

# ارتباط‌سنجی وضعیت اجتماعی - اقتصادی بیماران سوختگی و حیات با استفاده از

## تکنیک‌های داده‌کاوی

معصومه خدری<sup>۱\*</sup>، مرجان قاضی سعیدی<sup>۱</sup>، عباس شیخ طاهری<sup>۲</sup>، جواد زارعی<sup>۳</sup>

۱. گروه مدیریت اطلاعات سلامت، دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران.
۲. گروه مدیریت اطلاعات سلامت، دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران.
۳. گروه فناوری اطلاعات سلامت، دانشکده پیراپزشکی، دانشگاه علوم پزشکی جندی‌شاپور اهواز، اهواز، ایران.

مجله اطلاع‌رسانی پزشکی نوین؛ دوره سوم؛ شماره دوم؛ پاییز و زمستان ۱۳۹۶؛ صفحات ۲۵-۱۸

### چکیده

**هدف:** سوختگی در تمام جوامع یک چالش عمده بهداشتی است و با توجه به اینکه داده‌کاوی علم جدیدی است این مطالعه با هدف ارتباط‌سنجی وضعیت اجتماعی-اقتصادی بیماران سوختگی و حیات، با استفاده از تکنیک‌های داده‌کاوی انجام گرفت. تا رابطه معنادار بین وضعیت اقتصادی و اجتماعی افرادی که دچار سوختگی می‌شوند با نتیجه درمان آن‌ها را مشخص می‌کند.

**روش‌ها:** پژوهش حاضر یک مطالعه توصیفی تحلیلی و به صورت گذشته‌نگر است. جمعیت مورد مطالعه ۵۵۳ نفر شامل تمامی بیماران سوختگی بزرگسال بستری شده در بیمارستان آیت‌الله طالقانی اهواز در سال‌های ۱۳۹۰-۱۳۸۸ می‌باشد. داده‌ها از طریق چک‌لیست و پرسشنامه جمع‌آوری گردید و با استفاده از نرم‌افزار IBM SPSS Modeler و از طریق الگوریتم‌های داده‌کاوی C5.0 و CHAID و C&R مورد تجزیه و تحلیل قرار گرفتند. جهت ارزیابی مدل‌ها از ماتریس سردرگمی استفاده شد.

**نتایج:** مدل پیش‌بینی برای پیامد سوختگی بیماران که توسط الگوریتم‌های منتخب به دست آمد شامل عامل سوختگی، شغل، کپسول در خانه، منطقه سکونت، میزان درآمد، قصد، جنس و تحصیلات می‌باشند. در مقایسه صحت الگوریتم‌ها، بیشترین صحت مربوط به الگوریتم C&R (۰/۸۷)، بیشترین ویژگی مربوط به الگوریتم CHAID (۰/۹۸) و بیشترین حساسیت مربوط به الگوریتم C&R (۰/۵۰) می‌باشد. با توجه به برتری مدل C&R در مقایسه صحت و حساسیت مدل‌ها، این مدل به عنوان مدل برتر شناخته شد.

**نتیجه‌گیری:** با توجه به میانگین صحت مدل‌های پیشنهادی، مدل‌های داده‌کاوی C5.0 و CHAID و C&R معتبر هستند. نتایج این مطالعه می‌تواند برای پیش‌بینی تأثیر عوامل اجتماعی-اقتصادی بر روی پیامد سوختگی مفید باشد.

**کلیدواژه‌ها:** سوختگی‌ها، وضعیت اقتصادی-اجتماعی، داده‌کاوی، بیماران.

### نوع مقاله: پژوهشی

دریافت مقاله: ۹۵/۱۱/۱۱ اصلاح نهایی: ۹۶/۱۲/۲۲ پذیرش مقاله: ۹۶/۱۲/۲۵

ارجاع: خدری معصومه، قاضی سعیدی مرجان، شیخ طاهری عباس، زارعی جواد. ارتباط‌سنجی وضعیت اجتماعی - اقتصادی بیماران سوختگی و حیات با استفاده از تکنیک‌های داده‌کاوی. مجله اطلاع‌رسانی پزشکی نوین. ۱۳۹۶؛ ۳(۲): ۲۵-۱۸.

### مقدمه:

قربانی اثر می‌گذارد. اسکار سوختگی موجب بدشکلی می‌گردد و به دنبال آن افسردگی و اختلال استرس پس از حادثه در این بیماران شایع بوده و مشکلات تصویر ذهنی و فقدان اعتماد به نفس و احساس خجالت و گوشه‌گیری نیز قابل توجه است. بیشترین آمار سوختگی در مناطق توسعه نیافته و در حال توسعه دیده می‌شود که بیشتر به دلیل بی‌توجهی و ناآگاهی افراد جامعه و همچنین بی‌توجهی به نکات ایمنی و استانداردهای لازم هنگام نصب وسایل الکتریکی و گازسوز می‌باشد [۲].

سوختگی عبارت است از تخریب بخشی یا همه لایه‌های سلولی پوست در اثر برخورد با مایعات داغ یا مواد جامد [۱]. در حال حاضر سوختگی یکی از اصلی‌ترین موضوعات پزشکی است. سوختگی در تمام جوامع یک چالش عمده بهداشتی است و بیش از هر نوع ترومای دیگر برای بیمار و خانواده بیمار اختلال جسمی، روانی و بار بهداشتی و اقتصادی به دنبال دارد. سوختگی از حوادثی است که در جسم و روحیه

نویسنده مسئول:

معصومه خدری

دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران.

تلفن: +۹۸۹۱۶۳۵۶۲۷۰ پست الکترونیکی: masikhedri92@gmail.com

نوع، میزان و عوامل ایجادکننده سوختگی‌ها با توجه به گوناگونی سبک زندگی در مناطق مختلف جهان متفاوت است. بررسی‌های انجام شده حاکی از آن است که مواردی مانند سبک زندگی، سطح اجتماعی، اقتصادی و فرهنگی جامعه و نیز منبع انرژی مورد استفاده در محیط کار و زندگی می‌تواند میزان و نوع سوختگی‌ها را تغییر دهد [۳]. به طور کلی از مهم‌ترین عوامل اجتماعی اقتصادی تأثیرگذار در بروز سوختگی می‌توان به بیکاری، کیفیت زندگی، محل سکونت، نامناسب بودن شغل، تحصیلات کم، درآمد کم خانواده و غیره اشاره کرد. با افزایش توان اقتصادی و موقعیت اقتصادی اجتماعی افراد نه تنها از شیوع آسیب‌ها کاسته می‌شود بلکه شدت این آسیب‌ها نیز به طور متوسط کاهش می‌یابد [۴].

براساس آمار کمیته مرکزی اورژانس در میان ۱۵۰ علت شایع مرگ، سوختگی در رده ششم قرار دارد. سازمان پزشکی قانونی کشور با اعلام اینکه در سه ماهه نخست سال ۹۰، ۵۵۲ نفر که ۲۷۵ نفر زن و ۲۷۷ نفر مرد بودند، بر اثر سوختگی جان خود را از دست دادند گزارش کرد استان تهران با ۸۰ نفر، خوزستان با ۶۲ نفر و خراسان رضوی با ۵۳ تن بالاترین آمار مرگ ناشی از سوختگی را در اختیار دارند. آمار سازمان پزشکی قانونی درباره تعداد افراد فوت شده بر اثر سوختگی در کشورمان، نشان‌دهنده این است که مرگ و میر ناشی از سوختگی در ایران بسیار بالا است [۵]. حدود ۹۰ درصد از سوختگی‌ها در کشورهای با سطح اقتصادی متوسط و پایین رخ می‌دهد. براساس آمار موجود، بیشترین میزان سوختگی در شرق آسیا و کم‌ترین میزان مربوط به امریکا می‌باشد [۴]. داده‌ها عناصر تشکیل‌دهنده ساختمان دانش محسوب می‌شوند و پژوهشگران ناگزیر به ایجاد مفاهیم و مدل‌های جدید هستند تا نظم موجود بین آن‌ها را کشف کرده و در نهایت به دانش دست یابند [۶]. بنابراین محققان از شاخه‌ای تحت عنوان کشف دانش از پایگاه داده استفاده می‌کنند که متشکل از تئوری‌ها و ابزارهایی جهت تحلیل مقادیر زیاد داده‌ای و استخراج دانش پنهان آن‌هاست [۷]. واضح است که برای دسترسی به اطلاعات کاربردی موجود در حجم زیاد داده‌ها، روش‌های دستی و سنتی مناسب نمی‌باشد. در عوض به کارگیری روش‌های کشف دانش برای شناسایی الگوهای داخل داده‌ها و در نهایت دسترسی به اطلاعات کاربردی، کاملاً مناسب است [۸]. از آنجایی که داده‌های پزشکی نیز از نوع اداری، مالی، دموگرافی و بالینی حجیم می‌باشند، تحلیل و استخراج اطلاعات با معنی از انبوه این داده‌های خام، پیچیده و سخت می‌شود [۹]. بنابراین داده‌کاوی که علم رو به رشد و جدیدی است بر روی داده‌های پزشکی از اهمیت بالایی برخوردار است [۱۰]. از این رو این

### مواد و روش‌ها:

تحقیق حاضر از لحاظ هدف توصیفی تحلیلی است. در پژوهش حاضر جمعیت مورد مطالعه تمامی بیماران سوختگی بزرگسال بستری شده در بیمارستان آیت‌الله طالقانی شهر اهواز در سال‌های ۱۳۸۸-۱۳۹۰ است که تعداد ۵۵۳ بیمار را شامل می‌شود که در نهایت پرسشنامه تکمیل شده ۵۴۶ بیمار مورد بررسی قرار گرفت. معیار ورود به پژوهش وجود صدمه سوختگی در زمان پذیرش بود و پرونده‌های بیماران ترمیمی بستری شده به دلیل عوارض سوختگی قبلی مانند اسکار، مورد مطالعه قرار نگرفت. همچنین در این پژوهش بیماران سوختگی اطفال به دلیل لزوم همگن بودن جامعه آماری و متفاوت بودن علل و عوامل سوختگی در اطفال و بزرگسالان، مورد بررسی قرار نگرفتند و مطالعه تنها بر روی موارد سوختگی در بزرگسالان صورت پذیرفت. در این پژوهش نمونه‌گیری به روش سرشماری بود. بدین معنی که کلیه موارد در جامعه مورد بررسی قرار گرفتند و نمونه‌گیری انجام نشد.

استان خوزستان از جمله استان‌هایی است که به لحاظ وجود بنادر، پالایشگاه‌ها و صنایع بزرگ از جمله فولاد، نورد لوله، نیشکر و وجود خطوط انتقال نفت و گاز که تقریباً کل استان را احاطه کرده است از جمله نقاط حادثه‌خیز در زمینه آتش‌سوزی و سوختگی محسوب می‌شود. این بیمارستان نیز تنها مرکز تخصصی و فوق تخصصی سوختگی در جنوب غرب کشور بوده و علاوه بر ارائه خدمات به بیماران سوختگی استان، به بیماران دو استان هم‌جوار (لرستان و ایلام) خدمات ارائه می‌دهد. بنابراین به خاطر تنوع داده‌های بیماران از مکان‌ها و فرهنگ‌های مختلف، مدل ایجاد شده می‌تواند برای انواع مختلف بیماران بکار رود و در واقع از قابلیت تعمیم‌پذیری نسبتاً بالایی، برخوردار باشد.

در مرحله تعیین متغیرها، برای تعیین عوامل اجتماعی اقتصادی احتمالی و شناخته‌شده‌ی تأثیرگذار بر سوختگی، ابتدا متون و مقالات موجود در پایگاه‌های داده SID، Google Scholar، PubMed و Science Direct براساس کلیدواژه‌های سوختگی، وضعیت اجتماعی اقتصادی و داده‌کاوی مورد بررسی قرار گرفت [۱۱-۱۸]. سپس با استفاده از بررسی مقالات مربوطه، عوامل اجتماعی اقتصادی تأثیرگذار بر سوختگی در قالب سه بخش که شامل الف) ویژگی‌ها و عوامل

ارزیابی مدل‌ها و بررسی صحت آن‌ها، داده‌ها به دو دسته آموزش (۷۰ درصد) برای ساخت روش دسته‌بندی و داده‌های آزمایش (۳۰ درصد) برای آزمون مدل تقسیم شدند. در این پژوهش از ماتریس سردرگمی به علت سادگی و استفاده رایج در حوزه سلامت، برای ارزیابی مدل‌های به دست آمده استفاده شد [۱۸].

### یافته‌ها:

در یافته‌های مربوط به تعیین عوامل اجتماعی اقتصادی مؤثر بر سوختگی، متغیرها براساس منابع مرتبط شناسایی شدند. متغیرهای دموگرافیک شامل سن، جنس، وضعیت تأهل، منطقه سکونت، میزان درآمد، شغل، تحصیلات و تعداد جمعیت بود. متغیرهای بالینی شامل مدت اقامت، درصد سوختگی، درجه سوختگی، نتیجه‌ی درمان، عامل سوختگی، قصد سوختگی و محل سانحه، و متغیرهای تکمیلی شامل کپسول در خودرو، دوگانه‌سوز بودن خودرو، اقدام به دوگانه‌سوز کردن، کپسول در منزل و تلفن در خانه و اتومبیل در خانه بود.

یافته‌های مربوط به مرحله‌ی مدل‌سازی با استفاده از الگوریتم‌های منتخب C5.0، CHAID و C&R صورت گرفت و پس از ارزیابی مدل‌ها، مدل بهتر انتخاب گردید.

در الگوریتم C5.0 ابتدا با استفاده از گره انتخاب ویژگی متغیرهای مهم تعیین شد. سپس مدل‌سازی با ورود این متغیرها انجام گرفت. متغیرهای مهم به دست آمده از گره انتخاب ویژگی عبارت بودند از: عامل سوختگی، تلفن در خانه، میزان درآمد، کپسول در خانه، شغل، هشدار در خانه، تحصیلات، محل سانحه، جنس، منطقه سکونت و علت سوختگی در محل کار. پس از اجرای الگوریتم و سپری شدن مراحل آموزش و آزمایش روی مجموعه داده‌ها،

نتایج به دست آمده از جدول ۱ نشان داد صحت این الگوریتم برای داده‌های آموزش ۸۳/۹۱ درصد و برای داده‌های آزمایش ۸۳/۲۳ درصد می‌باشد. پس از اجرای الگوریتم، هیچ متغیری شناسایی نشد که نشان از عدم وجود هرگونه رابطه، بین متغیر هدف (پیامد سوختگی) و متغیرهای ورودی (عوامل اقتصادی اجتماعی) باشد.

جدول ۱- ارزیابی عملکرد الگوریتم C5.0 برای داده‌های آموزش و آزمایش

	آموزش		آزمایش	
	تعداد	درصد	تعداد	درصد
صحیح	۳۱۸	۸۳/۹۱	۱۳۹	۸۳/۲۳
اشتباه	۶۱	۱۶/۰۹	۲۸	۱۶/۷۷
مجموع	۳۷۹	۱۰۰	۱۶۷	۱۰۰

دموگرافیک. ب) اطلاعات بالینی مربوط به سوختگی و درمان آن. ج) اطلاعات تکمیلی در خصوص محل سانحه، شناسایی گردید. سپس این عوامل توسط گروه بررسی‌کننده، مورد بررسی قرار گرفتند. لازم به ذکر است که گروه بررسی‌کننده شامل دو گروه کارشناس (در مجموع ۸ نفر) بودند. معیار انتخاب این افراد آشنایی با بیماران سوختگی (به ویژه شرایط اقتصادی و اجتماعی آن‌ها) و نیز داده‌کاوی بیماری‌ها بود.

گروه اول، شامل مددکار، روانشناس بالینی و ریاست بیمارستان هرکدام با سابقه کار حداقل پنج سال در بیمارستان سوانح و سوختگی و دو پزشک متخصص در زمینه سوختگی، ترمیم و جراحی پلاستیک بودند. گروه دوم نیز شامل یک نفر دکترای انفورماتیک پزشکی و دو نفر دکترای مدیریت اطلاعات سلامت عضو هیئت‌علمی با سابقه انجام مطالعه مشابه در مورد داده‌کاوی بیماری‌ها بودند. متغیرهای مورد بررسی از نوع کمی و کیفی بودند و برخی از آن‌ها بر طبق وجود یا عدم وجودشان در بیمار به متغیرهای باینری (صفر و یک) و برخی دیگر براساس بخش‌های مختلفی که داشتند، کددهی شدند.

در مرحله گردآوری داده‌ها، داده‌ها از طریق چکلیست محقق ساخته براساس استانداردهای جهانی و پرسشنامه جمع‌آوری گردید. محقق به صورت حضوری به محیط پژوهش مراجعه و از طریق مطالعه پرونده بیماران و با استفاده از چکلیست طراحی شده براساس جدول متغیرها، به جمع‌آوری و استخراج داده‌ها پرداخت. لازم به ذکر است که متغیرهای سن، جنس، وضعیت تأهل، نوع منطقه سکونت، شغل، عامل سوختگی، درصد سوختگی، درجه سوختگی، مدت اقامت بیمار، نتیجه درمان، قصد حادثه، در پرونده بیمار ثبت بود. سایر متغیرها در پرونده بیمار مستند نبود و به صورت جداگانه بر طبق پرسشنامه، از بیمار یا خانواده وی به صورت تلفنی اخذ گردید. داده‌های جمع‌آوری شده وارد نرم‌افزار اکسل گردید. مجموعه داده این پژوهش شامل ۵۵۴ رکورد از اطلاعات بیماران سوختگی مراجعه‌کننده به بیمارستان آیت‌الله طالقانی شهر اهواز می‌باشد. هر رکورد از این پایگاه داده، دارای ۲۶ فیلد (ویژگی) است. سپس در مرحله پیش‌پردازش، فایل داده‌ها به نرم‌افزار IBM SPSS Modeler نسخه ۱۴/۲ وارد گردید و پس از بررسی‌های کمی و کیفی، مدل‌سازی بر روی داده‌ها انجام شد.

با توجه به اهداف پژوهش، نوع نمونه و متغیرها، از الگوریتم‌های پیش‌بینی‌کننده درخت تصمیم استفاده شد. در این پژوهش از متداول‌ترین و معتبرترین الگوریتم‌های دسته‌بندی در حوزه سلامت و پزشکی یعنی C5.0، CHAID و C&R به علت سادگی و قدرت تفهیم بالا برای ایجاد مدل استفاده شد [۱۹]. در مرحله ارزیابی به جهت امکان

## جدول ۲- ارزیابی عملکرد الگوریتم CHAID برای داده‌های آموزش و

آزمایش			
آموزش		آزمایش	
تعداد	درصد	تعداد	درصد
۳۲۰	۸۴/۴۳	۱۴۳	۸۵/۶
۵۹	۱۵/۵۷	۲۴	۱۴/۳
۳۷۹	۱۰۰	۱۶۷	۱۰۰

نتایج به دست آمده حاصل از اجرای الگوریتم CHAID بر روی متغیرهای مورد مطالعه نیز، در جدول ۳ آورده شده است.

در اجرای الگوریتم CHAID با استفاده از متغیرهای حاصل از انتخاب ویژگی، مدل‌سازی صورت گرفت. متغیرهای مهم شناسایی شده توسط این الگوریتم به ترتیب اهمیت شامل عامل سوختگی، میزان درآمد، منطقه سکونت و وجود کپسول در خانه بود. پس از اجرای الگوریتم و سپری شدن مراحل آموزش و آزمایش روی مجموعه داده‌ها، نتایج به دست آمده در جدول ۲ مربوط به ارزیابی عملکرد الگوریتم نشان می‌دهد صحت این الگوریتم برای داده‌های آموزش ۸۴/۴۳ درصد و برای داده‌های آزمایش ۸۵/۶۳ درصد می‌باشد.

## جدول ۳- نتایج به دست آمده حاصل از اجرای الگوریتم CHAID

متغیر	نتیجه درمان	
	ترخیص (زنده ماندن)	فوت
عامل سوختگی	۸۱٪	۱۹٪
	۹۸٪	۲٪
	۴۵٪	۵۵٪
کپسول در خانه	۸۸٪	۱۲٪
	۸۸٪	۱۲٪
	۷۵٪	۲۵٪
میزان درآمد	۱۰۰٪	۰٪
	۹۱٪	۹٪
منطقه‌ی سکونت	۶۲٪	۳۸٪
	۲۹٪	۷۱٪

همچنین نتایج به دست آمده حاصل از اجرای الگوریتم C&R بر روی متغیرهای مورد مطالعه نیز، در جدول ۵ آورده شده است.

## جدول ۴- ارزیابی عملکرد الگوریتم درخت C&amp;R برای داده‌های آموزش و

آزمایش			
آموزش		آزمایش	
تعداد	درصد	تعداد	درصد
۳۱۹	۸۴/۱۷	۱۴۵	۸۶/۸۳
۶۰	۱۵/۸۳	۲۲	۱۳/۱۷
۷۹	۱۰۰	۱۶۷	۱۰۰

در اجرای الگوریتم C&R متغیرهای حاصل از مرحله انتخاب ویژگی، وارد مدل گردیدند. متغیرهای مهم شناخته شده توسط این الگوریتم، به ترتیب اهمیت شامل عامل سوختگی، شغل، کپسول در خانه، محل سائحه، منطقه سکونت، تلفن در خانه، میزان درآمد، جنس و تحصیلات می‌باشند. پس از اجرای الگوریتم و سپری شدن مراحل آموزش و آزمایش روی مجموعه داده‌ها، نتایج به دست آمده از ارزیابی عملکرد الگوریتم در جدول ۴ ارائه شد. صحت این الگوریتم برای داده‌های آموزش ۸۴/۱۷ درصد و برای داده‌های آزمایش ۸۶/۸۳ درصد می‌باشد.

## جدول ۵- نتایج به دست آمده حاصل از اجرای الگوریتم C&amp;R

متغیر	نتیجه درمان	
	ترخیص (زنده ماندن)	فوت
عامل سوختگی	۶۵٪	۳۵٪
	۹۲٪	۸٪
	۴۷٪	۵۳٪
کپسول در خانه	۷۷٪	۲۳٪
	۷۷٪	۲۳٪
	۸۰٪	۲۰٪
میزان درآمد	۹۱٪	۹٪
	۸۰٪	۲۰٪
منطقه‌ی سکونت	۸۹٪	۱۱٪
	۳۵٪	۶۵٪

جنس، مدت اقامت، تأهل، شغل، تحصیلات، عامل سوختگی، قصد، درصد سوختگی، درجه سوختگی، نتیجه درمان، محل سکونت و محل سانحه با اپیدمیولوژیک سوختگی ارتباط داد که نتایج همراستا با یافته‌های این پژوهش می‌باشد [۱۱].

در مطالعه طبیعی و نخعی نیز یافته‌ها نشان داد عواملی چون جنس، سن و محل سوختگی با پیامد سوختگی ارتباط معناداری دارد. به طوری که میزان مرگومیر در زنان بالاتر از مردان بود. همچنین اکثر سوختگی‌ها در منزل اتفاق افتاده بود که این یافته نیز در راستای نتایج پژوهش حاضر و وجود ارتباط بین محل سانحه با پیامد سوختگی بود. این در حالی است که در این پژوهش سوختگی بین رده‌های سنی ۱۰ تا ۳۰ سال بیشتر بود؛ که با یافته‌های پژوهش حاضر مبنی بر نبود رابطه بین سن با پیامد سوختگی مغایرت داشت [۲۰].

عزیزی و همکاران در مطالعه خود با هدف بررسی عوامل موثر بر مرگومیر بیماران سوختگی با استفاده از الگوریتم داده‌کاوی درخت تصمیم نشان دادند جنس، سن، درصد سوختگی، درجه سوختگی، مدت اقامت بیمار و عامل سوختگی با مرگومیر بیماران سوختگی رابطه معناداری دارد. این نتایج با یافته‌های پژوهش حاضر همخوانی دارد [۲۱]. نتایج حاصل از مرحله مدل‌سازی در پژوهش نشان داد که با استفاده از مدل‌های منتخب، متغیرهای تأثیرگذار بر روی پیامد سوختگی شامل عامل سوختگی، تلفن در خانه، میزان درآمد، کپسول در خانه، شغل، تحصیلات، محل سانحه، جنس و منطقه سکونت می‌باشند؛ که مدل C&R به عنوان مدل برتر شناخته شد.

در مطالعه عاشوری و همکاران دو الگوریتم متفاوت از الگوریتم‌های استنتاج قانون به نام‌های C5.0 و CHAID روی داده‌ها اعمال گردید و صحت مدل‌های تولید شده مورد تایید قرار گرفت. به طوری که یافته‌ها نشان داد صحت مدل‌های C5.0 و CHAID روی داده‌های تحت بررسی به ترتیب ۴۵/۵۲ و ۲۸/۳۸ درصد بود که این نتایج برتری صحت مدل C5.0 را نشان می‌داد. در پژوهش حاضر نیز از الگوریتم‌های دسته‌بندی C5.0 و CHAID جهت ارائه مدل استفاده شد؛ که صحت الگوریتم CHAID بیشتر از الگوریتم C5.0 بود؛ که با نتایج به دست آمده از این مطالعه مغایرت دارد [۲۲].

نتایج پژوهش حاضر نشان داد که در استان خوزستان، اغلب متغیرهای اقتصادی اجتماعی که در این پژوهش مورد بحث قرار گرفتند، بر روی پیامد سوختگی (زنده یا فوت نمودن بیماران سوختگی) تأثیرگذار هستند. با توجه به میانگین صحت، ویژگی و حساسیت مدل‌های پیشنهادی، این مدل‌ها معتبر و قابل استناد هستند. در واقع نتایج این

در یافته‌های مربوط به مرحله ارزیابی مدل‌ها، مدل‌های به دست آمده براساس مؤلفه‌های صحت، ویژگی و حساسیت ارزیابی و مقایسه شدند. ارزیابی انجام شده، با استفاده از داده‌های حاصل از ماتریس اغتشاش به دست آمد. در مقایسه صحت الگوریتم‌ها، بیشترین صحت مربوط به الگوریتم C&R (۰/۸۷) می‌باشد. الگوریتم CHAID (۰/۸۶) در رده بعدی قرار گرفت و برای الگوریتم C5.0 هم مدل هیچ نوع خروجی ارائه نداد. در مقایسه ویژگی الگوریتم‌ها، بیشترین ویژگی مربوط به الگوریتم CHAID (۰/۹۸) می‌باشد. الگوریتم C&R (۰/۹۴) در رده بعدی قرار گرفت و برای الگوریتم C5.0 هم مدل هیچ نوع خروجی ارائه نداد. در مقایسه حساسیت الگوریتم‌ها، بیشترین حساسیت مربوط به الگوریتم C&R (۰/۵۰) می‌باشد. الگوریتم CHAID (۰/۲۱) در رده بعدی قرار گرفت و برای الگوریتم C5.0 هم مدل هیچ نوع خروجی ارائه نداد و با توجه به برتری مدل C&R در مقایسه صحت و حساسیت مدل‌ها، این مدل به‌عنوان مدل برتر شناخته شد.

## بحث و نتیجه‌گیری:

پس از بررسی مقالات در پایگاه‌های معتبر Google SID، PubMed Scholar و Science Direct مهم‌ترین متغیرهای شناسایی شده بر وقوع سوختگی از این مقالات استخراج شد و سپس بر طبق نظر کارشناسان مربوطه، متغیرهای دیگری نیز به این مجموعه اضافه گردید که شامل متغیرهای سن، جنس، وضعیت تأهل، نوع منطقه، محل بروز سانحه، شغل، تحصیلات بیمار، عامل سوختگی، میزان درآمد ماهیانه بیمار، تعداد جمعیت خانواده، قصد، درصد سوختگی، درجه سوختگی، مدت اقامت بیمار، نتیجه درمان، وجود کپسول اطفا حریق در منزل، وجود سیستم هشدار آتش‌سوزی در منزل، وجود کپسول اطفا حریق در خودرو، دوگانه‌سوز بودن خودرو، اقدام به دوگانه‌سوز کردن خودرو، علت سوختگی در محل کار، داشتن تلفن در خانه هنگام سوختگی، داشتن اتومبیل در خانه هنگام سوختگی، نوع وسایل گرمایشی، نوع وسیله پخت و پز و قومیت می‌باشند. مطالعات گوناگونی در زمینه بررسی سوختگی‌ها و عوامل خطر مرتبط با آن صورت گرفته است. اغلب مطالعات به بررسی تأثیر چند عامل محدود و مشترک در اغلب پژوهش‌ها و با استفاده از روش‌های آماری پرداخته‌اند. به نظر می‌رسد مطالعه جامعی برای بررسی عوامل اقتصادی اجتماعی مؤثر بر سوختگی با استفاده از تکنیک‌های داده‌کاوی، صورت نگرفته است.

در مطالعه‌ای که توسط مقرب و همکاران با هدف بررسی اپیدمیولوژیک بیماران سوختگی انجام شد، نشان داد عواملی چون سن،

می‌بایست اخذ می‌شد اشاره کرد. جهت رفع محدودیت اول بعضی از متغیرها به صورت جداگانه و بر طبق پرسشنامه طراحی شده، از بیمار یا خانواده وی به صورت تلفنی اخذ گردید و برای رفع محدودیت دوم پژوهشگر با توضیحات کامل و جامع در ارتباط با پژوهش، بیمار یا خانواده وی را متقاعد به ارائه اطلاعات می‌کرد.

### تشکر و قدردانی:

این مطالعه بخشی از پایان‌نامه کارشناسی ارشد رشته فناوری اطلاعات سلامت با عنوان ارائه الگوی ارتباط سنجی وضعیت اجتماعی اقتصادی بیماران سوختگی با استفاده از تکنیک‌های داده‌کاوی می‌باشد که در دانشگاه علوم پزشکی تهران به تصویب رسیده است. بدین‌وسیله از کارشناسان و مسئول بخش مدارک پزشکی بیمارستان آیت‌الله طالقانی شهر اهواز و نیز ریاست بیمارستان که ما را در انجام پژوهش یاری نمودند، تشکر و قدردانی می‌شود.

مطالعه یک الگوی بومی جامع، دقیق و معتبر برای پیش‌بینی تأثیر عوامل اجتماعی اقتصادی بر روی پیامد سوختگی را براساس متغیرهای فوق‌الذکر که بعضی از آن‌ها با نتایج سایر پژوهشگران مطابقت دارد، پیشنهاد می‌کند. بنابراین چنین مدلی می‌تواند به عنوان یک معیار مهم ارزیابی جهت پیش‌بینی نتیجه درمان بیماران سوختگی مورد استفاده قرار گیرد. در نهایت پیشنهاد می‌گردد که از الگوریتم‌های مورد استفاده در این پژوهش، جهت پیش‌بینی تأثیر عوامل اجتماعی اقتصادی بر فوت یا زنده ماندن بیماران سوختگی سایر استان‌های کشور، به منظور ارائه یک مدل ملی استفاده گردد. همچنین به نظر می‌رسد مقایسه مؤلفه‌های ارزیابی مدل‌ها در حوزه‌های گوناگون بخصوص در زمینه پیش‌بینی بیماری‌ها و حوادث، به دلیل اختلاف در ماهیت و پیچیدگی موضوع، معیار مناسبی برای برتری یک الگوریتم نسبت به الگوریتم دیگر نیست. همچنین از محدودیت‌های این پژوهش می‌توان به نبودن تمام متغیرهای مورد بررسی در پرونده‌ها برای تکمیل اطلاعات مورد نیاز و نیز عدم همکاری بیمار یا خانواده وی در پاسخگویی به مواردی که از طریق تماس تلفنی

### References

1. Basil A, Pruitt J, Cleon w, Goodwin A. Epidemiology of burn. In: Herendon D. Total burn care. London: Saunders; 2002.
2. Abbaspoor R. Coding quality compared with the case of burn injuries and death in Sina Hospital, Tabriz 2010-2011. [Dissertation]. Tehran: Tehran University of Medical Science; 2011. [In Persian]
3. Milner SM, Mottar R, Smith CE. The burn wheel: An innovative method for calculating the need for fluid resuscitation in burned patients. Am J Nurs. 2001; 101(11):35-7.
4. Lari AR, Alaghebandan R, Nikui R. epidemiological study of 3341 burns patients during three years in Tehran, Iran. Burns. 2000; 26(1):49-53.
5. Jonsdottir T, Hvanberg ET, Sigurdsson H, Sigurdsson S. The feasibility of constructing a predictive outcome model for breast cancer using the tools of data mining. Expert Syst Appl. 2008; 34(1):108-18.
6. Razavi A. Applications of knowledge discovery in quality registries - predicting recurrence of breast cancer and analyzing non-compliance with a clinical guideline. Sweden: LiU-Tryck, Linköping; 2007.
7. Richards G, Rayward Smith VJ, Sönksen PH, Carey S, Weng C. Data mining for indicators of early mortality in a database of clinical records. Artif Intell Med. 2001; 22(3):215-31.
8. Thongkam j, Xu G, Zhang Y, Huang F. Toward breast cancer survivability prediction models through improving training space. Expert Syst Appl. 2009; 36(10):12200-9.
9. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: A comparison of three data mining method. Artif Intell Med. 2005; 34(2):113-27.
10. Aghakhani k, Mohammadi S, Molanaei A, Memarian A, Ameri M. Epidemiologic study of scald burns in victims in Tehran burn hospital. Tehran Univ Med J. 2013; 71(7):452-7. [In Persian]
11. Mogharab M, Sabzehkar F, Sharifzadeh G, Azani M. An epidemiological study of hospitalised patients with burns in Imam Reza Hospital in Birjand between 2007 and 2013. J Birjand Univ Med Sci. 2014; 21(2):228-36. [In Persian]
12. Kashfi M, Yazdankhah M, Khani Jeihooni A, Karimi M. Evaluating the frequency of self-immolation and its relationship with social and demographic status of the patients referring to Ghotboddin E Shirazi during the years 2006 and 2011. J Fasa Univ Med Sci. 2015; 4(4):392-401. [In Persian]
13. Mistry RM, Pasisi L, Chong S, Stewart J, Wong She RB. Socioeconomic deprivation and burns. Burns. 2010; 36(3):403-8.
14. Al Qattan MM, Al Zahrani K. A review of burns related to traditions, social habits, religious activities,

- festivals and traditional medical practices. *Burns*. 2009; 35(4):476-81.
15. Atiyeh BS, Costagliola M, Hayek SN. Burn prevention mechanisms and outcomes: Pitfalls, failures and successes. *Burns*. 2009; 35(2):181-93.
  16. Park JO, Shin SD, Kim J, Song KJ, Peck MD. Association between socioeconomic status and burn injury severity. *Burns*. 2009; 35(4):482-90.
  17. Faramarzi H, Bagheri P, Mohammadi A, Hadizadeh E. Epidemiology of burn in Fars Province in 2009. *IRJE*. 2012; 8(2):54-64. [In Persian]
  18. Saniee Abadeh M, Mahmoudi S, Taherparvar M. Data mining application. Tehran: Niaz Danesh. 2012. [In Persian]
  19. Dormohammadi S, Alizadeh S, Asghari M, Shami M. Proposing a prediction model for diagnosing causes of infertility by data mining algorithms. *JHA*. 2014; 17(57):46-57. [In Persian]
  20. Tabiee S, Nakhaee M. Epidemiology of burn patients in Emam Reza Hospital, Birjand, 1998–2002. *J Shahrekord Univ Med Sci*. 2004; 6(1):43-51. [In Persian]
  21. Azizi A, Zarei J, Nabovati E, Vakili H, Abbasi E, Razavi A. Determining of the factors affecting mortality in burn patients using a decision tree data mining algorithm. *JHA*. 2014; 16(54):34-45. [In Persian]
  22. Ashuri M, Najimoghdam V, Alizadeh S, Safi M. Classification and clustering algorithm application for prediction of tablet numbers: Case study diabetes disease. *Health Inf Manage*. 2013; 10(5):739-49. [In Persian]

## Relationship between socioeconomic status and burn patients survive using data mining

Maasumeh Khedri<sup>1\*</sup> Marjan Ghazi Saeedi<sup>1</sup> Abbas Sheikhtaheri<sup>2</sup> Javad Zarei<sup>3</sup>

1. Department of Health Information Management, Faculty of Paramedicine, Tehran University of Medical Sciences, Tehran, Iran.

2. Department of Health Information Management, Faculty of Management and Medical Information, Iran University of Medical Sciences, Tehran, Iran.

3. Department of Health Information Technology, School of Para Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.

(Received 30 Jan, 2017

Accepted 16 Mar, 2018)

### Original Article

### Abstract

**Aim:** In all societies, burn is a health cause's physiologic mental economic and health disorder both for the patient and its family challenge and more than any other trauma. Cases such as lifestyle, social, economic, and cultural levels of society can change the type of burns. Given that data mining is growing and new, it is also of great importance to medical data. Therefore, in this project, the intention is to collect data using actual data mining techniques and in view of the prevalence of burns in developing countries, a model for establishing a meaningful relationship between the economic and social status of individuals develop burns as a result of their treatment.

**Methods:** The present study is analytic-descriptive based on retrospective nature. The population of the study consists of 553 patients all adult burn patients who were hospitalized in Ayatollah Taleghani hospital in Ahvaz for 3 years (1388-1390). data have been collected from check lists and questionnaire. The collected data were computed using spss22, IBM Modeler 14.2 and algorithm of C5.0, CHAID and CART.

**Results:** The predicted model for the outcome of burn on patients conducted through selected algorithms in order of priority including factors of burn, occupation, capsule at home, place of incident, residential area, and rate of income, intent, gender and education. By comparing the accuracy of the algorithms, the highest accuracy is for the C & R algorithm (0.87), the most characteristic of the CHAID algorithm (0.98) and the most sensitivity was obtained for the C & R algorithm (0.50). Given the superiority of the C & R model, this model was recognized as the top model in terms of the accuracy and sensitivity of the models.

**Conclusion:** Due to the average accuracy of suggested models, these models are both valid and attributable. The results of this study can be useful for the prediction of the impact of socioeconomic factors on burnout outcomes.

**Key Words:** Burns, Socioeconomic Status, Data Mining, Patients.

**Citation:** Khedri M, Ghazi Saeedi M, Sheikhtaheri A, Zarei J. Relationship between socioeconomic status and burn patients survive using data mining. J Mod Med Info Sci. 2017; 3(2):18-25.

Correspondence:

Maasumeh Khedri

Faculty of Paramedicine, Tehran University of Medical Sciences, Tehran, Iran.

Tel: +989166356270

Email: masikhedri92@gmail.com