مجله تخصصی اپیدمیولوژی ایران؛ ۱۳۹۳، دوره ۱۰، شماره ۲: صفحات ۳۹-۳۳.

مقاله پژوهشی

کاربرد روش‌های کلاس‌بندی رگرسیونی در شناسایی اثرات متقابل عوامل خطر فلج مغزی کودکان

سپیده زارع دلاور[۱]، عنایت‌اله بخشی[۲]، فرین سلیمانی[۳]، اکبر بیگلریان[۴]

[۱] فوق‌لیسانس آمار زیستی، گروه آمار زیستی، دانشگاه علوم بهزیستی و توان‌بخشی، تهران، ایران

[۲] استادیار، گروه آمار زیستی، دانشگاه علوم بهزیستی و توان‌بخشی، تهران، ایران

[۳] دانشیار و متخصص کودکان، مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران

[۴] استادیار، گروه آمار زیستی، دانشگاه علوم بهزیستی و توان‌بخشی، تهران، ایران

نویسنده رابط: اکبر بیگلریان، نشانی: تهران، اوین، بلوار دانشجو، بن‌بست کودکیار، دانشگاه علوم بهزیستی و توانبخشی، گروه آمار زیستی. تلفن: ۲۲۱۸۰۱۴۶ ، نمابر: ۲۲۱۸۰۱۴۶

پست الکترونیک: abiglarian@uswr.ac.ir

تاریخ دریافت: ۹۲/۵/۲۱؛ پذیرش: ۹۲/۱۰/۷

مقدمه و اهداف: شناسایی عوامل خطر و اثرات متقابل آن‌ها، در مطالعه‌های پزشکی بسیار مهم است. هدف از این مطالعه شناسایی اثرات متقابل عوامل خطر فلج مغزی در کودکان ۶-۱ ساله دارای فلج مغزی است.

روش کار: در این مطالعه مقطعی، اطلاعات ۲۲۵ کودک ۶-۱ ساله که در سال‌های ۸۷-۱۳۸۶جمع‌آوری شده بود؛ مورد استفاده قرار گرفت. برای شناسایی اثرات متقابل از روش‌های کلاس‌بندی رگرسیونی شامل درخت کلاس‌بندی رگرسیونی، آدابوست، بگینگ و الگوریتم C4.5 با استفاده از نرم‌افزار R نسخه ۱-۰-۳ استفاده شد.

نتایج: اثرات متقابل برگزیده از روش آدابوست (ازدواج فامیلی والدین: جنس کودک، حاملگی‌های قبلی: زایمان واژینال، ازدواج فامیلی والدین: زایمان نارس: جنس کودک، زایمان نارس: سابقه بیماری: آسفیکسی، ازدواج فامیلی والدین: آسفیکسی: جنس کودک، سابقه بیماری: جنس کودک: کوچک جثه نسبت به سن حاملگی، عفونت نوزادی: آسفیکسی: کوچکی جثه نسبت به سن حاملگی، سابقه بیماری: آسفیکسی: جنس کودک، زایمان نارس: زایمان واژینال: آسفیکسی)، روش بگینگ (ازدواج فامیلی والدین: آسفیکسی، ازدواج فامیلی والدین: زایمان نارس: آسفیکسی)، الگوریتم C4.5 (آسفیکسی: زایمان نارس، آسفیکسی: ازدواج فامیلی والدین: سابقه بیماری: زایمان نارس) و درخت کلاس‌بندی رگرسیونی (آسفیکسی: ازدواج فامیلی والدین) شناسایی شدند. حساسیت و ویژگی روش آدابوست بهتر از سایر روش‌های گفته شده، به ترتیب ۰/۰۲۹±۰/۹۴۱ و ۰/۰۳۰±۰/۹۵۱، به دست آمد.

نتیجه‌گیری: روش آدابوست عملکرد موفق‌تری در تشخیص و مدل‌سازی اثرات متقابل احتمالی میان عوامل خطر فلج مغزی کودکان داشت.

واژگان کلیدی: درخت کلاس‌بندی رگرسیونی، فلج مغزی، کودک، اثر متقابل

مقدمه

فلج مغزی[۱] به عنوان یکی از انواع معلولیت‌های جسمی، ناتوانی در کنترل حرکات و وضع اندامی بدن بر اثر آسیب مغزی است که این آسیب به طور معمول در حین فرایند تولد و یا مدت کوتاهی پس از آن رخ می‌دهد (۱). در واقع، فلج مغزی سندرومی است که دارای علایم زیر است: اختلالات حرکتی، اختلالات شناختی، تشنج و اختلال‌های رفتاری ناشی از آسیب مغزی. یک فرد ممکن است تنها یکی از نقایص آسیب مغزی مثل نقص حرکتی را از خود بروز دهد، در حالی که دیگری ممکن است دارای ترکیبی از این علایم باشد (۲).

فلج مغزی شایع‌ترین اختلال سیستم اعصاب مرکزی در دوران کودکی است (۳). شیوع فلج مغزی کودکان در ایران ۲ در هزار تولد زنده (۴) و شیوع آن در امریکا ۲/۴-۱/۴ در هزار تولد زنده برآورد شده است. علاوه بر این، ۵۰۰۰ نوزاد و ۱۵۰۰-۱۲۰۰ کودک پیش از مدرسه در امریکا هر ساله با فلج مغزی کودکان تشخیص داده می‌شوند (۵). با وجود پیشرفت‌های فن‌آوری در مراقبت‌های ویژه نوزادی[۲] و بهبود مراقبت‌های دوران بارداری در

[۱] Cerebral palsy

[۲] Neonatal intensive care

[Downloaded from journals.tums.ac.ir on 2024-06-17]

دو دهه‌ی اخیر، هنوز یکی از علت‌های اصلی و رایج اختلالات تکاملی دوران کودکی باقی مانده است (۶).

بیش‌تر موارد ابتلا، به جز موارد خفیف، در ۱۸ ماه اول زندگی مشخص می‌شوند. علایم به صورت ناتوانی در رسیدن به استانداردهای حرکتی، وجود اختلالاتی مثل عملکرد حرکتی خشن و نامتقارن و افزایش تون[۱] عضلات یا کاهش تون عضله است (۷). عواملی که باعث می‌شود کودکی مبتلا به فلج مغزی شود؛ قابل شناسایی می‌باشند، اما ممکن است به تنهایی با بیماری ارتباطی نشان ندهند، و در تعامل با عوامل خطر دیگر تأثیر مهم و قابل توجهی داشته باشند. به طور معمول، در بررسی هم‌زمان چندین عامل خطر در یک بیماری رایج، که بیش‌تر کیفی هستند، معمولاً نقش هر عامل به تنهایی، آن‌قدر کوچک است که این تأثیر با روش‌های کلاسیک آماری قابل تشخیص نمی‌باشند، اما بسیار بعید است که تنها یک عامل مسؤول بروز این بیماری باشد، بنابراین فرض بر این است که تعدادی عامل، به صورت هم‌زمان تغییرات مرتبط با خطر ابتلا به بیماری را تبیین می‌نمایند. از طرف دیگر بسیاری از این اثرات به صورت ضربی با بیماری مورد نظر ارتباط دارند (اثر متقابل هر عامل با عامل دیگر یا اثر متقابل هر عامل با عوامل محیطی) و این امر تشخیص عوامل مرتبط با بیماری و در نتیجه تعیین پیش آگاهی ابتلا به بیماری را مشکل می‌سازد. وجود چنین اثرات متقابلی، برای پزشکان و درمان‌گران مهم و شناسایی آن مهم‌تر است.

یکی از روش‌های شناسایی اثرات متقابل بین عوامل خطر، استفاده از روش‌های رگرسیونی است. برای پاسخ‌های دو حالتی، روش‌های رگرسیونی مانند رگرسیون لجستیک[۲] قابل استفاده است، اما وقتی که تعداد متغیرهای مستقل زیاد باشد، به خصوص این‌که متغیرهای مورد بررسی یا بیش‌تر آن‌ها از جنس کیفی باشند، رگرسیون لجستیک در شناسایی اثرات متقابل ناکارامد خواهد بود. در این حالت به کارگیری روش‌های جایگزین مانند روش‌های آدابوست[۳]، بگینگ، الگوریتم C4.5 و درخت کلاس‌بندی رگرسیون، مورد توجه و قابل استفاده خواهد بود. این تحقیق عملکرد روش‌های یادگیری ماشینی بر مبنای درخت را در تشخیص اثرات متقابل بیماری فلج مغزی کودکان بررسی می‌نماید.

[۱] Tone
[۲] Logistic regression
[۳] Adaptive Boosting

روش کار

این مطالعه، یک مطالعه مقطعی بود که در آن از داده‌های ۲۲۵ نفر کودک ۶-۱ ساله استفاده شد. داده‌های این مطالعه، برگرفته از طرحی است که در مرکز تحقیقات توان‌بخشی اعصاب اطفال دانشگاه علوم بهزیستی و توان‌بخشی انجام شده بود (۸). اطلاعات مربوط در خصوص عوامل مؤثر پری‌ناتال[۴] و نئوناتال[۵] از پرونده نوزادی، شرح حال و پرکردن پرسشنامه توسط متخصص کودکان آشنا به طب تکامل کودک به‌دست آمد. متغیرهای مرتبط با عوامل خطر فلج مغزی کودکان شامل: جنس کودک (دختر/پسر)، زایمان بریچ اختلال تنفس نوزادی، زردی نوزادی، عفونت نوزادی، سابقه بیماری در مادر، ازدواج فامیلی والدین، پارگی زودرس کیسه آب، سابقه درمان ناباروری، چند قلوزایی، خونریزی واژینال، آسفیکسی، جثه کوچک نسبت به سن حاملگی، زایمان نارس (سن حاملگی ۳۷ هفته یا کم‌تر) زایمان پرخطر (اولین زایمان یا بیش از ۴ زایمان)، سابقه سقط جنین، نوع زایمان، سن مادر (پرخطر: کم‌تر از ۱۶ سال یا بیش‌تر از ۴۰ سال)، جمع‌آوری و مورد استفاده قرار گرفت.

در مرحله تحلیل داده‌ها، تعداد ۲۳ نمونه دارای مقادیر گمشده در یک یا چند متغیر مورد مطالعه بودند که با حذف این نمونه‌ها تعداد ۲۰۲ نمونه، ۹۳ کودک دارای فلج مغزی و ۱۰۹ کودک سالم، وارد تحلیل شدند. برای شناسایی اثرات متقابل عوامل خطر از روش‌های آدابوست، بگینگ، الگوریتم C4.5 و درخت کلاس‌بندی رگرسیونی در نرم افزار R نسخه ۱-۰-۳ استفاده شد. لازم به ذکر است که به منظور اجتناب از بیش برازش[۶]، روش اعتبار سنجی ۱۰ قسمتی[۷] استفاده قرار گرفت و برای مقایسه روش‌ها از میانگین حساسیت، ویژگی، خطا و ضریب همبستگی متیو استفاده گردید. در عمل، روشی مطلوب است که به صورت هم‌زمان حساسیت و ویژگی آن بالاتر و خطای آن کم‌تر باشد.

یافته‌ها

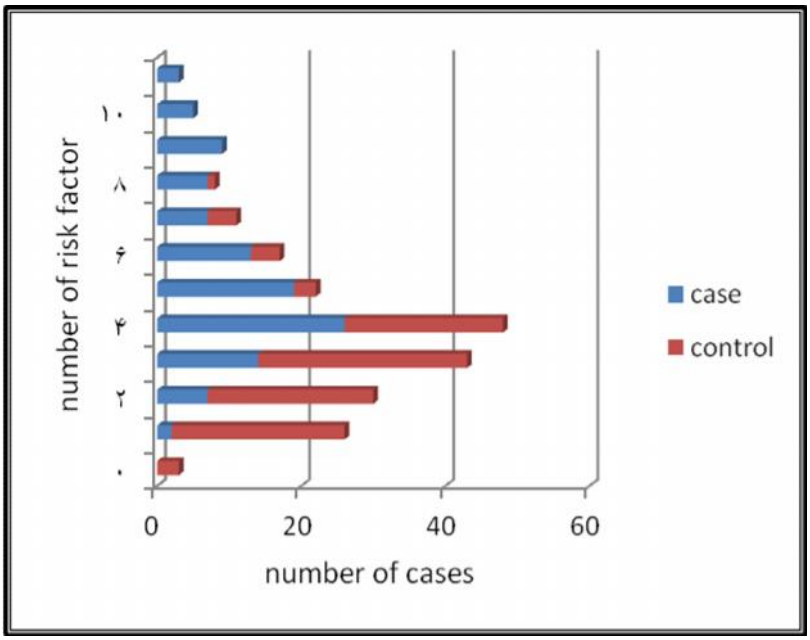
۵۲/۵ درصد از کودکان مورد بررسی پسر بودند. در کل، ۱۴/۷ درصد از کودکان دارای اختلال تنفسی، ۲۴/۴ درصد دارای زردی نوزادی، ۱۴/۶ درصد دارای عفونت نوزادی، ۱۸/۰ درصد دارای سابقه بیماری، ۷۱/۰ درصد دارای والدین با ازدواج فامیلی، ۹/۴

[۴] Prenatal
[۵] Neonatal
[۶] Over-fitting
[۷] 10-fold cross-validation

[Downloaded from journals.tums.ac.ir on 2024-06-17]

درصد دارای پارگی زودرس کیسه آب، ۴/۹ درصد دارای درمان کمکی ناباروری، ۴/۹ درصد دارای چند قلویی‌زایی، ۷/۲ درصد دارای سابقه خونریزی واژینال در دوران بارداری، ۵۲/۸ درصد دارای آسفیکسی، ۲۱/۳ درصد کوچکی جثه نسبت به سن حاملگی، ۱۹/۹ درصد دارای زایمان نارس، ۱/۴ درصد دارای زایمان پرخطر، ۱۷/۶ درصد دارای سابقه سقط جنین، ۱۹/۷ درصد دارای حاملگی‌های قبلی، ۵۸/۴ درصد دارای زایمان واژینال و ۷/۶ درصد از مادران دارای سن پر خطر بودند. کودکان سالم و فلج مغزی از نظر تعداد عوامل خطر، به طور هم‌زمان، متفاوت بودند. کودکان سالم از تعداد عوامل خطر کم‌تری برخوردار بودند و حداکثر ۴-۳ عامل خطر به طور هم‌زمان داشتند، اما در گروه کودکان فلج مغزی، بیش‌تر آن‌ها حداقل ۴ تا بیش از ۱۰ عامل خطر به طور هم‌زمان داشته‌اند (شکل شماره ۱).

برای شناسایی اثرات متقابل، از روش‌های آدابوست، بگینگ، الگوریتم C4.5 و درخت کلاس‌بندی رگرسیونی استفاده شد که نتیجه تحلیل در جدول شماره ۱ گزارش شده است. در ادامه، عملکرد روش‌های گفته شده با چهار ملاک حساسیت، ویژگی، خطا و ضریب همبستگی متیو (MCC) با استفاده از روش اعتبار سنجی متقابل ۱۰ قسمتی مقایسه شدند. روش آدابوست با خطای ۰/۰۶، حساسیت و ویژگی ۰/۹۴ و ۰/۹۵ در کل عملکرد موفق‌تری در شناسایی اثرات متقابل نسبت به سایر روش‌ها داشت و در ادامه الگوریتم C4.5 دارای عملکرد بهتری بود (جدول شماره ۲). برای مثال، اثرات متقابل دو طرفه آسفیکسی: زایمان نارس به این معنی است که اثر آسفیکسی بر فلج مغزی در سطوح مختلف زایمان نارس متفاوت است، یا اثر متقابل سه تایی آسفیکسی: ازدواج فامیلی والدین: زایمان نارس به این معنی است که اثر ازدواج فامیلی بر فلج مغزی در سطوح مختلف ایجاد شده از آسفیکسی: زایمان نارس برابر نیست.



شکل شماره ۱- توزیع تعداد عوامل خطر هم‌زمان در کودکان سالم (کنترل) و فلج مغزی (مورد)

[Downloaded from journals.tums.ac.ir on 2024-06-17]

جدول شماره ۱- اثرات متقابل تشخیص داده شده توسط روش‌های آدابوست، بگینگ، الگوریتم*C4.5* و درخت کلاس‌بندی رگرسیونی

روش		اثرات متقابل تشخیص داده شده
آدابوست	۲طرفه	ازدواج فامیلی والدین: جنس کودک حاملگی‌های قبلی: زایمان واژینال
	۳طرفه	ازدواج فامیلی والدین: زایمان نارس: جنس کودک زایمان نارس: سابقه بیماری: آسفیکسی ازدواج فامیلی والدین: آسفیکسی:جنس کودک سابقه بیماری: جنس کودک: کوچکی جثه نسبت به سن حاملگی عفونت نوزادی: آسفیکسی: کوچکی جثه نسبت به سن حاملگی سابقه بیماری: آسفیکسی: جنس کودک زایمان نارس: زایمان واژینال:آسفیکسی
بگینگ	۲طرفه	ازدواج فامیلی والدین: آسفیکسی
	۳طرفه	ازدواج فامیلی والدین: زایمان نارس: آسفیکسی
الگوریتم C4.5	۲طرفه	آسفیکسی: زایمان نارس
	۴طرفه	آسفیکسی: ازدواج فامیلی والدین: سابقه بیماری: زایمان نارس
درخت کلاس بندی رگرسیونی	۲طرفه	آسفیکسی: ازدواج فامیلی والدین

جدول شماره۲- میانگین شاخص‌های به دست آمده در روش‌های آدابوست، بگینگ، الگوریتم *C4.5* و درخت کلاس‌بندی رگرسیونی مبتنی بر اعتبارسنجی ۱۰ قسمتی

روش	حساسیت	ویژگی	خطا	ضریب همبستگی متیو
آدابوست	۰/۹۴۱±۰/۰۲۹	۰/۹۵۱±۰/۰۳۰	۰/۰۶۰±۰/۰۰۶	۰/۸۸۷±۰/۰۲۸
بگینگ	۰/۸۳۲±۰/۰۲۷	۰/۹۰۹±۰/۰۲۹	۰/۱۳۵±۰/۰۰۷	۰/۷۳۷±۰/۰۲۵
الگوریتم C4.5	۰/۸۹۲±۰/۰۳۰	۰/۹۲۷±۰/۰۲۹	۰/۰۹۵±۰/۰۰۹	۰/۸۱۳±۰/۰۲۹
درخت کلاس‌بندی رگرسیونی	۰/۸۵۶±۰/۰۳۰	۰/۸۹۴±۰/۰۲۸	۰/۱۱۵±۰/۰۱۰	۰/۷۶۹±۰/۰۳۰

* اعداد متن جدول به صورت انحراف معیار± میانگین گزارش شدند.

بحث

این پژوهش به منظور کاربرد روش‌های آدابوست، بگینگ، الگوریتم C4.5 و درخت کلاس‌بندی رگرسیونی، در یافتن اثرات متقابل عوامل خطر فلج مغزی کودکان صورت گرفت. بیش‌تر محققان تنها از ملاک‌های حساسیت و ویژگی برای سنجش عملکرد پیش‌بینی روش‌های به‌کار برده شده، استفاده نموده‌اند (۹،۱۰). در این مطالعه، مقایسه چهار روش با استفاده از میانگین شاخص‌های حساسیت، ویژگی، خطا و ضریب همبستگی متیو که توسط اعتبارسنجی متقابل ۱۰ قسمتی به دست آمد، انجام شد. بر این اساس، روش آدابوست و الگوریتم C4.5 در کل عملکرد موفق‌تری نسبت به سایر روش‌ها داشتند. در گزارش‌های مختلف اثری از معنی‌دار بودن متغیر ازدواج فامیلی والدین در بروز فلج مغزی کودکان مشاهده نشد و این امر شاید به علت ازدواج‌های فامیلی در کشور ما در مقایسه با سایر کشورها باشد و می‌توان این موضوع را به مبحث ژنتیکی فلج مغزی مرتبط نمود که ازدواج فامیلی والدین، وجود ژن‌های تأثیرگذار بر بروز فلج مغزی کودکان را تشدید می‌کند و بنابراین وجود این عامل در ترکیب متغیرهای معنی‌دار توجیه‌پذیر خواهد شد.

بهبود برازش مدل‌ها با استفاده از این روش‌ها در مطالعات متعددی گزارش شده است. از جمله در سال ۲۰۰۴ میلادی، شواندیر[1] و همکاران عملکرد روش‌های رندم فارست، درخت کلاس‌بندی رگرسیونی، بگینگ، لوجیت بوست، ماشین‌های برداری پشتیبان و k نزدیک‌ترین همسایه را در پیش‌بینی اثرگذاری ۲۵ پلی‌مورفیسم در بیماران مبتلا به سرطان سینه بررسی نمودند. آن‌ها معتقدند که این روش‌ها برای تحلیل پلی‌مورفیسم‌های تک‌نوکلئوتیدی روش‌های مقبولی هستند (۱۱). در سال ۲۰۰۸ میلادی، پومپانگ[2] و همکاران عملکرد روش‌های شبکه‌های بیزی، الگوریتم C4.5، جنگل تصمیم و نایو بیز را در داده‌های مرتبط با نمره‌های درسی دانشجویان برای سنجش برنامه

[1] Schwender H
[2] Pumpuang P

[Downloaded from journals.tums.ac.ir on 2024-06-17]

درسی داده شده، بررسی نمودند. ملاک مقایسه توان پیش‌بینی روش‌ها بود. نایو بیز بهترین عملکرد را در میان روش‌های به کار برده شده داشت (۱۲). در سال ۲۰۰۹ میلادی، هانگ[۱] و همکاران عملکرد پیش‌بینی روش‌های نایو بیز، ماشین‌های برداری پشتیبان و الگوریتم C4.5 را در داده‌های حاصل از ۴۲ پلی‌مورفیسم تک‌نوکلئوتیدی در اختلال خستگی مزمن بررسی نمودند. ملاک‌های مقایسه پیش‌بینی روش‌ها، سطح زیر منحنی، حساسیت و ویژگی بودند. از میان روش‌های به‌کار برده گرفته شده، نایو بیز بهترین عملکرد پیش‌بینی را داشت (۹). در سال ۲۰۱۰ میلادی، اوگوتو[۲] و همکاران عملکرد روش‌های آدابوست، لوجیت بوست، رندم فارست، L_2 بوستینگ، درخت تصمیم و ماشین برداری پشتیبان رادر بررسی جهش بیماری در ۲۶ پلی‌مورفیسم به‌کار بردند. ملاک مقایسه روش‌های معتبرسازی سیستماتیک بود. لوجیت بوست بهترین عملکرد را در پیش‌بینی پلی‌مورفیسم‌های تأثیرگذار بر بیماری داشت (۱۳). هم‌چنین در این سال، کیم[۳] و همکاران عملکرد سه روش رندم فارست، رگرسیون منطقی و رگرسیون خطی ساده را روی داده‌های شبیه سازی شده بررسی نمودند. ادعای آن‌ها این است که هیچ‌کدام از این سه روش بهترین عملکرد را نداشته و هر روش بسته به نوع داده می‌تواند عملکرد متفاوتی داشته باشد (۱۵). در سال ۲۰۱۳ میلادی، یائو[۴] و همکاران عملکرد پیش‌بینی روش‌های رندم فارست، MARS، روش‌های ترکیبی از این دو روش، الگوریتم C4.5 و ماشین بردار پشتیبان را در پیش‌بینی سرطان سینه با استفاده از ۳۲ متغیر بررسی نمودند. ملاک‌های مقایسه روش‌ها حساسیت ویژگی و دقت بودند. روش ترکیبی از رندم فارست و MARS بهترین عملکرد پیش‌بینی را داشت (۱۰).

این مطالعه جزء نخستین مطالعه‌هایی است که اقدام به شناسایی اثرات متقابل، مرتبه ۲ و بالاتر، عوامل خطر فلج مغزی کودکان با استفاده از روش‌های کلاس‌بندی رگرسیونی نموده است. این امر اطلاعات مفیدی در اختیار درمان‌گران قرار می‌دهد و پژوهش‌گران می‌توانند با عوامل خطر متعدد و نمونه‌های بیش‌تر مطالعات تکمیلی را انجام دهند.

نتیجه‌گیری

در بیش‌تر مطالعات انجام شده در خصوص فلج مغزی، گزارشی از شناسایی اثرات متقابل وجود نداشت. بدیهی است که اگر تعداد متغیرهای مستقل زیاد باشد تعیین اثرات متقابل می‌تواند در عمل مشکل یا حتی غیرممکن باشد. یافته‌های این پژوهش نشان داد که روش‌های آدابوست، بگینگ، الگوریتم C4.5 و درخت کلاس‌بندی رگرسیونی برای شناسایی اثرات متقابل کارامد بوده و روش ادابوست عملکرد بهتری داشت و بنابراین پیشنهاد می‌شود. پژوهش‌گران می‌توانند با به‌کارگیری این روش و سایر روش‌های گفته شده مواردی از اثرهای متقابل میان عوامل مؤثر بر یک حالت یا بیماری مانند فلج مغزی کودکان را شناسایی و مبتنی بر شواهد بالینی دسته‌بندی و گزارش نمایند.

تشکر و قدردانی

این مقاله برگرفته از طرح پژوهشی مصوب به شماره ۱۱۵۹۰ در معاونت تحقیقات و فناوری دانشگاه علوم بهزیستی و توان‌بخشی است که بدینوسیله از این معاونت تشکر و قدردانی می‌شود.

[۱] Huang LC
[۲] Ogutu J
[۳] Kim Y
[۴] Yao D

[Downloaded from journals.tums.ac.ir on 2024-06-17]

منابع

1. Villiams F. The Culture of exceptional children descriptive. (Translated by Behpajoh A, Alizadeh A, Yadegari F, Yousefi M) 1st ed. Tehran: Be'sat; 1996.
2. Kauffman JM, Halahan D. Exceptional children an introduction to special education. (Translated by Javadian M, Akhlaghi H and Saliani A) 2nd ed. Khorasan Razavi: Astan-e Qods Razavi Foundation Publishing; 1993.
3. Eliasson AC, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall AM, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006; 48: 549-54.
4. Razaviafzal ZS. A Survey on knowledge and application of Iranian caregivers regarding special care of children 1-5 years old with cerebral palsy. M.Sc. thesis, Occupational therapy group, Tehran: University of Social Welfare and Rehabilitation Science (USWRS); 2011.
5. Rogers SL. Common conditions that influence children's participation. In J. Case-Smith (Ed.), Occupational therapy for children. St. Louis, MO: Elsevier Mosby. 2005. P. 160-215.
6. Soleimani F, Vameqi R, Hemati S, Biglarian A, Surtchi H. A study types of cerebral palsy and associated disorders in Tehran northern and eastern. J Rehabilitation. 2011; 73: 72-79.
7. Marcdante K, Kliegman RM, Jenson HB. Nelson essentials of pediatrics, 6th ed. Philadelphia, PA: Saunders; 2011.
8. Soleimani F, Karimi H, Soleimani N, Biglarian A. A study of 1-6 years old children with Cerebral palsy in Asma Rehabilitation center in terms of type, associated disorders and perinatal and neonatal risk factors on 2007-2008. Deputy of researches and technology, University of Social Welfare and Rehabilitation Sciences (USWRS); Final report of the research: 2009.
9. Huang LC, Hsu SY, Lin E. A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data. J Transl Med. 2009; 7:81.
10. Yao D, Yang J, Zhan X. A Novel Method for Disease Prediction: Hybrid of Random Forest and Multivariate Adaptive Regression. Splines Journal of Computers. 2013; 8: 170-7.
11. Schwender H, Zucknick M, Ickstadt K, Bolt HM. A pilot study on the application of statistical classification procedures to molecular epidemiological data. Toxicol Lett. 2004; 151: 291-9.
12. Pumpuang P, Srivihok A, Praneetpolgrang P, editors. Comparisons of classifer algorithms: Bayesian network,C4.5, decision forest and NBTree for Registration Planning model of undergraduate students(Electronic versiuon). Conference on systems. Man and Cybernetics. 2008: 3647-51.
13. Ogutu JO, Piepho HP, Streeck TS. Comparison of Random Forest, Boosting and Support Vector Machines for Genomic Selection. BMC Bioinformatics. 2011; 5: S11.
14. Kim Y, Li Q, Cropp CD, Sung H, Cai J, Simpson CL, et al. Performance of Random Forest and Logic regression methods using mini-exome sequence data. BMC Boinformatics. 2011; 5: S104.

[Downloaded from journals.tums.ac.ir on 2024-06-17]

Iranian Journal of Epidemiology 2014; 10(2): 33-39.

Original Article

Classification Regression Methods in Identification of the Interactive Effects of the Risk Factors of Cerebral Palsy in Children

Zare Delavar S[1], Bakhshi E[2], Soleimani F[3], Biglarian A[4]

1- Student of M.Sc. in Biostatistics, Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
2- Assistant Professor, Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
3- Associate Professor, Pediatric Neurorehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
4- Assistant Professor, Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
Corresponding author: Biglarian A., abiglarian@uswr.ac.ir

Background & Objectives: The identification of risk factors and their interactions is important in medical studies. The aim of this study was to identify the interaction of risk factors of cerebral palsy in 1-6 years-old children with classification regression methods.

Methods: The data of this cross-sectional study which was conducted on 225 children aged 1-6 years was collected during 2008- 2009. Classification regression methods (classification and regression tree (CART), adapting boosting (AdaBoost), bagging, and C4.5 algorithm) were used to identify interactions between risk factors. Data analysis was carried out with R3.0.1 software.

Results: The identified interactions of the factors by a) the AdaBoost method were (consanguinity: sex, previous pregnancies: vaginal delivery, consanguinity: sex: preterm, history of the disease: preterm: asphyxia, consanguinity: sex: asphyxia, history of the disease: sex: small size relative to gestational age, neonatal infection: asphyxia: small size relative to gestational age, history of the disease: sex: asphyxia, preterm: asphyxia: vaginal delivery); by b) the bagging method were (consanguinity: asphyxia, consanguinity: preterm: asphyxia), by c) the C4.5 algorithm were (asphyxia: preterm, asphyxia: consanguinity: history of the disease: preterm), and by d) the CART method were (asphyxia: consanguinity). The sensitivity and specificity of the AdaBoost method was better than other methods (0.941 ± 0.029 and 0.951 ± 0.030, respectively).

Conclusion: The AdaBoost method could better recognize and model potential interactions between risk factors of cerebral palsy.

Keywords: Classification and regression tree, Cerebral palsy, Children, Interaction

[Downloaded from journals.tums.ac.ir on 2024-06-17]