادان در زمان تولد یکی از علل اصلی مرگ و میر نوزادان شناسایی شده و بالاترین خطر را برای افراد، خانواده‌ها و جامعه دارد. در مطالعه حاضر ما الگوریتم‌های داده‌کاوی مختلف را برای پیش‌بینی وزن کم نوزادان در زمان تولد به کار بردیم. این مدل‌ها برای داده‌های بخش اطفال بیمارستان امام علی (ع) دانشگاه علوم پزشکی زاهدان پیاده‌سازی شده‌اند، عملکرد هر یک از مدل‌ها مورد بررسی قرار گرفته و نتایج به دست آمده با یکدیگر مقایسه گردیده‌اند. اعتبارسنجی انجام شده بر روی مدل‌ها اثبات می‌کند که نتایج به دست آمده دقیق و قابل اعتماد بوده‌اند. در تحقیق حاضر و دیگر تحقیقات مورد بررسی، میزان تشخیص مدل‌ها کم‌تر از حساسیت می‌باشد، این بدین معنی است که مدل در تشخیص افراد بیمار نسبت به تشخیص افراد سالم، تواناتر است. درخت طبقه‌بندی در مقایسه با الگوریتم‌های دیگر بهتر است. Hosmer و Lemeshow به جای مدل رگرسیون لجستیک از روش رگرسیون گام به گام استفاده کردند(۱۰). نتایج این کار در مقایسه با رگرسیون لجستیک صحت پیش‌بینی در داده‌کاوی را بهبود داده است. بنابراین رگرسیون گام به گام بهتر از مدل رگرسیون لجستیک عمل کرده است. یک نگرش کلی به نتایج حاصل از مطالعه حاضر و سایر مطالعات انجام شده در این زمینه مشخص می‌سازد که وزن کم نوزادان به علل متعددی وابسته است و معلول یک علت نیست. با توجه به محدودیت‌های مطالعه ما از جمله انجام آن فقط در یک مرکز آموزشی درمانی و عدم دسترسی به اطلاعات سایر زایشگاه‌ها، نتایج مطالعه حاضر قابل تعمیم به کل کشور نمی‌باشد. لذا با عنایت به این مساله و نیز با توجه به نتایج مطالعه حاضر، لزوم اجرای یک برنامه مدون و دقیق و طراحی و اجرای برنامه‌هایی برای کاهش مرگ نوزادان ضروری می‌باشد. توصیه می‌شود در مطالعات آینده از تکنیک‌های محاسبات نرم‌افزاری برای بهبود صحت پیش‌بینی و ایجاد یک ابزار قدرتمند برای کمک به پزشکان در تصمیم‌گیری برای پیش‌بینی وزن کم نوزادان در زمان تولد استفاده شود.

References

1. Kavitha R. Health Care Industry in India. IJSRP 2012; 2(8): 1-4.
2. Sathiya Rani N, Vimala K, Kalaivani V. Health Care Monitoring for the CVD Detection using Soft Computing Techniques. IJFCST 2013; 3(4): 21-30.
3. Mosayebi Z, Fakhraee SH, Movahedian AH. Prevalence and risk factors of low birth weight infants in Mahdieh hospital, Tehran. Kaums 2004; 8(2): 58-67 (Persian).
4. Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data. CACM 1996; 39(11): 27-34.
5. Bellazzi R, Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. Int J Med Inform 2008; 77(2): 81-97.
6. Balakrishnan S, Savarimuthu N, Narayanaswamy R, Samikannu R. SVM Ranking with Backward Search for Feature Selection in Type II Diabetes Databases. Singapore :IEEE International Conference on Systems. Man and Cybernetics, 2008.
7. Sandhya J, Deepa Sh, Patnaik LM, Rrashmi PL, Venugopal KR, Vibhudendra Simha GG. Classification of Alzheimer's Disease and Parkinson's Disease by Using Machine Learning and Neural Network Methods. Second International Conference on Machine Learning and Computing. 2010; 218-222.
8. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ 1987; 65(5): 663–737.

www.SID.ir

Archive of SID

9. Bull J, Mulvihill C, Quigley R. Prevention of low birth weight: assessing the effectiveness of smoking cessation and nutritional interventions, 1st ed. London: HDA, 2003.
10. Hosmer DW, Lemeshow Jr SA, Sturdivant RX. Applied Logistic Regression, 3rd ed. Hoboken, NJ: Wiley; 2013.
11. Chan KY, Loh WY. LOTUS: An Algorithm for Building Accurate and Comprehensible Logistic Regression Trees. Journal of Computational and Graphical Statistics 2004; 13(4): 826-852.
12. Rao M, Mant D. Strengthening primary healthcare in India: white paper on opportunities for partnership Explore how India and the UK can work together on education. Professional development, affordable technologies, public-private partnerships, governance, and innovation in primary care in India. BMJ 2012; 344(3151): 1-14.
13. Raghavendra BK, Srivatsa SK, Evaluation of Logistic Regression and Neural Network Model with Sensitivity Analysis on Medical Datasets. International Journal of Computer Science and Security (IJCSS) 2011; 5(5): 503 -511.
14. Elfaki Yahia M, El-mukashfi El-taher M. A New Approach for Evaluation of Data Mining Techniques. IJCSI International Journal of Computer Science Issues. September 2010; 7(5): 181-186.
15. Jehad A, Rehanullah Kh, Nasir A. Imran Maqsood Random Forests and Decision Trees. IJCSI International Journal of Computer Science Issues 2012; 9(3): 272-278.
16. Santhosh Baboo S, Sasikala S. Multicategory Classification Using Support Vector Machine for Microarray Gene expression Cancer Diagnosis. Global Journal of Computer Science and Technology 2010;10(15): 38-44.
17. Koh HC, Tan G. Data Mining Applications in Healthcare. J Healthc Inf Manag 2005; 19(2): 64-72.
18. Parashar HJ, Vijendra S, Vasudeva N. An Efficient Classification Approach for Data Mining. International Journal of Machine Learning and Computing. 2012; 2(4):446-448.
19. Ozer P, Dr. I.G. Sprinkhuizen-Kuyper. Data Mining Algorithms for Classification. Thesis Artificial Intelligence. Radboud University Nijmegen. January 2008; 1-27.
20. Han J, Kamber M, Pei J. Data Mining: Concepts and Techniques, 3rd ed. San Francisco: Morgan Kaufman, 2011.
21. Senthilkumar D, Paulraj S. Diabetes Disease Diagnosis Using Multivariate Adaptive Regression Splines. IJET 2013; 5(5): 3922-3929.
22. Keller F, Crocker M. Connectionist and Statistical Language Processing. Universitat des Saarlandes. 2003.
23. University of Massachusetts Amherst, statistical software information, Index of Multiple Regression Datasets, Available online at: http://www.umass.edu/statdata/statdata/stat-rmult.html. Accessed May 2, 2015.
24. Friedman JH. Multivariate adaptive regression splines. The Annals of Statistics 1991; 19(1): 1-141.
25. Friedman JH, Roosen CB. An introduction to multivariate adaptive regression splines. Stat Methods Med Res 1995; 4(3): 197-217.
26. Craven P, Wahba G. Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. Numer Math 1979; 31(4): 377-403.
27. Ambalavanan N, Carlo WA, Bobashev G, Mathias E, Liu B, Poole K, et al. Prediction of Death for Extremely Low Birth Weight

www.SID.ir

Archive of SID

Neonates. Pediatrics 2005; 116(6): 1367-1373.

28. Nath Das R, Sanatombi Devi R, Kim J. Mothers' Lifestyle Characteristics Impact on Her Neonates' Low Birth Weight. IJWHR 2014; 2(4): 229-235.
29. van den Broek NR, Jean-Baptiste R, Nelison JP. Factors Associated with preterm, early preterm and late preterm birth in Malawi. PLoS One 2014; 9(3): 1-8.
30. Hajian K. A study of the Prevalence of low Birth weight and its Risk Factors in Babol, in 1998. J Mazandaran Univ Med Sci 2000; 10(26): 49-55 (Persian).
31. Ansari H, Parisae Z, Rahimi E, Rakhshaei F. Assessment of Relationship between violence Exposure pregnancy and low birth weight: a case–control study. Pars Journal of Medical Sciences (Jahrom Medical Journal) 2008; 6(2): 17-26 (Persian).
32. Hosseini SZ, Bahadori MH, Fallah Bagher Shaidaei H. Incidence of low birth weight and associated risk factors during March 2002-2003 in Tonekabon, Iran. J Mazandaran Univ Med Sci 2005; 15(49): 110-113 (Persian).
33. Haroon rashidi H, arjmandniya AK, Afrooz GhA, Beshlideh Q. Comparison of Psychological Distress of Low Birth Weight – Newborns, Mothers with those of normal weight-newborns' counterparts. J Ilam Univ Med Sci 2014; 21(7): 54-59 (Persian).
34. Khalilian AR, Hamta A, Farhadi R, Ranjbaran H. Investigation factors of Low Birth Weight infants with Structural Equation Model Approach. J Mazandaran Univ Med Sci 2012; 21(86): 108-114 (Persian).
35. Nkwabong E, Kamgnia Nounemi N, Sando Z, Mbu R, Mbede J. Risk factors and placental histopathological findings of term born low birth weight neonates. Placenta 2015; 36(2): 138-141.
36. Huang A, Jin X, Liu X, Gao S. A matched case-control study of preterm birth in one hospital in Beijing, China. Reprod Health 2015, 12: 1.
37. Yorifuji T, Naruse H, Kashima S, Takao S, Murakoshi T, Doi H, et al. Residential proximity to major roads and adverse birth outcomes: a hospital-based study. Environmental Health 2013; 12(1): 34.
38. Dasgupta A, Basu R. Determinants of low birth weight in a Blowk of Hooghly, West Bengal: A multivariate analysis. International Journal of Biological & Medical Research 2011; 2(4): 838-842.
39. Chen H-Y, Chuang Ch-H, Yang Y-J, Wu T-P. Exploring the risk factors of preterm birth using data mining. Expert Systems with Applications 2011; 38(5): 5384-5387.
40. Delaram M. The Incidence and Related Factors of Low Birth Weight. IJN 2010; 23(64): 29-36.
41. Nasreen HE, Kabir ZN, Forsell Y, Edhborg M. Low birth weight in offspring of women with depressive and anxiety symptoms during pregnancy: results from a population based study in Bangladesh. BMC Public Health 2010; 10: 515.
42. Neocleous C, Schizas CN, Anastasopoulos P, Nikolaides KH, Neokleous KC, Pattichis CS. Neural networks to estimate the influence of cervix length on the prediction of spontaneous preterm delivery before 37 weeks. Proceedings of the 5th International Conference on Information Technology and Application in Biomedicine. 2008; 423-425.

www.SID.ir