تعیین عوامل مؤثر بر مرگ و میر بیماران سوختگی با استفاده از الگوریتم داده‌کاوی درخت تصمیم

امیرعباس عزیزی[1]/ جواد زارعی[2]/ احسان نبوتی[1]/ حسن وکیلی ارکی[1] / ابراهیم عباسی[3]/ امیررضا رضوی[4]

چکیده

مقدمه: در سال‌های اخیر استفاده از روش‌های داده‌کاوی روی حجم زیادی از داده‌ها با هدف تولید مدل‌ها و الگوهای پیش‌بینی کننده در حیطه‌های متعدد پزشکی رواج یافته است. در این مطالعه پژوهشگران قصد دارند با استفاده از الگوریتم داده‌کاوی درخت تصمیم، عوامل موثر بر مرگ‌ومیر بیماران سوختگی را مشخص نمایند.

روش کار: این پژوهش توصیفی گذشته‌نگر است که مبتنی بر اطلاعات پرونده‌های بیمارستانی می‌باشد. در مجموع ٤٨٠٤ پرونده پزشکی مربوط به بیماران سوختگی، مورد بررسی قرار گرفت. داده‌های گردآوری شده با استفاده از نرم افزارهای تحلیل آماری SPSS-16 و داده‌کاوی Clementine-12 و از طریق الگوریتم (Chi-squared Automatic Interaction Detection: CHAID) مورد تجزیه و تحلیل قرار گرفتند.

یافته‌ها: مدل پیش‌بینی فوت و زنده‌ماندن برای بیماران سوختگی که کارایی آن توسط پژوهشگران اثبات گردید، به ترتیب اهمیت شامل درصد سوختگی، درجه سوختگی، مدت زمان بستری، جنسیت و سن می‌باشد. سایر متغیرهای مورد مطالعه شامل کشت خون، کشت زخم، کشت ادرار، عامل سوختگی و ماه بستری بیمار در کارایی مدل تاثیری نداشتند.

نتیجه‌گیری: با توجه به میانگین دقت مدل پیشنهادی(تقریباً ۹۵درصد) و میانگین دقت مدل روی داده‌های آزمون(تقریباً ۹۵درصد)، مدل پیشنهادی معتبر و قابل استناد می‌باشد. در حقیقت، نتایج این مطالعه یک مدل جامع، دقیق و معتبر پیش‌بینی فوت و زنده‌ماندن بیماران سوختگی را بر اساس متغیرهای فوق‌الذکر که بعضی از آن‌ها با نتایج سایر پژوهشگران مطابقت دارد پیشنهاد می‌کند. لذا چنین مدلی می‌تواند به عنوان یک معیار مهم ارزیابی جهت درمان موثر بیماران سوختگی مورد استفاده قرار بگیرد.

کلیدواژه‌ها: داده‌کاوی، درخت تصمیم، سوختگی

•وصول مقاله: ۹۲/۲/۱۸ •اصلاح نهایی:۹۲/۷/۲۷ •پذیرش نهایی:۹۲/۸/۲۹

۱. دانشجوی دکتری تخصصی انفورماتیک پزشکی، کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۲. دانشجوی دکتری تخصصی مدیریت اطلاعات سلامت، مرکز تحقیقات علوم مدیریت و اقتصاد سلامت، دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران

۳. دانشجوی دکتری تخصصی انفورماتیک پزشکی، عضو گروه پژوهشی انفورماتیک سرطان، مرکز تحقیقات سرطان پستان جهاد دانشگاهی، تهران، ایران

۴. استادیار گروه انفورماتیک پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران؛ نویسنده مسئول (razaviar@mums.ac.ir)

مقدمه

سوختگی آسیب بافتی در اثر تماس حرارتی، شیمیایی یا الکتریکی است که منجر به تغییر پروتئینی، اِدم و کاهش حجم مایعات داخل عروقی می‌شود [۱]. سازمان بهداشت جهانی در سال ۲۰۱۱ گزارش می‌کند که سالانه بیش از ۳۰۰۰۰۰ نفر بر اثر عوارض ناشی از سوختگی در جهان فوت می‌کنند که ۹۵ درصد آن را مصدومان کشورهای با درآمد پایین و متوسط تشکیل می‌دهند [۲]. بر اساس آمار سال ۲۰۰۰ سازمان جهانی بهداشت در ایران از هر صد هزار نفر، ۴/۰۱ تا ۴/۸ نفر بر اثر عوارض ناشی از سوختگی با آتش می‌میرند [۳]. آمار سوختگی در ایران با جمعیت ۷۵ میلیون نفری، سالانه ۱۵۰ هزار نفر است درحالی‌که این رقم در آمریکا با جمعیت ۳۲۰ میلیونی به ۵۰۰ هزار نفر می‌رسد که این امر نشان می‌دهد آمار سوختگی در کشور ما بسیار زیاد است. به علاوه، سالانه سه هزار نفر براثر سوختگی در کشور جان خود را از دست می‌دهند که در سال‌های اخیر این رقم به دو هزار نفر رسیده است [۴].

منظور از مدل پیش‌بینی مدلی می‌باشد که می‌تواند با استفاده از متغیرهای یک نمونه، پیامد آن را پیش‌بینی کند. در علم پزشکی مطالعات زیادی در زمینه پیش‌بینی در حوزه‌های مختلف انجام شده است. یکی از این حوزه‌ها، پیش‌بینی فوت یا زنده‌ماندن بیماران سوختگی می‌باشد. از سال ۱۹۴۹ که اولین مقاله مرتبط با این موضوع منتشر گردیده است، بیش از ۱۵ مدل گزارش شده است که خیلی از آن‌ها منسوخ، پیچیده و فاقد قدرت پیش‌بینی می‌باشند [۵]. بر اساس بررسی پژوهشگران، مشخص شد که مدل‌های پیش‌بینی پیامد در بیماران سوختگی که تاکنون انجام شده است از چند جنبه خاص، به شرح زیر، با هم متفاوت می‌باشند.

یکی از جنبه‌های تفاوت، تکنیک به کار برده شده در ساخت مدل می‌باشد. در مدل‌های اولیه از روش‌های سنتی استفاده می‌شد. در این روش‌ها برای ساخت مدل از تکنیک‌های آماری نظیر رگرسیون لجستیک و آنالیز پروبیت (Probit Analysis) استفاده می‌شد [۸-۶]. به هر حال استفاده از این تکنیک‌ها برای کاربران نهایی، زمان‌گیر می‌باشد و پیش‌بینی را در یک زمان قابل قبول فراهم نمی‌آورد. لذا سیستم‌های خبره برای غلبه بر این محدودیت‌ها ایجاد شدند [۶،۹]. اشکال عمده سیستم های خبره این است که کاربر باید برای رسیدن به یک پاسخ دقیق در یک فرایند تعاملی(Interactive) با کامپیوتر به چند سؤال (Query) پاسخ دهد؛ در حقیقت کاربر باید اطلاعات زیادی پیرامون مشکل داشته باشد و آن داده‌ها را به سیستم وارد نماید. در نتیجه، ورود نامناسب و ناکافی داده‌ها موجب کاهش همسانی نتایج آن‌ها می‌گردد. در نهایت، اخیراً پژوهشگران جهت فائق آمدن بر این مسائل به تکنیک‌های داده کاوی شامل الگوریتم‌های یادگیری ماشین، روی آورده‌اند. کارایی این تکنیک‌ها در ارائه مدل‌های پیش‌بینی به اثبات رسیده است [۱۰،۱۱].

به‌علاوه تفاوت دیگر در مدل‌های ارائه شده، تفاوت در تعداد و ماهیت متغیرهای تاثیرگذار در پیش‌بینی می‌باشد. اکثر پژوهشگران در مدل‌های ارائه شده سن و سطح سوختگی را گزارش داده‌اند و در سایر مدل‌ها به متغیرهایی نظیر صدمات استنشاقی، پنومونی، درجه سوختگی و صدمات و بیماری‌های همراه اشاره شده است [۴،۱۰،۱۱].

سومین تفاوت در حجم نمونه مورد مطالعه برای ایجاد مدل پیش‌بینی می‌باشد. بعضی از مدل‌ها بر اساس پایگاه‌های داده‌ای بسیار کوچک شامل ۹۰ بیمار [۱۲] و ۱۸۰ بیمار [۱۰] انجام شده است. به‌علاوه از سال ۱۹۹۰ چهار مطالعه با حداقل حجم نمونه ۱۰۰۰ بیمار انجام شده است[۸-۶،۱۳]. البته مطالعه‌ای هم وجود دارد که بر روی حجم نمونه‌ای بسیار بزرگ شامل ۶۸۶۶۱ بیمار انجام شده است [۱۱]. چهارمین تفاوت مربوط به نوع مطالعه از لحاظ زمانی می‌باشد. اکثر مطالعات به صورت گذشته‌نگر بر روی اطلاعات پرونده پزشکی بیماران انجام شده است [۱۰،۱۱]. البته مطالعاتی هم وجود دارند که به صورت آینده‌نگر بر روی بیماران انجام شده‌اند [۱۲]. پنجمین تفاوت مربوط به تنوع داده‌ها می‌باشد. اطلاعات بعضی از مطالعات صرفاً از یک مرکز درمانی بیماران سوختگی استخراج شده است [۱۰،۱۲]. این در حالیست که بعضی از

آن‌ها دارای تنوع بسیار زیادی از اطلاعات موجود از چند مرکز تخصصی درمانی بیماران سوختگی می‌باشند [۱۱].

ششمین تفاوت مربوط به عینیت مدل می باشد. بعضی از مدل‌ها به صورت کمی ارائه شده‌اند و یک شاخص نمره‌ای برای آن مدل در نظر گرفته شده مثل سیستم نمره‌دهی Baux و آپاچی سه و بعضی دیگرمثل مطالعه مک‌گوین به صورت کیفی گزارش شده‌اند [۱۱]. هفتمین تفاوت مربوط به پیامد مورد پیش‌بینی می‌باشد. اکثر مطالعات، فوت بیماران را پیش-بینی کرده‌اند[۱۲،۱۱،۵].ولی بعضی دیگر میزان بقاء را پیش-بینی کرده‌اند [۱۴،۱۰]. در نهایت اینکه، بعضی از مدل‌ها بسیار ساده در قالب دو یا سه متغیر گفته شده و بعضی دیگر به صورت پیچیده در قالب متغیرهای بیشتری بیان شده‌اند. برخی از مدل‌های ساده عبارتند از مدل Griffe بر اساس سن و سطح سوختگی[۱۳]، مدل‌های Ryan و Smith بر اساس سن و سطح سوختگی، صدمات استنشاقی[۸،۶] و مدل O'Keefe بر اساس جنس، سن و سطح سوختگی و صدمات استنشاقی [۷] شاید پیچیده ترین مدلی که تاکنون ارائه شده است مربوط به مطالعه مک‌گوین و همکاران [۱۱] باشد که در آن مدل به متغیرهایی مثل سن، سطح سوختگی، پنومونی، صدمات و بیماری های همراه و صدمات استنشاقی به عنوان متغیرهای اصلی توجه شده است.

ابزار پیش‌بینی تحت عناوین مدل و یا سیستم‌های نمره‌دهی (Scoring System) منتشر شده‌اند. رایج‌ترین و شناخته‌ترین مدل‌های پیش‌بینی فوت بیماران سوختگی که تاکنون منتشر شده‌اند عبارتند از: سیستم Baux (که بر اساس سن و سطح سوختگی محاسبه می‌شود)، سیستم نمره‌دهی ABST (که بر اساس جنسیت، سن، درجه سوختگی و سطح سوختگی محاسبه می‌شود)، شاخص Roi (که بر اساس سن، جنسیت، درجه سوختگی و سطح سوختگی محاسبه می‌شود) و سیستم APACHE III همانطور که ملاحظه می گردد، بعضی از مدل‌ها با نام خاصی منتشر شده‌اند. البته مدل‌هایی هم وجود دارد که نامی برای آن‌ها انتخاب نشده است.

مدل‌های پیش‌بینی بیماری‌ها بر اساس داده‌های جامعه مشخصی ایجاد می‌شوند، بنابراین، بهترین مورد کاربرد آن‌ها برای همان جامعه و سایر جوامعی است که جمعیت آن‌ها ویژگی‌های مشابهی با جمعیت جامعه اصلی داشته باشند. بدیهی است ویژگی‌های جمعیت کشور ما به دلایل متفاوت اقلیمی، نژادی، مذهبی و فرهنگی با ویژگی‌های جمعیت سایر کشورها بخصوص کشورهای اروپایی و آمریکایی تفاوت‌های زیادی دارد. لذا استفاده از مدل‌های پیش‌بینی آن‌ها برای جمعیت کشور ما هیچگاه روش مناسبی نبوده و منجر به دستیابی به دقت قابل قبولی نخواهد شد. بنابر این، بومی‌سازی مدل‌های پیش‌بینی بر اساس ویژگی‌های منحصر بفرد کشورمان ضروری به نظر می‌رسد.

امروزه با توسعه روزافزون کاربردهای فن‌آوری اطلاعات در حیطه پزشکی، جمع‌آوری و ذخیره‌سازی حجم زیادی از داده‌ها تسهیل شده، و امکان بکارگیری روش‌های داده‌کاوی در این حوزه نیز برای پژوهشگران فراهم شده است. روش‌های داده‌کاوی از الگوریتم‌های متعدد یادگیری برای تولید مدل-های پیش‌بینی استفاده می‌کنند که با توجه به برخی عوامل نظیر ویژگی‌های داده‌ها و حیطه کاربردی مدل‌ها، برخی الگوریتم‌ها نسبت به سایر آن‌ها در اولویت قرار می‌گیرند. در حوزه پزشکی، الگوریتم‌های ساخت درخت تصمیم، به علت سادگی و فهم آسان‌تر محبوبیت بیشتری نسبت به سایر الگوریتم‌های یادگیری ماشین دارند؛ الگوریتم‌های یادگیری متعددی برای ساخت درخت تصمیم استفاده می‌گردند. عملیات ایجاد درخت از مجموعه داده‌های آموزش را ساخت درخت (Tree Induction/Building/Growing) می-گویند. اغلب الگوریتم‌هایی که برای این منظور به طور رایج استفاده می شوند عبارتند از: درخت تصمیم (Decision Tree/DT)، ID3 (Iterative Dichotomiser 3)، C4.5 (توسعه یافته ID3) [۱۵،۱۶]، CART (Classification And Regression Tree) [۱۷] و CHAID (Chi-squared Automatic Interaction Detector) می‌باشد. در این مطالعه پژوهشگران برای تولید مدل پیش‌بینی از الگوریتم CHAID استفاده نمودند.

در این مطالعه، پژوهشگران قصد دارند با دخالت دادن متغیرهایی که نقش بعضی از آن‌ها در پیش‌بینی فوت و زنده-

ماندن بیماران سوختگی توسط سایر پژوهشگران در این زمینه به اثبات رسیده است، مدلی بر اساس پایگاه داده‌ای با تعداد زیادی رکورد از بیماران سوختگی ارائه دهند. جهت ایجاد این مدل، از یک پایگاه داده‌ای با تعداد زیاد رکورد از بیماران سوختگی مرکز آموزشی-درمانی آیت اله طالقانی اهواز استفاده گردید. 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 حادثه خیز در مورد آتش‌سوزی و سوختگی محسوب می‌شود. این بیمارستان 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. لذا به خاطر تنوع داده‌های بیماران از مکان ها و فرهنگ‌های مختلف، مدل ایجاد شده می‌تواند برای انواع مختلف بیماران به کار رود و در واقع از قابلیت تعمیم پذیری نسبتاً بالایی، برخوردار باشد. بعلاوه، انتظار می رود این مدل بتواند در تصمیم‌گیری بالینی به متخصصین مراقبت بالینی و سایر کادر درمان درگیر در مراقبت موثر بیماران سوختگی کمک شایانی نماید.

روش کار

این پژوهش توصیفی گذشته‌نگر است که مبتنی بر اطلاعات پرونده‌های بیمارستانی می‌باشد. جمعیت مورد مطالعه شامل بیماران سوختگی بستری شده در مرکز آموزشی-درمانی آیت اله طالقانی اهواز از سال ۱۳۸۰ تا سال ۱۳۸۶ می‌بود. در مجموع ۴۸۰۴ پرونده پزشکی مربوط به بیماران سوختگی، مورد بررسی قرار گرفت. ابزار جمع‌آوری داده‌ها چک‌لیست پژوهشگرساخته بود که مطابق با اهداف ویژه پژوهش طراحی شده بود. داده‌ها مطابق با این چک لیست گردآوری شدند. در این مطالعه صرفاً پرونده‌های بیمارانی که در زمان پذیرش دچار صدمه سوختگی با یکی از عوامل سوختگی شده بودند، بررسی شدند و پرونده‌های بیمارانی که به علل عوارض ناشی از سوختگی مثل اسکار سوختگی، پذیرش شده بودند از مطالعه حذف گردید.

در مورد تحلیل و پردازش داده ها باید گفت که ابتدا داده‌ها با استفاده از آمارهای توصیفی (فراوانی، درصد و میانگین) از طریق نرم‌افزار SPSS 16 تحلیل شدند، سپس داده‌های گردآوری شده با استفاده از نرم‌افزار Clementine 12.0 و از طریق الگوریتم CHAID که از انواع الگوریتم‌های درخت تصمیم می‌باشد، مورد تحلیل قرار گرفت. به‌علاوه، اولین مرحله در ایجاد هر مدلی بر اساس تکنیک‌های داده‌کاوی، پردازش اولیه (Preprocessing) می‌باشد. در این مرحله مقادیر داده‌ها از لحاظ وجود مقادیر گمشده (Missing) بررسی می‌گردد و رکوردهای دارای مقادیر گمشده حذف خواهند شد. همچنین متغیر کد شناسایی بیمارحذف گردید و اسامی بیماران به صورت ناشناس(Anonymous) در پایگاه داده قرارگرفت. در حقیقت از طریق این مرحله، داده‌ها آماده تحلیل می‌گردند.

یکی از عوامل مهم در پیچیدگی مدل‌های پیش‌بینی تولید شده توسط الگوریتم‌های یادگیری ماشین، تعداد متغیرهای پیش-بینی می‌باشد. برخی پژوهشگران برای جلوگیری از پیچیدگی مدل، از تعداد متغیرهای پیش‌بینی کاسته و تنها متغیرهای مهمتر را برای تولید مدل‌ها استفاده می‌کنند. اما از آنجائیکه متغیرهای پیش‌بینی انواع متفاوتی داشته و هرکدام در پیش‌بینی نتیجه نقش متفاوتی ایفا می‌کنند، بنابراین بهتر است از تمامی آن ها در تولید مدل‌های پیش‌بینی استفاده گردد. الگوریتم CHAID قابلیت بکارگیری تمامی متغیرها را در تولید مدل پیش‌بینی دارد. داده‌های آماری به ویژه داده‌های حوزه پزشکی در حجم بالا عاری از مقادیر گمشده نیستند. وجود این مقادیر تأثیر بسزایی در عملکرد الگوریتم‌های متعدد یادگیری ماشین دارد. الگوریتم CHAID از معدود الگوریتم‌هایی است که رفتار مناسبی را در مواجهه با مقادیر گمشده از خود نشان می-دهد. درخت تولید شده توسط این الگوریتم لزوماً دودویی نمی‌باشد. این خود از ویژگی‌های مهم دیگر این الگوریتم می‌باشد. بدین ترتیب امکان شناخت و درک مدل برای متخصصین سوختگی افزایش می‌یابد و انعطاف‌پذیری بیشتری را به ایشان در استفاده از مدل در تصمیم‌گیری‌های مهم بالینی می‌دهد.[۱۸] در مورد روش اجرای درخت تصمیم CHAID، باید گفت که این الگوریتم روش مدل‌سازی است که برای

مطالعه روابط بین یک متغیر وابسته و تعداد زیادی متغیرهای مستقل به کار می‌رود. متغیرهای پیش‌بینی کننده می‌توانند کیفی یا کمی باشند. این روش برای بررسی نقش متغیرهای مستقل کیفی از آنالیز Chi-Square و برای بررسی نقش متغیرهای مستقل کمی از روش های آنالیز واریانس استفاده می‌کند. این الگوریتم متغیرهای تاثیرگذار در پیش‌بینی متغیر نتیجه را بر اساس P-Value انتخاب می‌کند[۱۹]. در مورد روش ایجاد مدل پیش‌بینی و ارزیابی کارایی آن می توان گفت ابتدا باید با استفاده از تکنیکی مجموعه داده‌ها را به زیرمجموعه‌هایی جداگانه برای ایجاد و تست مدل‌ها تفکیک کنیم. برای کاهش تورش در مدلسازی استفاده از روش اعتبارسنجی K-Fold برای این تفکیک توصیه شده است[۲۲-۲۰].

در این مطالعه، روش اعتبارسنجی k-Fold، مقدار k برابر پنج در نظر گرفته شد، یعنی رکوردها را به صورت تصادفی به پنج گروه نسبتاً مساوی تقسیم گردید و پنج بار عملیات آموزش (Train) برای چهار گروه و عملیات آزمون (Test) برای گروه پنجم تکرار گردید الگوریتم درخت تصمیم روی زیرمجموعه آموزش، اعمال و مدل پیش‌بینی ایجاد گردید. سپس مدل ایجاد شده روی زیرمجموعه باقیمانده آزمایش و کارایی آن‌ها با استفاده از معیارهای حساسیت (Sensitivity)، ویژگی (Specificity) و دقت (Accuracy) ارزیابی شدند. همچنین درصد اهمیت و وزن هر کدام از متغیرها در پیش‌بینی پیامد مشخص گردید. حساسیت و ویژگی در آمار دو شاخص برای ارزیابی نتیجه یک دسته بندی دودویی (دوحالته) هستند. زمانی که بتوان داده ها را به دو گروه مثبت و منفی تقسیم کرد، دقت نتایج یک آزمایش که اطلاعات را به این دو دسته تقسیم می کند با استفاده از شاخص‌های حساسیت و ویژگی قابل اندازه‌گیری و توصیف است. حساسیت به معنی نسبتی از موارد مثبت است که آزمایش آن‌ها را به درستی به عنوان مثبت علامت‌گذاری می‌کند. ویژگی به معنی نسبتی از موارد منفی است که آزمایش آن‌ها را به درستی به عنوان منفی علامت‌گذاری می‌کند. به بیان ریاضی، حساسیت حاصل تقسیم موارد مثبت واقعی(True Positive) به حاصل جمع موارد مثبت واقعی و موارد منفی کاذب (False Negative/FN) است. ویژگی نیز حاصل تقسیم موارد منفی واقعی (True Negative) به حاصل جمع موارد منفی واقعی و مثبت کاذب (False Positive) است. دقت، اشاره به میزان نزدیکی مقدار اندازه‌گیری شده به مقدار واقعی دارد [۲۳]. همچنین از مقدار FN (به تعداد موارد مثبتی که به اشتباه منفی بیان شده‌اند) [۲۴]، برای سنجش بهینه بودن مدل استفاده گردید.

یافته ها

در مرحله اولیه تعداد ۶۰۲۴ پرونده مربوط به بیماران سوختگی بررسی گردید و اطلاعات بیماران بر اساس متغیرهای از قبل تعریف شده استخراج گردید. متغیرهای استخراج شده، از پرونده بیماران شامل ۱۲ متغیر (کد، سن، جنسیت، درصد سوختگی، درجه سوختگی، عامل سوختگی، مدت زمان بستری، کشت ادرار، کشت خون، کشت زخم، ماه بستری و وضعیت ترخیص) بودند. متغیر پیامد (Outcome) در این مطالعه وضعیت ترخیص بیمار (زنده ماندن یا فوت) می‌باشد.

با حذف رکوردهایی که مقادیر وضعیت ترخیص آن‌ها تهی بود و یا به میل شخصی مرخص شده بودند، تعداد ۴۸۰۴ رکورد باقی ماند که از این تعداد ۳۴۱۶ بیمار (۷۱.۱۱ درصد) زنده و بقیه آن ها یعنی ۱۳۸۸ بیمار (۲۸.۸۹ درصد) فوت کرده بودند که تعداد بیماران فوت شده جنس مؤنث نسبت به بیماران فوت شده جنس مذکر بیشتر بوده است (۹۲۷ زن در برابر ۴۶۱ مرد). بیشترین میزان سوختگی‌ها و همچنین فوت شده‌ها مربوط به ماه بهمن می‌باشد، بیشترین میزان سوختگی‌ها مربوط به سوختگی‌های درجه دو، ولی بیشترین میزان فوت ناشی از سوختگی درجه سه بوده است. رایجترین علت سوختگی نیز آتش بوده که منجر به بیشترین تعداد فوت شده و بیشترین میزان فوت نیز در رده سنی ۱۵ تا ۳۰ سال بوده است. جدول شماره یک شاخص های آمار توصیفی از متغیرهای عددی پیوسته را نشان می دهد.

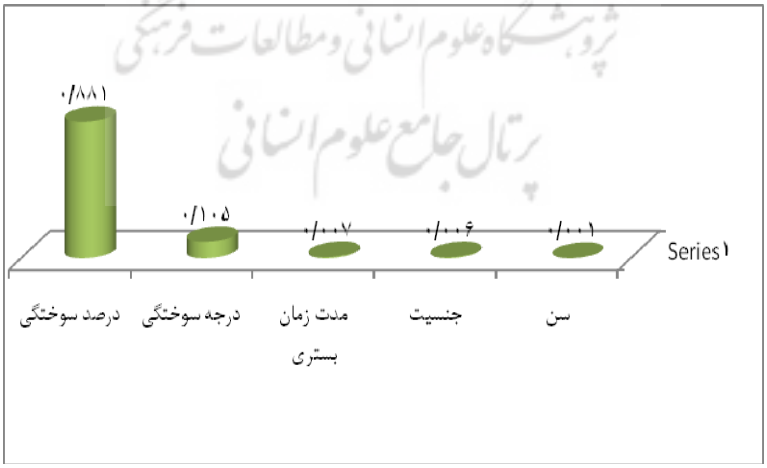
جدول ۱: آمار مربوط به متغیرهای عددی پیوسته

نام متغیر	کمترین	بیشترین	میانگین	انحراف معیار
سن (سال)	۱	۹۱	۲۳/۲	۱۶/۹۹
درصد سوختگی	۰/۵	۱۰۰	۴۰/۸	۳۰/۸۵
مدت زمان بستری (روز)	۱	۱۰۰	۱۴/۷	۱۴/۴۷

متغیرهای کیفی و گسسته مورد استفاده در ایجاد مدل پیش-بینی و مقادیر آن‌ها، عبارت بودند از:

جنس (مذکر، مونث)، درجه سوختگی (یک، دو، سه، هیچکدام)، عامل سوختگی (مایعات، الکتریسیته، آتش، قیر، مواد شیمیایی، مواد مذاب)، ماه پذیرش (۱۲ ماه سال)، کشت خون (No Culture, Staphylococcus, Negative Culture, Pseudomonas, Enterobacter, E-Coli, Klebsiella, Bacillus Proteus, Streptococcal, Actinobacter, none)، کشت زخم (No Culture, Negative Culture, Staphylococcus, Pseudomonas, Enterobacter, E-Coli, Klebsiella, Bacillus Proteus, Streptococcal, Actinobacter, none)، کشت ادرار (No Culture, Negative Culture, Staphylococcus, Pseudomonas, Enterobacter, E-Coli, Klebsiella, Bacillus Proteus, Streptococcal, Actinobacter, none)، وضعیت ترخیص (زنده، فوت).

پس از مدل‌سازی با استفاده از درخت تصمیم CHAID، پنج متغیر به عنوان متغیرهای تاثیرگذار در پیش‌بینی مشخص شدند. برای پیش‌بینی فوت و زنده‌ماندن، الگوریتم مدل‌سازی از بین ۱۱ متغیر موجود، تنها از پنج متغیر استفاده کرده است. نمودار یک درصد وزن و اهمیت هر کدام از متغیرهای تاثیرگذار در پیش‌بینی را نشان می‌دهد. هر چند متغیرهای سن، جنس، مدت زمان بستری و درجه سوختگی نسبت به متغیر درصد سوختگی، تاثیر کمتری در مرگ یا بقاء بیماران داشتند اما جزء متغیرهای تاثیرگذار بودند. متغیر سن در کودکان کمتر از یک سال و افراد بالای ۵۰ سال بر پیامد سوختگی تأثیرگذار بود. متغیر جنس با درصد سوختگی ارتباط داشت و سوختگی‌های با درصد بالا در جنس مونث بیشتر بود. بین مدت زمان بستری و بروز برخی عفونت ها در برخی از بیماران ارتباط وجود داشت. درجه سوختگی، که نشان دهنده شدت سوختگی است نیز بر پیامد سوختگی در بیماران تاثیرگذار بود.



نمودار ۱: درصد اهمیت و وزن متغیرهای تأثیرگذار در پیش بینی

با توجه به مدل پیش‌بینی ایجاد شده در این مطالعه، جداول دو و سه، قواعد (Rules) پیش‌بینی توسط برگ‌هایی (Leaf) (در ساختار درخت تصمیم، به نودهای پایین‌ترین سطح که هیچ فرزندی ندارند، برگ گفته می‌شود که هر کیس به طور اختصاصی مربوط به یک برگ می‌باشد) که بالاترین توانایی در پیش‌بینی فوت و زنده‌ماندن را دارند نشان می‌دهد.

جدول ۲: شرایط پیش بینی توسط برگ هایی با بالاترین توانایی در پیش بینی فوت

شماره برگ	قواعد پیش‌بینی فوت
۲۶	تمامی افراد جنس مونث که درصد سوختگی بالای ۹۳ درصد داشته‌اند فوت شده‌اند.
۲۲	تمامی بیمارانی که درصد سوختگی آن‌ها بین ۶۰ تا ۹۳ درصد بوده و مدت زمان بستری آن‌ها کمتر یا مساوی سه روز است فوت شده-اند.
۳۵	تمامی بیمارانی که درصد سوختگی آن‌ها بین ۵۴ و ۷۶ درصد بوده و درجه سوختگی آن‌ها بیشتر از دو و همچنین مدت زمان بستری آن‌ها کمتر یا مساوی ۱۰ روز بوده، فوت شده‌اند.
۳۸	۹۹ درصد از افراد جنس مونث که درصد سوختگی آن‌ها بین ۷۶ و ۹۳ درصد بوده و مدت زمانی بین سه تا ۱۶ روز بستری بوده‌اند، فوت شده‌اند.
٤۰	۹۸.۳ درصد از بیمارانی که درصد سوختگی بین ۷۶ تا ۹۳ درصد داشته‌اند، مدت زمانی بیش از ۱۶ روز بستری شده و درجه سوختگی آن‌ها بیش از دو بوده فوت شده‌اند
۲۵	۹۸ درصد از افراد جنس مذکر که درصد سوختگی بالای ۹۳ درصد داشته‌اند فوت شده‌اند.

جدول ۳: شرایط پیش‌بینی توسط برگ‌هایی با بالاترین توانایی در پیش‌بینی زنده‌ماندن

شماره برگ	قواعد پیش بینی زنده ماندن
۲۸	۹۹.۹ درصد از افرادی که درصد سوختگی ۱۸ یا کمتر داشته، نوع کشت خون آن‌ها یکی از موارد،Ecoli, Pseudomonas یا No culture بوده و درجه سوختگی آن‌ها بیش از یک بوده است زنده مانده‌اند.
۲۹	۹۹.۳ درصد از افرادی که درصد سوختگی آن‌ها بیشتر از ۱۸ و کوچکتر مساوی ۳۰ بوده، مدت زمان بستری بیشتر از سه روز داشته‌اند و سن ایشان کمتر یا مساوی ۴۵ سال بوده است زنده مانده‌اند.
۱٤	تمامی افرادی که درصد سوختگی آن‌ها بیشتر از ۳۰ و کوچکتر مساوی ۳۹ بوده و درجه سوختگی آن‌ها دو بوده است زنده مانده‌اند.
۳۲	تمامی افرادی که درصد سوختگی آن‌ها بیشتر از ۳۹ و کوچکتر مساوی ۵۴ بوده و درجه سوختگی آن‌ها دو بوده و همچنین مدت زمان بستری آن‌ها بیشتر از ۱۶ روز بوده زنده مانده‌اند.

چنانچه برگ‌های درخت را بر اساس درجه اهمیت برای پیش‌بینی در پنج گروه دسته‌بندی کنیم مهمترین برگ‌ها در گروه اول قرار می‌گیرند. بیمارانی که در برگ‌های گروه اول قرار می‌گیرند به احتمال بالای ۹۲ درصد، سوختگی آن‌ها منتج به فوت می‌شوند. جدول چهار آمار مربوط به برگ‌های درخت را در قالب این پنج گروه نشان می‌دهد.

جدول٤: آمار مربوط به پنج گروه از برگ‌ها

شماره گروه	گروه (برگ‌ها)	کل رکوردهای گروه	تعداد رکوردها با نتیجه مرگ در گروه	درصد فوت در گروه	شاخص
۱	۲۶، ۲۲، ۳۵، ۳۸، ۴۰، ۲۵	۹۶۱	۹۵۲	۹۲	۳۴۳
۲	۳۳، ۳۶، ۳۷، ۱۸، ۳۹، ۳۴، ۱۵، ۱۱، ۱۶، ۳۱، ۱۳	۹۶۱	۴۰۰	۴۱/۶	۲۴۳
۳	۳۰، ۲۰، ۱۰، ۲۷	۹۶۰	۳۲	۳	۱۶۶
۴	۲۹	۹۶۱	۴	۰/۴	۱۲۵
۵	۲۸، ۱۴، ۳۲	۹۶۱	۰	۰	۱۰۰

در جدول ۴ شاخص ۳۴۳ برای رکوردهای گروه اول نشان می‌دهد که چنانچه رکوردی از برگ‌های این گروه انتخاب شود، امکان فوت وی ۳.۴ برابر بیشتر از رکوردی است که به صورت تصادفی از بین کل رکوردهای مجموعه داده‌ای انتخاب گردد؛ بنابراین گره‌های واقع در گروه یک، گره‌های مهمی در پیش‌بینی نتیجه فوت پس از سوختگی می‌باشند.

برای ارزیابی مدل پیش‌بینی ابتدا مدل تولید شده با استفاده از کل رکوردها ارزیابی گردید و بعد از آن با استفاده از روش اعتبارسنجی k-Fold مدل مورد ارزیابی قرار گرفت. پارامترهای مورد استفاده برای ارزیابی مدل حساسیت، ویژگی و دقت می‌باشند. جدول پنج، ماتریس طبقه‌بندی اشتباه مربوط به مدل تولید شده با کل رکوردها را نشان می‌دهد.

جدول۵. ماتریس طبقه‌بندی اشتباه (Misclassification) مدل پیش‌بینی کننده بر اساس کل رکوردها

آزمون / واقعیت	بله	خیر
مثبت	۱۲۲۵	۸۶
منفی	۱۶۳	۳۳۳۰

حساسیت = ۸۸/۲٪ ویژگی = ۹۷/۵٪ دقت = ۹۴/۸٪

جدول٦. ارزیابی دقت مدل پیش‌بینی با استفاده از روش اعتبارسنجی k-Fold

گروه	آموزش		آزمون	
	تعداد رکورد	دقت مدل (درصد)	تعداد رکورد	دقت مدل(درصد)
گروه اول	۳۸۳۶	۹۵	۹۶۸	۹۴
گروه دوم	۳۸۴۲	۹۵	۹۶۲	۹۴
گروه سوم	۳۸۱۷	۹۵	۹۸۷	۹۵
گروه چهارم	۳۸۵۸	۹۵	۹۴۶	۹۵
گروه پنجم	۳۸۶۳	۹۵	۹۴۱	۹۶
میانگین	۳۸۴۳	۹۵	۹۶۱	۹۵

بحث و نتیجه گیری

در این مطالعه، پژوهشگران از تکنیک‌های داده‌کاوی برای ساختن مدل پیش‌بینی دقیق و موثر استفاده نمودند. تعداد کل متغیرها، ۱۱ مورد بودند که الگوریتم درخت تصمیم صرفاً از پنج مورد آن‌ها در ساختن مدل استفاده نمود. درخت تصمیم بر اساس CHAID ساخته شد که منجر به تولید ۳۲ برگ گردید. یافته‌ها نشان داد که متغیرهای موثر در پیش‌بینی فوت و زنده‌ماندن بیماران سوختگی به ترتیب اهمیت عبارتند از سطح سوختگی، درجه سوختگی، مدت زمان بستری، جنسیت و سن. در مدل ارائه شده توسط پژوهشگران، به متغیری چون مدت زمان بستری توجه شده است که در مطالعات قبلی [۱۴-۵] گزارش نشده است. از آنجاییکه ویژگی مدل پیشنهادی بالاتر از حساسیت آن می باشد بنابراین قدرت بیشتری در پیش‌بینی زنده‌ماندن نسبت به پیش‌بینی فوت دارد.

سن و سطح سوختگی تنها متغیرهایی هستند که تقریباً در همه مدل‌های پیش‌بینی ارائه شده در مطالعات موجود تا حال حاضر گزارش شده اند. اهمیت سن در فوت بیماران به عملکرد سیستم ایمنی آن‌ها مربوط می‌شود. در حقیقت، عملکرد سیستم ایمنی در افراد مسن ضعیف می باشد، لذا امکان بهبود زخم در آن‌ها پایین است و این می‌تواند در افزایش فوت آن‌ها نقش مهمی داشته باشد[۲۵].

در این مطالعه مشخص گردید که میزان فوت بیماران بعد از سوختگی، در جنس مؤنث بیش از جنس مذکر می‌باشد. این نتیجه با نتایج اکثر مطالعات در این زمینه، مطابقت دارد[۲۸-۲۶،۱۲-۱۰،۷]. در واقع از بین تمام مطالعات، صرفاً یک مطالعه گزارش شده است که میزان فوت جنس مذکر نسبت به جنس مونث به علت سوختگی بیشتر است. آن‌ها در مقاله گزارش داده‌اند که این افزایش مرگ‌ومیر در آقایان به خاطر وضعیت اقتصادی ضعیف و میزان بیکاری بالا (۴۵ درصد در سال ۲۰۰۴) است که می‌تواند منجر به صدمات سوختگی به علت سوء مصرف الکل باشد[۵]. در مطالعات داخلی انجام شده در مناطق مختلف کشور نیز مرگ‌ومیر ناشی از سوختگی در زنان بیشتر گزارش شده است [۳۱-۲۹].

حسین و همکاران در یک مطالعه مروری به بررسی ۴۵ مدل ارایه شده برای پیش‌بینی زنده‌ماندن در بیماران سوختگی پرداختند، آن‌ها نتیجه‌گیری کردند که تنها هشت مدل از ساختار و روش‌شناسی استاندارد برای پیش‌بینی بقاءدر بیماران سوختگی برخوردار بودند. یکی از عوامل مهم در مدل‌های پیش‌بینی بقاء در بیماران سوختگی استفاده از روش‌هایی برای ارزیابی اعتبار مدل است[۳۲].

در این مطالعه، مدل ایجاد شده با استفاده از روش‌های ارزیابی داده‌کاوی جهت سنجش دقت و اعتبار آن‌ها مورد بررسی قرار گرفت. اعتبار مدل به روش اعتبارسنجی k-Fold ارزیابی گردید. همچنین، یکی از شاخص‌های ارزیابی کیفیت مدل-های ایجاد شده از تکنیک‌های داده‌کاوی، مقدار FN می‌باشد. در واقع هر اندازه که مقدار آن پایین باشد، نشانه بهینه‌بودن مدل می‌باشد. به علاوه شاخص سنجش بالا یا پایین‌بودن FN، مقدار حساسیت مدل می‌باشد[۱۰]. از آنجایی‌که مقدار حساسیت مدل ارائه شده در این مطالعه بالا می‌باشد (بیش از ۸۸ درصد) لذا می توان گفت FN آن پایین بوده و دال بر بهینه‌بودن مدل می‌باشد.

با توجه به میانگین دقت مدل پیشنهادی و میانگین دقت مدل روی داده‌های آزمون، مدل پیشنهادی دقیق، معتبر و قابل استناد می‌باشد.در حقیقت، نتایج این مطالعه یک مدل بومی جامع، دقیق و معتبر پیش‌بینی فوت و زنده‌ماندن بیماران سوختگی را بر اساس متغیرهای فوق‌الذکر پیشنهاد می کند. که بعضی از آن‌ها توسط سایر مطالعات گزارش شده است. لذا، چنین مدل بومی می تواند به عنوان یک معیار مهم ارزیابی جهت پیش‌بینی نتیجه درمان بیماران سوختگی مورد استفاده قرار بگیرد. همچنین جهت تریاژ صحیح بیماران سوختگی نقش بسزایی داشته و در نهایت در تصمیم‌گیری بالینی به متخصصین مراقبت بالینی و سایر افراد درگیر در درمان بیمار کمک شایانی می‌نماید. به‌علاوه می‌تواند در مدیریت و بهره-وری بهینه از منابع نقش بسزایی داشته باشد؛ در حقیقت از طریق آن تعداد نیروی انسانی مورد نیاز اعم از پزشکان، پرستاران، تکنسین‌ها و همچنین تجهیزات پزشکی ضروری قابل پیش‌بینی می‌باشد.

البته دقت یافته‌های پژوهش می‌تواند با در نظر گرفتن متغیرهایی که در این پژوهش به آن توجه نشده است بهبود یابد. این متغیرها عبارتند از صدمات استنشاقی، صدمات و بیماری‌های همراه و پنومونی. به‌علاوه، برای ارائه مدل صرفاً از داده‌های بیماران سوختگی مراجعه کننده به مرکز آموزشی-درمانی آیت اله طالقانی اهواز شامل استان خوزستان و سایر استان‌های همجوار استفاده شده است؛ لذا، در تعمیم‌پذیری نتایج مطالعه به بیماران سایر شهرها، باید دقت کرد. به عبارت دیگر، با ادغام پایگاه داده‌ای بیماران سوختگی سایر استان‌های کشور با این پایگاه داده، می‌توان به بومی‌سازی کردن مدل بیش‌بینی فوت و زنده‌ماندن در بیماران سوختگی در سطح ملی اقدام نمود. در نهایت، داده‌های مورد نیاز مطالعه از روی پرونده‌های پزشکان بیماران به صورت گذشته‌نگر و با هدفی دیگر، غیر از هدف مورد نظر پژوهشگران، استخراج شده‌اند و لذا پژوهشگران به بعضی از داده‌های ضروری جهت تحلیل، دسترسی پیدا نکردند، که این موارد از محدودیت‌های این پژوهش بودند.

با این وجود، پیشنهاد می‌گردد که از الگوریتم مورد استفاده در این مطالعه روی داده‌های سوختگی بیماران سایر استان‌های دیگر، به منظور مقایسه و ارائه مدلی کاملاً بومی، استفاده گردد. به‌علاوه، سایر الگوریتم‌های داده‌کاوی روی داده‌های سوختگی بیماران مورد مطالعه جهت مقایسه دقت، حساسیت و ویژگی، به کار برده شوند. در نهایت پیشنهاد این است که از روش‌های منطق فازی با هدف بهره‌برداری از دانش انسانی جهت ایجاد مدل پیش‌بینی استفاده گردد.

تشکر و قدردانی

نویسندگان مقاله بدینوسیله مراتب سپاس و قدردانی خود را از پرسنل بخش مدارک پزشکی مرکز آموزشی-درمانی آیت اله طالقانی اهواز، به‌خاطر مشارکت در گردآوری داده‌ها، اعلام می‌دارند

References

1. Seyedi M. [Burns]. Mashhad: Astan Quds Razavi Puplication; 1991. [Persian]
2. World Health Organization. Burn Prevention Success Stories Lessons Learned. Geneva: Switzerland: WHO; 2011.
3. World Health Organization. Injuries, and Violence Prevention Deptartment. The injury chart book: A graphical overview of the global burden of injuries. Geneva: WHO; 2002.
4. Burn Statistics. Iranian Stem Cell New Agency. Available from:http://www.bonyannews.ir. Access Date: 30/9/2013 [Persian]
5. Brusselaers N, Juhasz I, Erdei I, Monstrey S, Blot S . Evaluation of mortality following severe burns injury in Hungary: external validation of a prediction model developed on Belgian burn data. Burns 2009; 35(7): 1009-14.
6. Ryan CM, Schoenfeld DA, Thorpe WP, Sheridan RL, Cassem EH, Tompkins RG. Objective estimates of the probability of death from burn injuries. N Engl J Med 1998; 338(6): 362–6.
7. O'Keefe GE, Hunt J, and Purdue GF. An evaluation of risk factors for mortality after burn trauma and the identification of gender-dependent differences in outcomes. J Am Coll Surg 2001; 192(2): 153–60.
8. Smith DL, Cairns BA, Ramadan F, Dalston JS, Fakhry SM, Rutledge R, et al. Effect of inhalation injury, burn size, and age on mortality: A study of 1447 consecutive burn patients. J Trauma 1994; 37: 655–659.
9. Saffle JR. Predicting outcomes of burns. N Engl J Med 1998; 388 (6): 387–8.
10. Patil BM, Joshi RC, Toshniwal D, Biradar S. A new approach: role of data mining in prediction of survival of burn patients. J Med Syst 2011; 35(6): 1531-42.
11. McGwin G Jr, George RL, Cross JM, Rue LW. Improving the ability to predict mortality among burn patients. Burns 2008; 34(3): 320-7.
12. MotamedAl-shariati M, Rezaei E, Beiraghi-Toosi A. [Comparison between acute physiology and chronic health evaluation system and abbreviated burn severity index (ABSI) in prediction of burn patient's mortality]. Zahedan Journal of Research in Medical Sciences 2012; 13(9): 33-8. [In Persian]
13. Griffe O, Gartner R, Captier G, Brabet M, Baro B, Selloumi D, et al. [Evaluation of prognostic factors in the burned patient]. Annales de chirurgie plastique et esthetique 2001; 46(3): 167-72.[French]
14. Estahbanati HK, Bouduhi N. Role of artificial neural networks in prediction of survival of burn patients-a new approach. Burns 2002;28(6):579-86.
15. Hammann F, Drewe J. Decision tree models for data mining in hit discovery. Expert Opin Drug Discov 2012 Apr;7(4):341-52.
16. Quinlan JR. C4.5: Programs for machine learning. Morgan Kaufman:San Mateo;1993.
17. Breiman L, Friedman JH, Olshen RA, and Stone CJ. Classification and regression trees. Wadsworth, Belmont, CA, 1984.

18. Clementine® 12.0 Algorithms Guide: SPSS Inc.; 2007.

19. Han J, Kamber M. Data Mining: Concepts and Techniques. Second ed. United States of America: Morgan Kufman Publishers; 2006.

20. Cios KJ, Pedrycz W, Swiniarski RW, Kurgan LA. Data Mining: A Knowledge Discovery Approach: Springer; 2007.

21. Olson DL, Delen D. Advanced Data Mining Techniques: Springer; 2008.

22. Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition: Elsevier; 2005.

23. Sensitivity and specificity. Available from: http://en.wikipedia.org/wiki/Sensitivity_and_specificity. Access Date: 15/10/2013

24. Type I and type II errors. Available from: http://en.wikipedia.org/wiki/Type_I_and_type_II_errors. Access Date: 16/10/2013

25. Demling RH. Care of geriatric patients. In: Herndon DN ed. Total burn care,2nd edition. London:WB Saunders; 2001.

26. Germann G, Barthold U, Lefering R, Raff T, Hartmann B. The impact of risk factors and pre-existing conditions on the mortality of burn patients and the precision of predictive admission-scoring systems. Burns 1997; 23:195–203.

27. Othman N, Kendrick D. Epidemiology of burn injuries in the East Mediterranean Region: a systematic review. BMC public health 2010; 10(1), 83.

28. George RL, Mcgwin G, Schwacha MG, Metzger J, Cross JM, Chaudry IH, et al. The association between sex and mortality among burn patients as modified by age. J Burn Care Rehabil 2005; 26: 416–21.

29. Rastegar Lari AA, Alaghehbandan R, Nikui RA. Epidemiological study of 3341 burns patients during three years in Tehran, Iran. Burns 2000; 26 (1): 49-53.

30. Aghakhani N, Rahbar N, Feizi A, Karimi H, Vafa Shoar N. [Epidemiology of Hospitalized Patients in Burn Ward of Imam Khomeini Hospital in Urmia.(2005)]. Behbood 2008; 12(2 (37)); 140-150. [In Persian]

31. Rezaei E, Safari H, Motarnedolshariati SM, Aghaei-Monavar A. [Analysis of mortality in a burn center]. Medical Journal of Mashhad University of Medical Sciences 2010; 52 (4): 239-243.[In Persian]

32. Hussain A, Choukairi F, Dunn K. Predicting survival in thermal injury: A systematic review of methodology of composite prediction models. Burns 2013[Article in Press]

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

Determining of the Factors Affecting Mortality in Burn Patients Using a Decision Tree Data Mining Algorithm

Azizi AA[1]/ Zarei J[2]/ Nabovati E[1]/ Vakili-Arki H[1]/ Abbasi E[3]/ Razavi AR[4]

Abstract

Introduction: Using data mining to develop prediction models from large amounts of data in medicine has become increasingly popular during the recent years. In this study, we intend to use a decision tree data mining algorithm in order to identify factors influencing mortality in burn patients.

Methods: The present retrospective descriptive study is based on burning patients' records. Overall, the medical records of 4804 patients were scrutinized. The collected data were analyzed using statistical software (SPSS version 16), data mining software (Clementine version 12), and CHAID algorithm.

Results: The resulting model for predicting survival and mortality of burning patients included the percentage of Total Burn Surface Area (TBSA %), degree of burn, length of stay, gender, and age of patients. Other variables including blood cultures, wound cultures, urine cultures, and the months of patient hospitalization had no effect on improving the efficiency of the model.

Conclusion: The proposed model is valid and reliable due to its accuracy (approximately 95%). In fact, the results of this study, some of which are consistent with the results of other studies, can propose a comprehensive, accurate, and reliable local model for predicting mortality and survival of burning patients based on the mentioned variables. Thus, this local model can be used as an important criterion for evaluating the effective treatment of burn patients.

Keywords: Data Mining, Decision Tree, Burn

● **Received:** 8/May/2013 ● **Modified:** 19/Oct/2013 ● **Accepted:** 20/Nov/2013

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1. PhD Candidate of Medical Informatics, Student Research Committee, Mashhad University of Medical Sciences, Mashhad, Iran
2. PhD Student of Health Information Management, Health Management and Economics Research Center, School of health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran
3. PhD Candidate of Medical Informatics, Member of Cancer Informatics Research Group, Breast Cancer Research Center, ACECR, Tehran, Iran
4. Assistant Professor of Medical Informatics Department, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran; Corresponding Author (razaviar@mums.ac.ir)