

مدل ترکیبی داده‌کاوی برای تحلیل الگوهای حاکم بر تصادفات عابرین پیاده در ایران

مقاله علمی - پژوهشی

علی توکلی کاشانی، دانشیار، دانشکده مهندسی عمران و مرکز تحقیقات ایمنی کاربردی حمل‌ونقل جاده‌ای،

دانشگاه علم‌و‌صنعت ایران، تهران، ایران

محمد مهدی بشارتی*، استادیار، دانشکده مهندسی حمل‌ونقل، دانشگاه صنعتی اصفهان، اصفهان، ایران

*پست الکترونیکی نویسنده مسئول: besharati@iut.ac.ir

دریافت: ۱۴۰۴/۰۱/۰۸ - پذیرش: ۱۴۰۴/۰۵/۰۱

صفحه ۱۵۴-۱۳۵

چکیده

عابرین پیاده به عنوان یکی از مهم‌ترین کاربران آسیب‌پذیر در تصادفات ترافیکی، بخش قابل‌توجهی از فوتی‌های ناشی از تصادفات را به خود اختصاص می‌دهند. هدف مطالعه حاضر بررسی الگوهای حاکم بر تصادفات عابرین پیاده و شناسایی مجموعه عواملی است که با قرار گرفتن در کنار یکدیگر منجر به افزایش احتمال فوت عابرین در تصادفات می‌گردد. برای این منظور، داده‌های تصادفات ترافیکی عابرین پیاده در کشور ایران طی سال‌های ۱۳۹۵ تا ۱۴۰۱ در قالب یک چارچوب داده‌کاوی ترکیبی مورد تحلیل قرار گرفت. در گام اول، داده‌های مذکور با بهره‌گیری از روش «خوشه‌بندی» به چهار خوشه همگن تقسیم شده و ویژگی‌های توصیفی هر خوشه مورد تحلیل قرار گرفت. سپس، با کمک روش «کشف قواعد پیوندی»، مجموعه عوامل و شرایطی که احتمال فوت عابر و نیز احتمال مقصر بودن عابر در تصادفات را افزایش می‌دهد، شناسایی گردید. یافته‌های این پژوهش نشان می‌دهد که اثرات تعاملی متغیرهایی مانند تاریکی هوا، تیره بودن رنگ لباس عابر، سن عابر بالاتر از ۶۵، عابرین مذکر، وقوع تصادف در آزادراه‌ها و بزرگراه‌های برون‌شهری، تصادف با وسیله نقلیه سنگین و عبور از عرض خیابان در محل غیر مجاز، از جمله شرایطی است که با افزایش احتمال فوت عابر پیاده همراه بوده است. بر این مبنا، می‌توان گفت انسداد مسیر عبور عرضی همسطح عابرین از بزرگراه‌ها و آزادراه‌ها و ایجاد تسهیلات عبور غیر همسطح عابر در این راه‌ها از اهمیت ویژه‌ای برخوردار است. همچنین، ایجاد گذرگاه‌های عرضی با آشکارسازی و روشنایی کافی و تشویق عابرین به استفاده از این گذرگاه‌ها می‌تواند احتمال وقوع و شدت تصادفات را کاهش دهد.

واژه‌های کلیدی: تصادفات فوتی، عابر پیاده، خوشه‌بندی، قواعد پیوندی

۱-مقدمه

ایمنی کاربران راه، به خصوص عابرین پیاده یکی از مهم‌ترین دغدغه‌های مدیران حوزه حمل‌ونقل است. از سوی دیگر، رشد چشمگیر تکنولوژی‌های کامپیوتری و تجهیزات نگهداری داده‌های تصادفات، موجب پدیدآمدن پایگاه‌های بزرگ از داده‌های تصادفات شده است. در این پایگاه‌های داده، اطلاعات

طبق آمار سازمان پزشکی قانونی کشور، طی سال‌های ۱۳۹۵ تا ۱۴۰۰، بیش از ۹۸ هزار و ۴۰۰ نفر در تصادفات ترافیکی در کشور ایران فوت شده‌اند که حدود ۲۱ درصد از آنان را عابرین پیاده تشکیل می‌دهند. بنابراین، تلاش برای شناسایی عوامل موثر بر تصادفات و تلفات ترافیکی به منظور کاهش این آمار و ارتقای

افزایش سرعت برخورد یکی از عوامل مهم در افزایش شدت جراحات و فوت عابر در تصادفات ترافیکی می‌باشد (Ankunda, Ali, and Mohanty, 2024). همچنین، در مطالعات پیشین، به نقش تیره بودن رنگ لباس و اثر تجمعی آن به همراه تاریک بودن هوا بر روی افزایش احتمال وقوع و شدت تصادفات عابر اشاره شده است (Hossain et al., 2024). به‌علاوه، نتایج مطالعات پیشین نشان می‌دهد مشخصات هندسی و محیطی راه از قبیل: عرض و تعداد خطوط جاده، خیس بودن سطح روسازی، نوع راه و کاربری زمین‌های اطراف محل تصادف نیز از جمله عواملی هستند که بر شدت جراحات عابریاده تاثیرگذارند (Lee and Abdel-Aty, 2005; Mohamed et al., 2013; Zajac and Ivan, 2003).

از نقطه نظر روش‌شناسی، مطالعات متعددی در سراسر جهان (Iranitalab and Khattak, 2017; Santos, Dias, and Amado, 2022) و از جمله ایران (Keymanesh and Zabihian, 2023; Moslemi mehni, Nadera, and Naser-Alavi, 2023; Nadimi et al., 2024)، از روش‌های مختلف داده‌کاوی برای تحلیل و مدل‌سازی فراوانی و شدت تصادفات ترافیکی استفاده کرده‌اند. برای مثال، ندیمی و همکارانش اخیراً در مطالعه خود با استفاده از تکنیک استخراج قواعد پیوندی به بررسی الگوهای مشترک پرتکرار در تصادفات کاربران آسیب‌پذیر در شهر تهران طی سال‌های ۱۳۹۷ تا ۱۳۹۹ پرداختند (Nadimi et al., 2024). یکی از مهمترین مشکلاتی که در تحلیل فراوانی و شدت جراحات ناشی از تصادفات وجود دارد، ناهمگنی پنهان موجود در میان داده‌های تصادفات می‌باشد. اگر این ناهمگنی در نظر گرفته نشود، می‌تواند موجب بروز خطا در استنباط آماری شود (Kim, 2023).

یکی از رویکردهایی که اخیراً محققین برای حل مشکل ناهمگنی داده‌های تصادفات پیشنهاد کرده‌اند؛ استفاده از تکنیک‌های خوشه‌بندی برای تقسیم‌بندی تصادفات در گروه‌های همگن است (Alikhani, Nedaie, and Ahmadvand, 2013). در این رویکرد ابتدا داده‌های تصادفات در گروه‌های همگن خوشه‌بندی شده و سپس به کمک روش‌های مختلف آماری و یا داده‌کاوی، تصادفات مربوط به هر خوشه مورد تحلیل قرار می‌گیرد. در این راستا، برخی از مطالعات قبلی از تکنیک خوشه‌بندی به همراه یک مدل آماری مانند مدل رگرسیون لوجستیک (Dai, 2012) و یا لوجیت چندجمله‌ای (Depaire, Wets, and Vanhoof, 2008) استفاده

مختلفی از جمله عوامل انسانی، محیطی، ویژگی‌های خودرو و مشخصات هندسی راه، برای هر تصادف ثبت شده است. با توجه به حجم زیاد داده‌های خام و متغیرهای متنوع موجود در پایگاه‌های داده تصادفات، تکنیک‌های داده‌کاوی و یادگیری ماشین از جمله روش‌های کارایی هستند که در سال‌های اخیر در حوزه تجزیه و تحلیل تصادفات مورد استفاده قرار گرفته‌اند تا بتوان به کمک این روش‌ها، الگوهای پنهان موجود درون داده‌های تصادفات را استخراج نمود. هدف پژوهش حاضر، شناسایی الگوهای پنهان، اثرات تعاملی میان متغیرها و شرایطی است که منجر به افزایش احتمال فوت عابریاده در تصادفات می‌گردد. برای این منظور از دو تکنیک داده‌کاوی خوشه‌بندی و کشف قواعد پیوندی به صورت توأم استفاده شده است.

۲- پیشینه تحقیق

مرور مطالعات پیشین حاکی از آن است که عوامل مختلفی می‌تواند در احتمال وقوع و شدت تصادفات عابریاده موثر باشد. نتایج برخی از پژوهش‌ها نشان می‌دهد که عواملی همچون مصرف الکل توسط عابر، مسن بودن (بالای ۶۵ سال)، کم سن بودن (زیر ۵ سال و زیر ۱۵ سال) و مذکر بودن عابر ریسک فوت عابر را به میزان قابل توجهی افزایش می‌دهد (Mohamed et al., 2013; Moudon et al., 2011; Salehian et al., 2023). به‌علاوه، مشخصات راننده از جمله سن راننده، جنسیت راننده و تجاوز از سرعت مجاز و عدم توجه به تجهیزات کنترل ترافیک (چراغ راهنمایی، تابلو توقف، حق تقدم و غیره) در شدت مصدومیت عابر اثرگذار است (Kim et al., 2008; Kim et al., 2010; Salehian et al., 2023).

طبق یافته‌های مطالعات پیشین، ابعاد خودرو از جمله ویژگی‌های وسیله‌نقلیه است که می‌تواند بر شدت جراحات عابر پیاده موثر باشد. تصادف عابر با وسیله نقلیه سنگین و نیمه‌سنگین می‌تواند احتمال فوت عابر بر اثر تصادف را افزایش دهد (Kim et al., 2008; Salehian et al., 2023).

همچنین در اکثر مطالعات قبلی از تاریکی محل تصادف به‌عنوان عامل موثر در افزایش شدت مصدومیت و احتمال فوت عابر نام برده شده است (Aziz, Ukkusuri, and Hasan, 2013; Salehian et al., 2023). تاریکی محل تصادف می‌تواند منجر به تاخیر در شناسایی عابریاده توسط راننده، افزایش زمان عکس‌العمل راننده و در نتیجه افزایش سرعت برخورد خودرو با عابر گردد. مطالعات قبلی نشان داده‌اند که

الگوریتم خوشه‌بندی دو مرحله‌ای از دو گام تشکیل شده است. در گام اول، تمامی رکوردهای پایگاه داده بررسی شده و پس از شناسایی مجموعه رکوردهای مشابه، اطلاعات آن‌ها به عنوان «نواحی متراکم» ذخیره می‌شوند. منظور از نواحی متراکم داده‌ای، مجموعه‌ای از رکوردهاست که مقادیر تعدادی از متغیرها برای آن‌ها یکسان است. مجموعه نواحی متراکم شناسایی شده در مرحله اول به عنوان ورودی در مرحله دوم استفاده می‌شوند. در این مرحله، با استفاده از یک الگوریتم سلسله‌مراتبی، این نواحی متراکم دوباره با یکدیگر ترکیب شده و در نهایت داده‌ها براساس معیار فاصله لگ-درستمنامی، به تعداد خوشه مناسب تقسیم می‌شوند. برای این منظور ابتدا هریک از زیرخوشه‌های (نواحی متراکم) تولید شده در گام اول به عنوان یک خوشه مستقل در نظر گرفته می‌شود. سپس، تمامی خوشه‌ها با یکدیگر مقایسه گردیده و جفت خوشه‌ای که کمترین فاصله از یکدیگر دارند انتخاب و با هم ترکیب می‌شوند. پس از ادغام این دو خوشه، مجموعه جدید خوشه‌ها مجدداً با یکدیگر مقایسه شده و عملیات ترکیب خوشه‌های نزدیک به هم، تکرار می‌شود. این فرآیند تا زمانی که براساس معیار فاصله هیچ کدام از خوشه‌های باقی مانده را نتوان با دیگری ترکیب کرد، ادامه می‌یابد.

در روش خوشه‌بندی دو مرحله‌ای از معیار فاصله لگ-درستمنامی برای تعیین میزان فاصله میان مراکز خوشه‌ها استفاده می‌شود. در این الگوریتم، فاصله بین دو خوشه به میزان کاهش در مقدار لگ-درستمنامی پس از ترکیب دو خوشه با هم، بستگی دارد. در محاسبه مقدار لگ-درستمنامی، فرض می‌شود متغیرهای طبقه‌ای از توزیع چندجمله‌ای و متغیرهای پیوسته از توزیع نرمال پیروی می‌کنند. فرض کنید قرار است خوشه C_i با خوشه C_j ترکیب شده و خوشه جدید C_{ij} نامیده شود. میزان کاهش در مقدار لگ-درستمنامی برابر است با:

$$d(i, j) = \xi_i + \xi_j - \xi_{(i, j)} \quad (1)$$

که در آن،

$$\xi_v = -N_v \left(\sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2) + \sum_{k=1}^{K^B} \hat{E}_{vk} \right) \quad (2)$$

و $v = i, j, (i, j)$ به ازای

$$\hat{E}_{vk} = - \sum_{l=1}^{L_k} \frac{N_{vkl}}{N_v} \log \frac{N_{vkl}}{N_v} \quad (3)$$

و K^A تعداد متغیرهای پیوسته، K^B تعداد متغیرهای گسسته، L_k تعداد سطوح k -مین متغیر گسسته، N_v تعداد رکوردهای

کرده‌اند. از طرف دیگر، برخی دیگر از مطالعات قبلی از رویکرد استفاده توأم از تکنیک خوشه‌بندی و یک تکنیک داده‌کاوی دیگر برای شناسایی عوامل موثر بر تصادفات استفاده کرده‌اند (de Oña et al., 2013; Ma, Mei, and Cuomo, 2021; Raihan, Hossain, and Hasan, 2018). برای نمونه، دی اونا و همکارانش از الگوریتم خوشه‌بندی کلاس پنهان به عنوان ابزاری برای گروه‌بندی داده‌های تصادف استفاده نموده و در مرحله بعد، با استفاده از تکنیک شبکه‌های بیزین مهم‌ترین عوامل موثر بر شدت تصادفات را شناسایی کردند (de Oña et al., 2013). در یک مطالعه دیگر، علیخانی و همکارانش از یک روش ترکیبی خوشه‌بندی-دسته‌بندی برای تحلیل داده‌های تصادفات بهره بردند (Alikhani, Nedaie, and Ahmadvand, 2013). در پژوهش حاضر، از یک چارچوب داده‌کاوی ترکیبی برای مطالعه تصادفات عابرین پیاده در کشور ایران طی سال‌های ۱۳۹۵ تا ۱۴۰۱ استفاده شده است. برای این منظور ابتدا از تکنیک خوشه‌بندی برای تقسیم‌بندی تصادفات در گروه‌های همگن استفاده گردیده و سپس از تکنیک کشف قواعد پیوندی برای شناسایی اثرات تعاملی متغیرها و شرایطی که می‌تواند موجب افزایش احتمال «فوت عابر» و یا «مقصر شناخته شدن عابر» در تصادف گردد، استفاده شده است.

۳- روش شناسایی پژوهش

۳-۱- خوشه‌بندی

خوشه‌بندی به معنای تقسیم داده‌ها در گروه‌های مشابه است. روش خوشه‌بندی، یکی از روش‌های یادگیری بدون ناظر است که با یادگیری داده‌ها براساس اصل حداکثر کردن شباهت داخل گروه‌ها و حداقل کردن شباهت بین گروه‌ها، انجام می‌شود. در پژوهش حاضر، به منظور خوشه‌بندی داده‌های تصادفات از الگوریتم «خوشه‌بندی دو مرحله‌ای» استفاده شده است. یکی از مزیت‌های این الگوریتم آن است که برخلاف سایر الگوریتم‌های رایج در ادبیات ایمنی ترافیک مانند k -میانگین و خوشه‌بندی کلاس پنهان، الگوریتم دو مرحله‌ای می‌تواند به طور خودکار تعداد بهینه خوشه‌ها را محاسبه کند و کمتر تحت تأثیر قضاوت پژوهشگر است. تعیین تعداد خوشه‌ها در ابتدای عملیات خوشه‌بندی یکی از چالش‌های مهم تحلیل‌گران است.

۳-۲- کشف قواعد پیوندی

کشف قواعد پیوندی که در زبان فارسی به نام‌های دیگری از جمله قواعد تلازمی و یا قواعد انجمنی نیز ترجمه شده است، یکی از پرکاربردترین تکنیک‌های داده‌کاوی ناپارامتریک است. هدف قواعد پیوندی، شناسایی روابط و همبستگی‌های موجود میان سطوح مختلف متغیرها در یک پایگاه داده بزرگ است. مهم‌ترین مزیت‌های این تکنیک، سادگی تفسیر نتایج آن و نیز امکان یافتن رابطه بین هر یک از متغیرهاست. به عبارت دیگر در این رویکرد صرفاً یک متغیر خروجی خاص وجود ندارد؛ بلکه امکان تعریف چندین متغیر به‌عنوان متغیر خروجی وجود خواهد داشت. پژوهشگران حوزه ایمنی ترافیک از این تکنیک در چندین مورد برای شناسایی همبستگی‌ها و اندرکنش‌های موجود میان عوامل دخیل در تصادفات ترافیکی بهره برده‌اند (Montella et al., 2012; et al., 1403).

طبق تعریف، منظور از کشف قواعد پیوندی، شناسایی مجموعه اقلامی (مثلاً مجموعه عوامل دخیل در شدت جراحت کاربر راه) است که در یک رویداد (مثلاً یک تصادف ترافیکی) به‌طور همزمان و پرتکرار مشاهده می‌شوند. قواعد پیوندی براساس مقایسه فراوانی نسبی دفعات مشاهده این مجموعه اقلام به تنهایی و یا با هم در یک پایگاه داده، شناسایی می‌شوند. به عبارت دیگر، تکنیک قواعد پیوندی به دنبال مجموعه عواملی است که «تعداد دفعاتی که باهم رخ داده‌اند» به‌طور قابل ملاحظه‌ای بیشتر از تعداد دفعاتی باشد که «اگر مستقل از هم بودند»، در رویدادها مشاهده می‌شدند. شکل نمایش یک قاعده پیوندی به صورت «پیشین(ها) ← نتیجه» است. توجه شود که قسمت پیشین می‌تواند شامل بیش از یک مورد باشد. در مطالعه حاضر متغیرهای دخیل به‌عنوان «پیشین»، و دو متغیر «فوت عابر در تصادف» و «مقصر بودن عابر پیاده»، به‌عنوان «نتیجه» قواعد در نظر گرفته شده است. برای شناسایی و استخراج قواعد پیوندی از سه معیار پشتیبان، اطمینان و لیفت استفاده می‌شود. در ادامه هر یک از این سه معیار به‌طور مختصر معرفی می‌شوند.

«پشتیبان» یک قاعده عبارتست از درصدی از کل رکوردهای پایگاه داده که قاعده موردنظر درمورد آن‌ها صدق می‌کند (رابطه ۵).

$$\text{support}(A \rightarrow B) = \frac{\#(A \cap B)}{N} \quad (5)$$

که در آن N تعداد کل رویدادها (در اینجا تعداد کل تصادفات)، و صورت کسر، تعداد رویدادهایی است که در آن‌ها آیت A

خوشه v ، N_{vkl} تعداد رکوردهای خوشه v که به l -مین سطح از k -مین متغیر گسسته تعلق دارد، $\hat{\sigma}_{kl}^2$ واریانس برآورد شده برای k -مین متغیر پیوسته به‌ازای تمامی رکوردها، $\hat{\sigma}_{vkl}^2$ واریانس برآورد شده برای k -مین متغیر پیوسته به‌ازای رکوردهای خوشه v -م و (i, j) اندیس نمایانگر خوشه جدیدی است که از ترکیب دو خوشه i و j تولید شده است.

در پژوهش حاضر، مدل خوشه‌بندی با استفاده از شاخص «ضرب سیلهوت» ارزیابی می‌شود. ضرب سیلهوت براساس مقایسه میزان شباهت و تمایز میان هریک از خوشه‌ها شاخصی را برای ارزیابی اعتبار خوشه‌بندی ارائه می‌کند. برای هر رکورد i مقدار سیلهوت با استفاده از رابطه ۴ محاسبه می‌شود.

$$s(i) = \frac{B(i) - A(i)}{\max(A(i), B(i))} \quad (4)$$

که در آن $A(i)$ فاصله رکورد i از مرکز خوشه‌ای که به آن تعلق دارد و $B(i)$ فاصله این رکورد از مرکز نزدیک‌ترین خوشه (غیر از خوشه‌ای که به آن تعلق دارد) است. سپس مقدار متوسط سیلهوت (که به آن ضرب سیلهوت (SC) نیز می‌گویند) برای تمامی خوشه‌ها به‌صورت متوسط مقدار $s(i)$ ها برای تمامی رکوردهای پایگاه داده محاسبه می‌شود. مقدار ضرب سیلهوت بین ۱- تا ۱ می‌باشد. اگر مقدار SC برابر با ۱ باشد بدین معناست که تمامی رکوردها بر روی مرکز خوشه خود قرار دارند و مقدار ۱- بدین معناست که تمامی رکوردهای هر خوشه بر روی مرکز خوشه‌های دیگر قرار گرفته‌اند (تمامی رکوردها به‌صورت اشتباه خوشه‌بندی شده‌اند). به‌علاوه، مقدار صفر بدین معناست که به‌طور متوسط فاصله رکوردها از مرکز خوشه‌ای که به آن تعلق دارند با فاصله آنها از مرکز نزدیک‌ترین خوشه برابر است.

به‌طورکلی، هرچه مقدار ضرب سیلهوت بیشتر باشد، نشان دهنده آنست که الگوریتم خوشه‌بندی توانسته است ساختار خوشه‌ای قوی‌تری را در میان داده‌ها کشف کند. براساس نظام رتبه‌بندی پیشنهادشده توسط روسو و کافمن (Kaufman and Rousseeuw, 2009)، در صورتی که ضرب سیلهوت برای یک تحلیل خوشه‌بندی در بازه ۰/۵ تا ۰/۷ باشد، می‌توان گفت الگوریتم توانسته است یک ساختار خوشه‌ای «خوب» در میان داده‌ها کشف کند. همچنین، اگر مقدار این ضرب برای یک تحلیل خوشه‌بندی در بازه ۰/۷۱ تا ۱ باشد، می‌توان گفت الگوریتم توانسته است یک ساختار خوشه‌ای «قوی» در میان داده‌ها کشف کند.

۳-۳- پایگاه داده

داده‌های تصادفات ترافیکی عابرین پیاده در کل کشور مربوط به سال‌های ۱۳۹۵ تا ۱۴۰۱ که توسط پلیس راهور ثبت و نگهداری می‌شود، برای تحلیل‌های پژوهش حاضر مورد استفاده قرار گرفت. پس از پایش و آماده‌سازی داده‌های مذکور، در نهایت ۳۴۷۱۹۷ تصادف مرتبط با عابرین پیاده برای انجام تحلیل شناسایی شده و مورد استفاده قرار گرفت. فهرست متغیرهای در نظر گرفته شده در این پژوهش در جدول ۱ ارائه شده است.

۴- نتایج و تفسیر آن‌ها

۴-۱- خوشه‌بندی دو مرحله‌ای

در مرحله اول، با استفاده از ترکیبات مختلف از متغیرهای معرفی شده در جدول ۱، چندین مدل خوشه‌بندی کاندیدا ساخته شد. در نهایت، مناسب‌ترین مدل خوشه‌بندی (بر مبنای شاخص سیلهوت) از میان مدل‌های کاندیدا انتخاب گردید که شامل ۴ خوشه و با استفاده از سه متغیر نوع منطقه، جنسیت عابر و نوع وسیله‌نقلیه درگیر در تصادف است. مقدار ضریب سیلهوت برای این مدل، برابر با ۰/۷ می‌باشد که براساس توضیحات ارائه شده در بخش روش‌شناسی، این مقدار نشان دهنده کشف یک ساختار خوشه‌ای خوب در میان داده‌ها توسط الگوریتم خوشه‌بندی است.

در مرحله بعد، همانند مطالعات پیشین (de Oña et al., 2013; Depaire, Wets, and Vanhoof, 2008)، خوشه‌ها براساس توزیع فراوانی سطوح مختلف متغیرها، نامگذاری شدند. برای مثال، اگر ۹۵ درصد تصادفات درون یک خوشه در مناطق برون‌شهری رخ داده باشد، درحالی‌که توزیع فراوانی تصادفات در مناطق مختلف (سطوح مختلف متغیر نوع منطقه) در سایر خوشه‌ها تقریباً یکسان باشد؛ می‌توان خوشه موردنظر را «خوشه تصادفات برون‌شهری» نامید.

همان‌طور که قبلاً اشاره شد، پس از بررسی ترکیبات مختلف از متغیرهای در دسترس، سه متغیر (نوع منطقه، نوع خودرو درگیر در تصادف، و جنسیت عابر) برای خوشه‌بندی و نامگذاری خوشه‌ها انتخاب شدند. همچنین نسبت فراوانی هر یک از سطوح متغیرها برای هر یک از این ۴ خوشه نیز در جدول ۲ قابل مشاهده است. توجه شود که تنها سطوحی که در تعیین ویژگی‌های خوشه‌ها اهمیت داشته‌اند در این جدول آورده شده است. ویژگی‌های کلی خوشه‌های شناسایی شده عبارتست از؛ خوشه ۱- تصادفات درون‌شهری عابرین مذکر با خودرو سواری (۱۳۱۱۸۶ تصادف معادل ۳۸ درصد از کل تصادفات)،

به‌عنوان پیشین و آیم B به‌عنوان نتیجه است. با در نظر گرفتن یک مقدار حداقلی برای معیار پشتیبان، تنها قواعدی که فراوانی قابل قبولی دارند، به‌عنوان قواعد پیوندی نهایی انتخاب می‌شوند. این کار موجب حذف قواعدی می‌گردد که تنها به‌صورت شانسی کشف شده‌اند. «اطمینان» یک قاعده عبارتست از احتمال مشاهده نتیجه، به شرطی که بدانیم اقلام پیشین مشاهده شده‌اند.

$$\text{Confidence} = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A)} \quad (6)$$

این معیار، میزان قوت یک قاعده پیوندی را اندازه‌گیری می‌کند. «لیفت»، برابر است با اطمینان یک قاعده تقسیم بر اطمینان مورد انتظار، با فرض اینکه رخ دادن اقلام موردنظر مستقل از هم باشد.

$$\text{Lift} = \frac{\text{confidence}(A \rightarrow B)}{\text{support}(B)} = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A) \times \text{Support}(B)} \quad (7)$$

به‌عبارت دیگر، اگر یک قاعده پیوندی را به‌صورت «اگر پیشین(ها)، آن‌گاه نتیجه» داشته باشیم؛ می‌توان گفت که مقدار لیفت برابر با مقدار افزایش احتمال رخ دادن نتیجه است به شرط اینکه بدانیم پیشین(ها) به صورت همزمان مشاهده شده‌اند. برای مثال، مقدار لیفت برابر با ۲/۳۳ بدین معناست که اگر همه رخداد‌های پیشین مشاهده شوند، آنگاه احتمال رخ دادن نتیجه، ۲/۳۳ برابر بیشتر از احتمال مشاهده نتیجه در میان تمامی رخداد‌هاست. در مورد ارزیابی قواعد پیوندی همان‌طور که قبلاً اشاره شد، می‌توان گفت هرچه مقدار دو معیار پشتیبانی و اطمینان بیشتر بوده و مقدار لیفت نیز هرچه بیشتر از ۱ باشد، قاعده مورد نظر قابل‌اتکاتر خواهد بود. اما از آنجا که تصادفات فوتی در نسبت با تصادفات جرحی و خسارتی، فراوانی کمتری دارند؛ پشتیبانی و اطمینان برای برخی از قواعد ممکن است مقدار کمی داشته باشد. بنابراین در این‌گونه مطالعات، بهتر است قواعد استخراج شده براساس مقادیر لیفت، ارزیابی شوند (Montella et al., 2012). طبق تفسیر ارائه شده برای معیار لیفت در بخش قبل، پس از استخراج قواعد پیوندی، از این معیار برای تفسیر قواعد پیوندی و شناسایی شرایطی که موجب افزایش احتمال فوت کاربران راه بر اثر تصادفات می‌گردد، استفاده شده است. به‌علاوه، برای اطمینان از این که الگوهای شناسایی شده توسط قواعد، فراوانی و دقت قابل قبولی دارند؛ مقادیر حداقلی برای دو معیار پشتیبانی و اطمینان نیز تعیین می‌شوند. با توجه به مقادیر حداقلی در نظر گرفته شده برای این سه معیار در مطالعات قبلی، در پژوهش حاضر مقدار حداقل برای سه معیار پشتیبانی، اطمینان و لیفت به ترتیب برابر با ۱ درصد، ۲ درصد و ۲ در نظر گرفته شده است.

خوشه ۲- تصادفات برون شهری (۶۵۱۴۳) تصادف معادل ۱۹ درصد از کل تصادفات)،
 خوشه ۳- تصادفات درون شهری عمدتاً با موتورسیکلت، وانت بار و وسایل نقلیه سنگین (۷۷۵۲۵) تصادف معادل ۲۲ درصد از کل تصادفات)،
 خوشه ۴- تصادفات درون شهری عابرین مونث با خودرو سواری (۷۳۳۴۳) تصادف معادل ۲۱ درصد از کل تصادفات).

جدول ۱. متغیرهای استفاده شده در پژوهش و سطوح هر متغیر

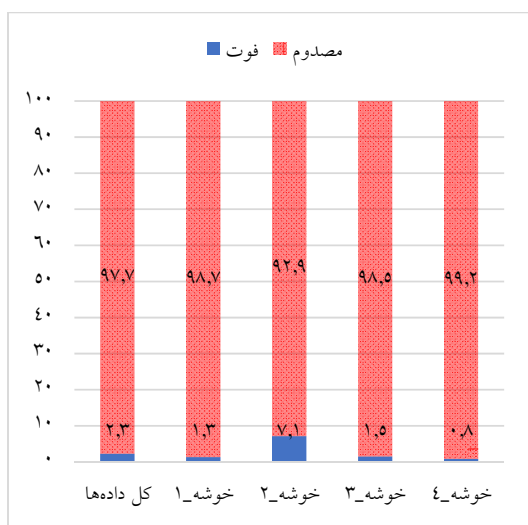
متغیر	سطوح متغیر	متغیر	سطوح متغیر	متغیر	سطوح متغیر
شدت جراحت عابر	جراحت	جنسیت راننده	مذکر	ساعت وقوع تصادف	۱۲ شب تا ۵ صبح
					مقصر
وضعیت مقصر بودن عابر	مقصر	گروه سنی راننده	زیر ۱۹	وضعیت آب و هوا	۸ تا ۱۴ ظهر
			۱۹-۲۴		۱۴ تا ۱۹ ظهر
جنسیت عابر	مذکر	نوع وسیله نقلیه	۲۴-۳۵	وضعیت سطح جاده	۱۹ تا ۱۲ شب
			۳۵-۴۴		روشن
گروه سنی عابر	زیر ۶	نوع وسیله نقلیه	۴۴-۵۴	وضعیت آب و هوا	تاریک
			۵۴-۶۴		صاف
			۶۴-۷۵		مه
			بالای ۷۵		باران
نوع رنگ لباس عابر	۶-۱۴	نوع وسیله نقلیه	سواری	وضعیت سطح جاده	برف
			موتورسیکلت		ابری
			وانت بار		سایر
			انواع کامیون		خشک
کاربری زمین	۱۵-۲۴	نوع راه	اتوبوس/مینی بوس	هندسه راه	خیس
			سایر		یخ
			آزادراه		سایر
			بزرگراه		مستقیم و مسطح
کاربری زمین	۲۵-۳۴	نوع راه	سایر	هندسه راه	مستقیم و غیر مسطح
			اتوبوس/مینی بوس		انحنای افقی و مسطح
			آزادراه		انحنای افقی و غیر مسطح
			بزرگراه		درون شهری
کاربری زمین	۳۵-۴۴	نوع راه	سایر	نوع منطقه	برون شهری
			اتوبوس/مینی بوس		درون شهری
			آزادراه		درون شهری
			بزرگراه		برون شهری
کاربری زمین	۴۵-۵۴	نوع راه	اتوبوس/مینی بوس	نوع منطقه	درون شهری
			سایر		برون شهری
			آزادراه		درون شهری
			بزرگراه		برون شهری
کاربری زمین	۵۵-۶۴	نوع راه	اتوبوس/مینی بوس	نوع منطقه	درون شهری
			سایر		برون شهری
			آزادراه		درون شهری
			بزرگراه		برون شهری
کاربری زمین	بالای ۶۵	نوع راه	اتوبوس/مینی بوس	نوع منطقه	درون شهری
			سایر		برون شهری
			آزادراه		درون شهری
			بزرگراه		برون شهری

جدول ۲. خلاصه‌ای از درصد فراوانی سطوح متغیرها در خوشه‌های مختلف

متغیر	خوشه سطح	۱	۲	۳	۴
نوع منطقه	برون شهری	۰	۶۸	۰	۰
	درون شهری	۱۰۰	۳۲	۱۰۰	۱۰۰
جنسیت عابر	مونث	۰	۳۰	۳۴	۱۰۰
	مذکر	۱۰۰	۷۰	۶۶	۰
نوع وسیله برخوردکننده با عابر	سواری	۱۰۰	۶۹	۱۰	۱۰۰
	موتورسیکلت	۰	۱۰	۵۱	۰
	وانت بار	۰	۱۳	۲۵	۰
	سایر	۰	۸	۱۴	۰

انواع وسایل نقلیه سنگین حمل بار در راه‌های برون‌شهری بیشتر از راه‌های درون‌شهری است.

نوع وسیله نقلیه درگیر در تصادف در هر خوشه در نمودار شکل ۲ نمایش داده شده است. همان‌طور که مشاهده می‌شود، دو خوشه ۲ و ۳ از لحاظ سهم موتورسیکلت و وانت بار، وضعیت خاصی دارند، در خوشه ۲، حدود ۱۰ درصد از تصادفات مربوط به موتورسیکلت و ۱۳ درصد مربوط به وانت بار بوده است. همچنین، در خوشه ۳، ۵۱ درصد از کل تصادفات مربوط به موتورسیکلت و ۲۵ درصد مربوط به وانت بار بوده است.



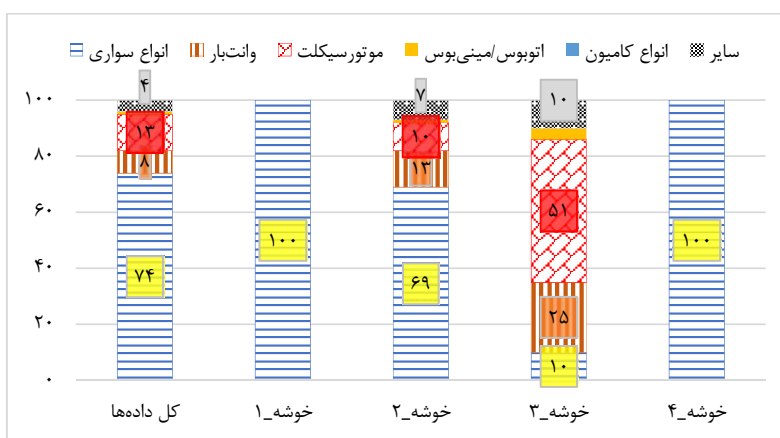
شکل ۱. وضعیت مصدومیت عابر پیاده در تصادفات هر یک از خوشه‌ها

به منظور بررسی ویژگی‌های هر خوشه به صورت جزئی‌تر، نسبت تصادفات فوتی و جرحی در هر یک از خوشه‌ها در جدول ۳ ارائه شده است. همان‌طور که مشاهده می‌شود، درصد تصادفات فوتی در خوشه ۲ به میزان قابل توجهی بیشتر از سایر خوشه‌هاست (۸ درصد از تصادفات این خوشه فوتی بوده است). این موضوع نشان دهنده بیشتر بودن ریسک فوت در تصادفات این خوشه است.

جدول ۳. درصد تصادفات فوتی و جرحی در هر یک از خوشه‌ها

	خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴
جرحی	۹۸/۶	۹۲	۹۸/۳	۹۹/۱
فوتی	۱/۴	۸	۱/۷	۰/۹
مجموع	۱۰۰	۱۰۰	۱۰۰	۱۰۰

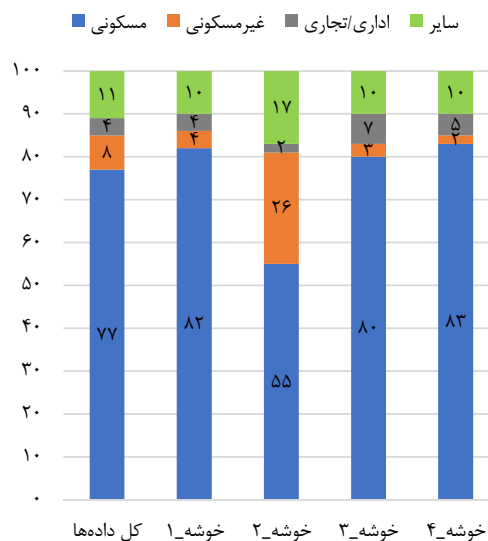
تصادفات خوشه ۲ برای عابرین پیاده نیز بسیار مرگبارتر از تصادفات سایر خوشه‌ها بوده است. همان‌طور که در نمودار شکل ۱ مشاهده می‌شود، ۷/۱ درصد از عابرین درگیر در تصادفات این خوشه فوت شده‌اند؛ که معادل است با ۵۷ درصد از کل آمار عابرین فوت شده در تصادفات مورد مطالعه در پژوهش حاضر. با توجه به اینکه بیش از دوسوم از تصادفات این خوشه در راه‌های برون‌شهری بوده است، به نظر می‌رسد بالا بودن سرعت تردد و بالتبع سرعت برخورد خودروها با عابرین پیاده منجر به افزایش احتمال فوت شده است. همچنین، مواجهه عابرین با



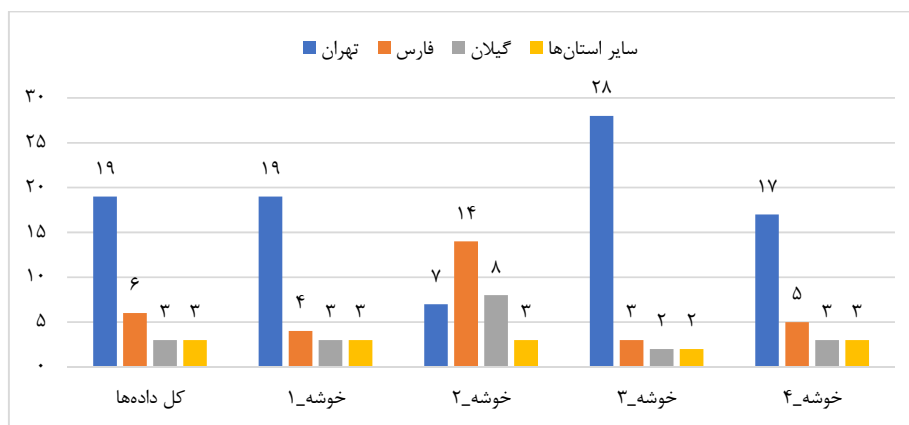
شکل ۲. نوع وسیله نقلیه درگیر در تصادفات هر خوشه

سهم استان‌ها از تصادفات هر خوشه نیز در نمودار شکل ۴ نمایش داده شده است. به منظور جلوگیری از شلوغی نمودار، صرفاً استان‌های خاص نمایش داده شده و متوسط سایر استان‌ها نیز به صورت یک میله ارائه گردیده است. همان‌طور که مشاهده می‌شود، در خوشه ۲، درحالی‌که متوسط سهم استان‌ها از تصادفات این خوشه حدود ۳ درصد است؛ اما استان فارس ۱۴ درصد از تصادفات این خوشه را به خود اختصاص داده است. این سهم برای استان گیلان نیز قابل توجه بوده و برابر با ۸ درصد است. از سوی دیگر، درحالی‌که سهم استان تهران از کل تصادفات در نظر گرفته شده در این پژوهش ۱۹ درصد بوده، اما در خوشه ۲، تنها ۷ درصد از تصادفات به استان تهران تعلق داشته است. همچنین، در خوشه ۳، ۲۸ درصد از تصادفات به استان تهران تعلق داشته است. سهم هریک از گروه‌های سنی رانندگان درگیر در تصادفات هر خوشه نیز در نمودار شکل ۵ نمایش داده شده است. نکته قابل توجه در این نمودار آن است که سهم رانندگان زیر ۱۹ سال (بدون گواهینامه) در تصادفات خوشه ۳ برابر با ۸ درصد بوده که به میزان قابل توجهی بیشتر از سایر خوشه‌هاست. بررسی دقیق‌تر تصادفات خوشه ۳ نشان داد که ۶ واحد از این ۸ درصد، مربوط به راکبین موتورسیکلت زیر ۱۹ سال است. به عبارت دیگر، ۶ درصد از کل تصادفات خوشه ۳ مربوط به راکبین موتورسیکلت زیر ۱۹ سال (و بالتبع، بدون گواهینامه) است.

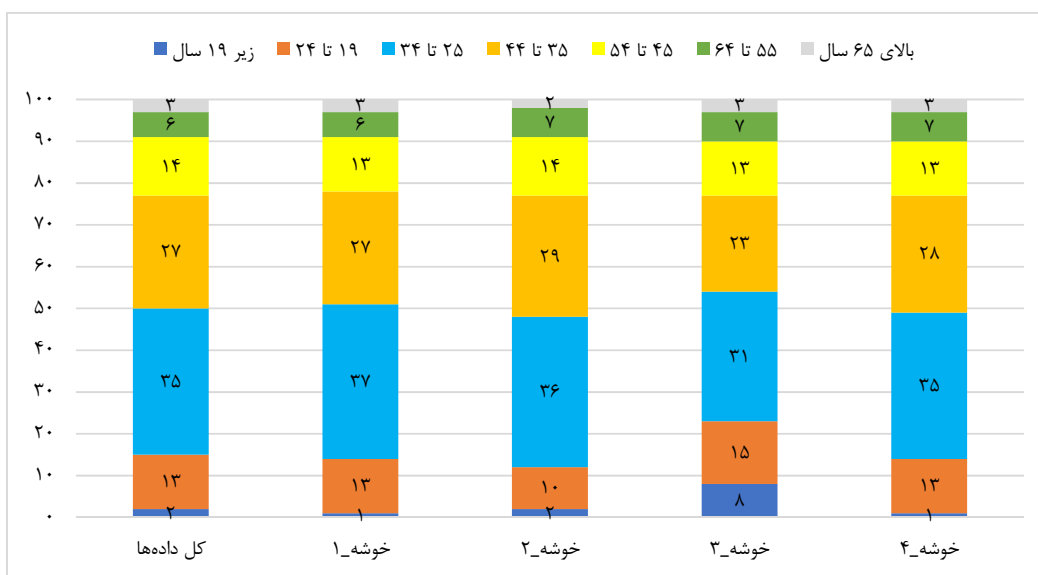
توزیع نوع کاربری زمین در محل تصادفات هر خوشه در شکل ۳ ارائه شده است. در رابطه با این متغیر نیز خوشه ۲ وضعیت ویژه‌ای دارد. در حالی‌که در تمامی سه خوشه دیگر، کمتر از ۴ درصد از تصادفات در مناطق غیرمسکونی رخ داده، اما در این خوشه، در ۲۶ درصد از موارد، تصادف در یک منطقه غیرمسکونی رخ داده است.



شکل ۳. نوع کاربری زمین در محل تصادفات هر خوشه



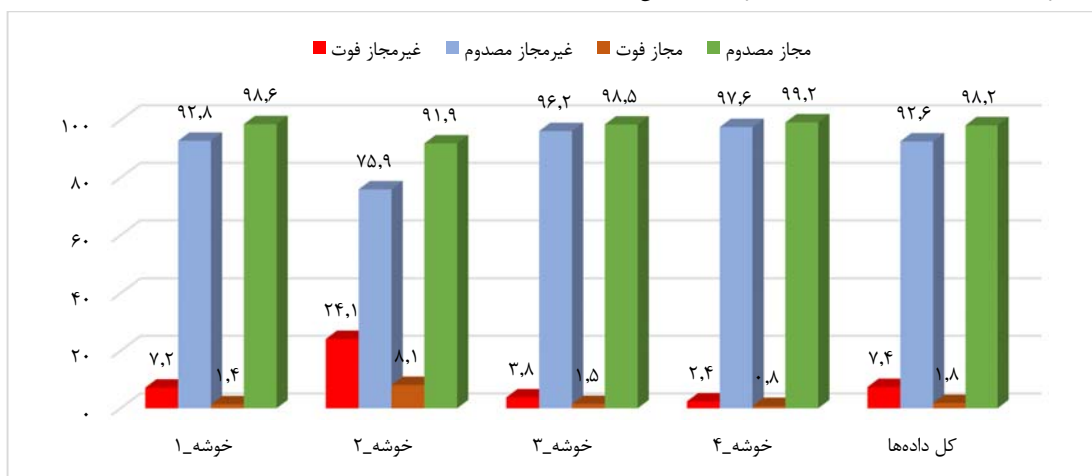
شکل ۴. سهم استان‌های منتخب از تصادفات هر خوشه



شکل ۵. سهم گروه سنی رانندگان درگیر در تصادفات هر خوشه

که این وضعیت در تصادفات خوشه ۲ بسیار شدیدتر بوده است. در این خوشه از تصادفات که عمدتاً در راه‌های برون‌شهری رخ داده، بیش از ۲۴ درصد از تصادفاتی که عابر در حال تردد از محل غیرمجاز بوده، منجر به فوت عابر شده است. به عبارت دیگر، در تصادفات این خوشه، از هر ۴ تصادفی که عابر در حال تردد از محل غیرمجاز بوده، یک مورد منجر به فوت عابر شده است.

سوال دیگری که می‌تواند مطرح شود عبارتست از اینکه آیا تردد از محل‌های غیرمجاز رابطه‌ای با فوتی‌های عابرین پیاده دارد؟ برای بررسی این موضوع، وضعیت مجاز بودن تردد عابر در محل تصادف به تفکیک نوع مصدومیت عابر در هر خوشه در شکل ۶ نمایش داده شده است. همان‌طور که مشاهده می‌شود، در مجموع کل داده‌ها، در تصادفاتی که عابر در حال تردد از محل مجاز بوده، ۱/۸ درصد از موارد منجر به فوت عابر شده، اما در تصادفاتی که عابر در حال تردد از محل غیرمجاز بوده، ۷/۴ درصد از موارد منجر به فوت شده است. بنابراین، می‌توان گفت به طور کلی، تردد از محل غیرمجاز منجر به افزایش احتمال فوت عابر در تصادف شده است. از سوی دیگر، مشاهده می‌شود

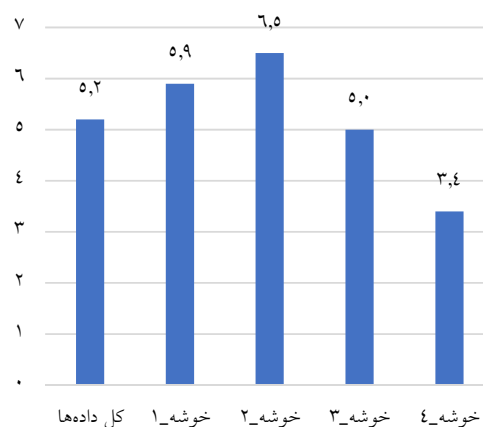


شکل ۶. وضعیت مجاز بودن تردد عابر در محل تصادف به تفکیک نوع مصدومیت عابر در هر خوشه

جدول‌ها، یک قاعده‌پیوندی را به همراه مقادیر لیفت، پشتیبان و اطمینان مربوطه نشان می‌دهد. این قواعد براساس مقدار لیفت و به‌صورت نزولی مرتب شده‌اند. در ادامه به بحث در مورد برخی از این قواعد پرداخته شده‌است.

قواعد پیوندی مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه اول «تصادفات درون‌شهری عابرین مذکر با خودرو سواری» در جدول ۴ نمایش داده شده است. همان‌طور که در شکل ۱ نشان داده شد، ریسک فوت عابرین در تصادفات این خوشه ۱/۳ درصد می‌باشد. این مقدار در واقع احتمال فوت عابر در شرایط عمومی در تصادفات این خوشه است. بدین ترتیب، قواعد ارائه شده در جدول ۴، نشان‌دهنده شرایط و عواملی است که در صورت برقراری، موجب افزایش احتمال فوت عابر در تصادفات این خوشه می‌شود. اولین قاعده در این جدول به تصادفات به وقوع پیوسته در بزرگراه‌های عبوری از مناطق غیرمسکونی اشاره می‌کند که راننده خودرو مذکر بوده است. در این شرایط، احتمال فوت عابر ۶/۷ برابر افزایش می‌یابد. بر مبنای قاعده دوم در صورت وقوع تصادف در بزرگراه‌های عبوری از مناطق غیرمسکونی در استان تهران، احتمال فوت عابر ۶/۶ برابر (نسبت به شرایط عمومی تصادفات این خوشه) بیشتر می‌شود. احتمال فوت در تصادف عابرین زیر ۶ سال در استان تهران و درحالیکه راننده مذکر بوده، ۵/۱ برابر افزایش می‌یابد (قاعده ۴). وقوع تصادف در بزرگراه و در بازه زمانی ۷ تا ۱۲ شب نیز با افزایش ۴/۹ برابری احتمال فوت عابر همراه است (قاعده ۶). به طور کلی بر مبنای قاعده ۹، وقوع تصادف در بزرگراه احتمال فوت عابر را ۴/۵ برابر افزایش می‌دهد. استان‌های اصفهان و کرمانشاه نیز از جمله استان‌هایی هستند که احتمال فوت عابر در برخی شرایط در این استان‌ها به میزان قابل توجهی بیشتر از مقدار متوسط این خوشه است (قواعد ۱۶، ۱۸، ۱۹ و ۲۱). همچنین، در شرایط مواجهه عابرین گروه سنی بالای ۶۵ سال با رانندگان ۱۹ تا ۲۴ سال (قاعده ۲۰) و ۲۵ تا ۳۴ سال (قاعده ۲۱) احتمال فوت عابر در تصادف به ترتیب ۲/۲ و ۲/۱ افزایش یافته است.

درصد موارد مقصر شناخته شدن عابر در تصادفات هر خوشه نیز در نمودار شکل ۷ نمایش داده شده است. همان‌طور که مشاهده می‌شود، خوشه ۲ (عمدتاً تصادفات برون‌شهری) دارای بیشترین و خوشه ۴ (تصادفات درون‌شهری عابرین مونث) دارای کمترین درصد موارد مقصر شناخته شدن عابر در تصادفات بوده‌اند. به عبارت دیگر، در راه‌های برون‌شهری (احتمالاً به دلیل فراهم نبودن زیرساخت‌های تردد غیرهمسطح عابرین از عرض راه‌ها و یا عدم استفاده عابرین از تسهیلات موجود)، درصد بیشتری از عابرین در وقوع تصادفات مقصر شناخته شده‌اند.



شکل ۷. درصد موارد مقصر شناخته شدن عابر در تصادفات هر خوشه

۴-۲- قواعد پیوندی

در مرحله دوم از داده‌کاوی، برای هر یک از خوشه‌های چهارگانه تصادفات شناسایی شده در مرحله اول، به‌طور جداگانه قواعد پیوندی استخراج شد. برای این کار، تمامی متغیرهای معرفی شده در جدول ۱ به‌عنوان «پیشین» و دو متغیر «شدت مصدومیت عابر» و «مقصر بودن عابر» به‌عنوان «نتیجه» در قواعد پیوندی انتخاب شدند.

برای هر یک از خوشه‌ها، تعداد مختلفی قواعد پیوندی کشف شد. مهمترین قواعد شناسایی شده برای هر خوشه در جدول‌های ۴ تا ۱۱ نمایش داده شده است. هرکدام از ردیف‌های این

جدول ۴. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه ۱

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	کاربری محل = غیرمسکونی و نوع راه = بزرگراه و جنسیت راننده = مرد	فوت	۶/۷	۱/۲	۸/۷
۲	کاربری محل = غیرمسکونی و نوع راه = بزرگراه و استان = تهران بزرگ	فوت	۶/۶	۱/۱	۸/۶
۳	کاربری محل = غیرمسکونی و نوع راه = بزرگراه	فوت	۶/۵	۱/۳	۸/۴
۴	گروه سنی عابر = زیر ۶ سال و استان = تهران بزرگ و جنسیت راننده = مرد	فوت	۵/۱	۱/۱	۶/۶
۵	نوع راه = بزرگراه و بازه زمانی = ۷ تا ۱۲ شب و جنسیت راننده = مرد	فوت	۴/۹	۱/۲	۶/۳
۶	نوع راه = بزرگراه و بازه زمانی = ۷ تا ۱۲ شب	فوت	۴/۹	۱/۳	۶/۳
۷	نوع راه = بزرگراه و وضعیت روشنایی راه = شب و جنسیت راننده = مرد	فوت	۴/۷	۱/۱	۶/۱
۸	نوع راه = بزرگراه و استان = تهران بزرگ	فوت	۴/۵	۲/۸	۵/۸
۹	نوع راه = بزرگراه	فوت	۴/۵	۴/۱	۵/۸
۱۰	کاربری محل = غیرمسکونی و بازه زمانی = ۷ تا ۱۲ شب	فوت	۳/۵	۱/۳	۴/۶
۱۱	کاربری محل = غیرمسکونی	فوت	۳/۵	۳/۸	۴/۵
۱۲	نوع راه = بزرگراه و کاربری محل = مسکونی	فوت	۳/۴	۲/۴	۴/۳
۱۳	کاربری محل = غیرمسکونی و گروه سنی راننده = ۲۵ تا ۳۴ سال و جنسیت راننده = مرد	فوت	۳/۲	۱/۱	۴/۲
۱۴	بازه زمانی = ۱۲ شب تا ۵ صبح	فوت	۲/۹	۴/۱	۳/۸
۱۵	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و نوع راه = خیابان اصلی	فوت	۲/۸	۱/۶	۳/۶
۱۶	استان = اصفهان و گروه سنی عابر = بالای ۶۵ سال	فوت	۲/۷	۱/۰	۳/۵
۱۷	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و کاربری محل = مسکونی	فوت	۲/۴	۱/۷	۳/۱
۱۸	استان = کرمانشاه	فوت	۲/۳	۳/۳	۲/۹
۱۹	استان = کرمانشاه و کاربری محل = مسکونی	فوت	۲/۲	۲/۳	۲/۹
۲۰	گروه سنی عابر = بالای ۶۵ سال و گروه سنی راننده = ۱۹ تا ۲۴ سال	فوت	۲/۲	۱/۵	۲/۹
۲۱	استان = کرمانشاه و گروه سنی راننده = ۲۵ تا ۳۴ سال و جنسیت راننده = مرد	فوت	۲/۲	۱/۱	۲/۸
۲۲	استان = تهران بزرگ و رنگ لباس عابر = تیره و وضعیت روشنایی راه = شب و جنسیت راننده = مرد	فوت	۲/۱	۱/۱	۲/۸

و به خصوص در سطح بزرگراه‌ها انتظار مواجهه با عابر را ندارند، سرعت برخورد نیز افزایش یافته و بالتبع، احتمال فوت عابر افزایش می‌یابد. بر مبنای قاعده ۹، اندرکنش مناطق غیرمسکونی و وقوع تصادف در شب نیز احتمال فوت عابر را ۳/۲ برابر بیشتر می‌کند. قواعد پیوندی مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه دوم «تصادفات برون‌شهری» در جدول ۶ نمایش داده شده است. با توجه به شکل ۱، ریسک فوت عابری در تصادفات این خوشه ۷/۱ درصد می‌باشد. این مقدار در واقع احتمال فوت عابر در شرایط عمومی در تصادفات این خوشه است. بر این مبنای قواعد ارائه شده در جدول ۶، بیانگر شرایطی است که در صورت برقراری، موجب افزایش احتمال فوت عابر در تصادفات این خوشه می‌شود.

قواعد کشف شده که در جدول ۵، نشان‌دهنده شرایطی است که احتمال مقصر شناخته شدن عابر در تصادفات خوشه اول بیشتر می‌شود. برای مثال، بر مبنای قاعده ۱، در تصادفات رخ داده در قطعاتی از بزرگراه‌های استان تهران که از مناطق غیرمسکونی عبور می‌کنند، احتمال مقصر شناخته شدن عابر ۹/۲ برابر بیشتر است. قبلاً در قالب قاعده دوم از جدول ۴ مشاهده کرده بودیم که در صورت وقوع تصادف در بزرگراه‌های عبوری از مناطق غیرمسکونی در استان تهران، احتمال فوت عابر ۶/۶ برابر می‌شود. بنابراین، از تجمیع این دو قاعده می‌توان نتیجه گرفت که در این گونه موقعیت‌ها، هم احتمال فوت عابر زیاد است و هم احتمال مقصر بودن عابر (احتمالاً عبور عرضی همسطح از سطح بزرگراه). با توجه به اینکه رانندگان در مناطق غیرمسکونی

تصادف در بزرگراه و در بازه زمانی ۷ تا ۱۲ شب نیز از جمله شرایط افزایش احتمال فوت عابر است که در چندین مورد از قواعد کشف شده به همراه عوامل دیگر قرار گرفته است (قواعد ۸، ۱۲، ۱۵، ۱۶). تصادف به وقوع پیوسته در بازه زمانی ۱۲ شب تا ۵ صبح در مناطق غیرمسکونی نیز با افزایش قابل ملاحظه احتمال فوت عابر همراه است (قاعده ۲۵). به علاوه، تردد عابرین بالای ۶۵ سال در مناطق غیرمسکونی (قاعده ۲۸)، به خصوص در جاده‌های اصلی برون‌شهری (قاعده ۱۸) از جمله شرایط افزایش احتمال فوت عابر است.

به طور کلی، به نظر می‌رسد به دلیل عدم انتظار رانندگان برای مواجهه با عابر در حال تردد در مناطق غیرمسکونی، تصادفات در این مناطق (به خصوص در قطعات آزادراهی و بزرگراهی عبوری از این مناطق) با شدت بیشتری همراه است.

بر مبنای قاعده ۱، برخورد عابر با کامیون کشنده در منطقه برون‌شهری، احتمال فوت عابر را ۴ برابر افزایش می‌دهد. به طور کلی در صورت برخورد عابر با کامیون کشنده احتمال فوت عابر ۳/۹ برابر افزایش می‌یابد (قاعده ۶).

این قواعد بیانگر بالا بودن احتمال فوت عابر در تصادف با وسایل نقلیه سنگین است. این یافته مطابق با یافته‌های مطالعات پیشین است (Salehian et al., 2023). قاعده دوم بیان می‌کند که در صورت وقوع تصادف در بزرگراه برون‌شهری در اوقات شب، احتمال فوت عابر ۳/۹ برابر افزایش می‌یابد. قاعده ۱۰ بیان می‌کند که در آزادراه‌های برون‌شهری که از زمین‌های غیرمسکونی عبور می‌کند، تصادف با عابر احتمال فوت را ۳/۷ برابر افزایش می‌دهد.

جدول ۵. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال مقصر بودن عابر در تصادفات خوشه ۱

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	کاربری محل = غیرمسکونی و نوع راه = بزرگراه و استان = تهران بزرگ	عابر مقصر	۹/۲	۱/۱	۵۲/۴
۲	کاربری محل = غیرمسکونی و نوع راه = بزرگراه	عابر مقصر	۸/۲	۱/۳	۴۶/۵
۳	کاربری محل = غیرمسکونی و استان = تهران بزرگ	عابر مقصر	۷/۹	۱/۳	۴۴/۹
۴	نوع راه = بزرگراه و استان = تهران بزرگ	عابر مقصر	۷/۵	۲/۸	۴۲/۸
۵	نوع راه = بزرگراه و وضعیت روشنایی راه = شب	عابر مقصر	۶/۵	۱/۱	۳۷/۱
۶	نوع راه = بزرگراه و استان = تهران بزرگ و کاربری محل = مسکونی	عابر مقصر	۶/۵	۱/۶	۳۷/۰
۷	نوع راه = بزرگراه و کاربری محل = مسکونی	عابر مقصر	۵/۰	۲/۴	۲۸/۴
۸	کاربری محل = غیرمسکونی	عابر مقصر	۳/۴	۳/۸	۱۹/۵
۹	کاربری محل = غیرمسکونی و وضعیت روشنایی راه = شب	عابر مقصر	۳/۲	۱/۲	۱۸/۰
۱۰	استان = تهران بزرگ و وضعیت روشنایی راه = شب	عابر مقصر	۲/۸	۳/۴	۱۶/۲

جدول ۶. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه ۲

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	نوع خودرو = کامیون کشنده و نوع منطقه = برون شهری	فوت	۴/۰	۱/۳	۲۸/۱
۲	نوع راه = بزرگراه و وضعیت روشنایی راه = شب و نوع منطقه = برون شهری	فوت	۳/۹	۱/۳	۲۷/۱
۳	نوع خودرو = کامیون کشنده و جنسیت عابر = مرد	فوت	۳/۸	۱/۲	۲۶/۹
۴	نوع راه = بزرگراه و وضعیت روشنایی راه = شب	فوت	۳/۸	۱/۴	۲۶/۸
۵	نوع راه = بزرگراه و وضعیت روشنایی راه = شب و نوع منطقه = برون شهری و جنسیت عابر = مرد	فوت	۳/۸	۱/۱	۲۶/۷
۶	نوع خودرو = کامیون کشنده	فوت	۳/۸	۱/۴	۲۶/۵

۲۶/۵	۱/۷	۳/۸	فوت	نوع راه= آزاد راه و کاربری محل= غیرمسکونی و جنسیت عابر= مرد
۲۶/۵	۱/۱	۳/۸	فوت	نوع راه= بزرگراه و بازه زمانی= ۷ تا ۱۲ شب و نوع منطقه= برون شهری و نوع خودرو= سواری و جنسیت راننده= مرد
۲۶/۳	۱/۱	۳/۸	فوت	نوع راه= بزرگراه و وضعیت روشنایی راه= شب و جنسیت عابر= مرد
۲۵/۹	۱/۹	۳/۷	فوت	نوع راه= آزاد راه و کاربری محل= غیرمسکونی و نوع منطقه= برون شهری
۲۵/۶	۱/۹	۳/۷	فوت	نوع راه= آزاد راه و کاربری محل= غیرمسکونی
۲۵/۶	۱/۳	۳/۷	فوت	نوع راه= بزرگراه و بازه زمانی= ۷ تا ۱۲ شب و جنسیت عابر= مرد و جنسیت راننده= مرد
۲۵/۵	۱/۱	۳/۶	فوت	نوع خودرو= کامیون و کاربری محل= غیرمسکونی و نوع منطقه= برون شهری
۲۵/۳	۱/۱	۳/۶	فوت	نوع خودرو= کامیون و کاربری محل= غیرمسکونی
۲۵/۲	۱/۷	۳/۶	فوت	نوع راه= بزرگراه و بازه زمانی= ۷ تا ۱۲ شب
۲۵/۱	۱/۶	۳/۶	فوت	نوع راه= بزرگراه و بازه زمانی= ۷ تا ۱۲ شب و نوع منطقه= برون شهری
۲۴/۶	۱/۰	۳/۵	فوت	نوع راه= بزرگراه و رنگ لباس عابر= تیره
۲۴/۳	۱/۱	۳/۵	فوت	گروه سنی عابر= بالای ۶۵ سال و کاربری محل= غیرمسکونی و نوع راه= جاده اصلی و نوع منطقه= برون شهری
۲۴/۱	۱/۳	۳/۵	فوت	نوع راه= آزاد راه و کاربری محل= غیرمسکونی و نوع منطقه= برون شهری و نوع خودرو= سواری
۲۳/۶	۲/۶	۳/۴	فوت	نوع راه= آزاد راه و نوع منطقه= برون شهری و جنسیت عابر= مرد
۲۳/۵	۲/۵	۳/۴	فوت	نوع راه= آزاد راه و نوع منطقه= برون شهری و جنسیت عابر= مرد و جنسیت راننده= مرد
۲۳/۱	۱/۴	۳/۳	فوت	رنگ لباس عابر= تیره و وضعیت روشنایی راه= شب و کاربری محل= غیرمسکونی و نوع منطقه= برون شهری
۲۲/۸	۲/۷	۳/۳	فوت	نوع راه= آزاد راه و جنسیت عابر= مرد
۲۲/۸	۱/۷	۳/۳	فوت	بازه زمانی= ۱۲ شب تا ۵ صبح و کاربری محل= غیرمسکونی
۲۲/۶	۳/۱	۳/۲	فوت	نوع راه= آزاد راه و نوع منطقه= برون شهری
۲۲/۳	۱/۶	۳/۲	فوت	رنگ لباس عابر= تیره و وضعیت روشنایی راه= شب و کاربری محل= غیرمسکونی
۲۱/۸	۱/۸	۳/۱	فوت	وضعیت روشنایی راه= شب با نور ناکافی و کاربری محل= غیرمسکونی
۲۱/۸	۲/۶	۳/۱	فوت	گروه سنی عابر= بالای ۶۵ سال و کاربری محل= غیرمسکونی

بود. قواعد مشابهی نیز برای استان‌های تهران و قزوین شناسایی شده است (قواعد ۱۱، ۱۲، ۱۴، ۱۵).

قواعد پیوندی مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه سوم «تصادفات درون شهری عابرین با موتورسیکلت، وانت‌بار و وسایل نقلیه سنگین» در جدول ۸ نمایش داده شده است. احتمال فوت عابرین در تصادفات این خوشه به طور متوسط ۱/۵ درصد می‌باشد (شکل ۱).

همان‌طور که در جدول ۸ مشاهده می‌شود، در قواعد این خوشه نیز برخورد با کامیون موجب افزایش احتمال فوت عابر شده است (قواعد ۱ تا ۹). به خصوص در بازه زمانی ۸ صبح تا ۲ بعد از ظهر (قاعده ۱)، در مناطق مسکونی (قاعده ۲) و در برخورد با عابرین مذکر (قاعده ۷). شایان توجه است که بیشتر بودن احتمال

قواعد کشف شده که در جدول ۷، نشان‌دهنده شرایطی است که احتمال مقصر شناخته شدن عابر در تصادفات خوشه ۲ را بیشتر می‌کند. برای مثال، قاعده اول در این جدول بیان می‌کند که در تصادفات عابرین مذکر در آزادراه و در قطعات عبوری از مناطق غیرمسکونی، احتمال مقصر شناخته شدن عابر ۸/۲ برابر بیشتر از مقدار متوسط برای تصادفات این خوشه است. بر مبنای قاعده ۷، در این خوشه برای تصادفاتی که در «استان البرز» رخ داده است احتمال مقصر شناخته شدن عابر (نسبت به متوسط تصادفات این خوشه) ۳/۱ برابر بیشتر بوده است. تصادفات این خوشه در استان البرز در صورتی که در راه‌های برون‌شهری رخ داده باشد (قاعده ۵) احتمال مقصر بودن عابر ۵/۹ برابر خواهد

برابر افزایش می‌دهد. تصادفات به وقوع پیوسته در مناطق مسکونی در استان مازندران نیز با افزایش احتمال تصادف همراه بوده است (قاعده ۱۱). همچنین، وقوع تصادف عابر مذکر در بزرگراه‌های استان تهران نیز به افزایش ۳ برابری احتمال فوت عابر انجامیده است (قاعده ۱۵). تصادفات گروه سنی بالای ۶۵ سال در استان اصفهان نیز با افزایش قابل توجه در احتمال فوت همراه بوده است (قاعده ۱۹).

فوت عابر مذکر می‌تواند به دلیل بیشتر بودن ریسک‌پذیری مردان نسبت به زنان باشد که موجب می‌شود عابریین مذکر رفتارهای پرریسک‌تری را انجام داده و شرایط وقوع تصادف فوتی را فراهم نمایند. البته این فرضیه نیازمند بررسی در یک مطالعه جداگانه است اما منطبق بر یافته‌های مطالعات پیشین است (Salehian et al., 2023). استان مازندران نیز حضور قابل ملاحظه‌ای در قواعد این خوشه دارد. برای مثال قاعده ۱۳ بیان می‌کند که وقوع تصادف عابر مذکر در استان مازندران احتمال فوت عابر را ۳/۶

جدول ۷. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال مقصر بودن عابر در تصادفات خوشه ۲

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	نوع راه = آزاد راه و کاربری محل = غیر مسکونی و جنسیت عابر = مرد	عابر مقصر	۸/۲	۱/۷	۵۰/۱
۲	نوع راه = آزاد راه و کاربری محل = غیر مسکونی و نوع منطقه = برون شهری	عابر مقصر	۸/۲	۱/۹	۵۰/۰
۳	نوع راه = آزاد راه و نوع منطقه = برون شهری و جنسیت عابر = مرد	عابر مقصر	۷/۱	۲/۶	۴۳/۴
۴	نوع راه = آزاد راه	عابر مقصر	۶/۵	۳/۳	۳۹/۳
۵	استان = البرز و نوع منطقه = برون شهری	عابر مقصر	۵/۹	۱/۲	۳۵/۹
۶	استان = البرز و جنسیت عابر = مرد	عابر مقصر	۳/۵	۱/۷	۲۱/۱
۷	استان = البرز	عابر مقصر	۳/۱	۲/۳	۱۸/۸
۸	بازه زمانی = ۱۲ شب تا ۵ صبح و کاربری محل = غیر مسکونی و نوع منطقه = برون شهری	عابر مقصر	۲/۹	۱/۶	۱۷/۶
۹	بازه زمانی = ۱۲ شب تا ۵ صبح و کاربری محل = غیر مسکونی و جنسیت عابر = مرد	عابر مقصر	۲/۹	۱/۵	۱۷/۴
۱۰	بازه زمانی = ۱۲ شب تا ۵ صبح و کاربری محل = غیر مسکونی	عابر مقصر	۲/۸	۱/۷	۱۶/۸
۱۱	استان = تهران و نوع منطقه = برون شهری	عابر مقصر	۲/۷	۱/۲	۱۶/۳
۱۲	استان = قزوین و نوع منطقه = برون شهری و جنسیت عابر = مرد	عابر مقصر	۲/۷	۱/۱	۱۶/۲
۱۳	استان = تهران و نوع منطقه = برون شهری	عابر مقصر	۲/۴	۱/۰	۱۴/۸
۱۴	استان = قزوین و نوع منطقه = برون شهری	عابر مقصر	۲/۴	۱/۴	۱۴/۵

جدول ۸. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه ۳

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	نوع خودرو = کامیون و بازه زمانی = ۸ صبح تا ۲ بعدازظهر	فوت	۷/۰	۱/۶	۱۰/۵
۲	نوع خودرو = کامیون و کاربری محل = مسکونی	فوت	۶/۹	۲/۶	۱۰/۳
۳	نوع خودرو = کامیون و بازه زمانی = ۸ صبح تا ۲ بعدازظهر و کاربری محل = مسکونی	فوت	۶/۸	۱/۲	۱۰/۲
۴	نوع خودرو = کامیون	فوت	۶/۷	۳/۴	۱۰/۱
۵	نوع خودرو = کامیون و جنسیت عابر = مرد و نوع راه = خیابان اصلی و کاربری محل = مسکونی	فوت	۶/۷	۱/۳	۱۰/۰
۶	نوع خودرو = کامیون و نوع راه = خیابان اصلی و کاربری محل = مسکونی	فوت	۶/۶	۱/۸	۹/۸
۷	نوع خودرو = کامیون و جنسیت عابر = مرد	فوت	۶/۶	۲/۶	۹/۸
۸	نوع خودرو = کامیون و بازه زمانی = ۸ صبح تا ۲ بعدازظهر و جنسیت عابر = مرد	فوت	۶/۳	۱/۲	۹/۴
۹	نوع خودرو = کامیون و نوع راه = خیابان اصلی	فوت	۶/۲	۲/۲	۹/۳

۵/۷	۱/۱	۳/۸	فوت	نوع خودرو = مینی بوس	۱۰
۵/۶	۱/۳	۳/۷	فوت	استان = مازندران و جنسیت عابر = مرد و کاربری محل = مسکونی	۱۱
۵/۴	۱/۰	۳/۶	فوت	کاربری محل = غیرمسکونی و نوع راه = بزرگراه	۱۲
۵/۳	۱/۶	۳/۶	فوت	استان = مازندران و جنسیت عابر = مرد	۱۳
۴/۹	۱/۰	۳/۳	فوت	نوع راه = بزرگراه و بازه زمانی = ۷ تا ۱۲ شب	۱۴
۴/۴	۲/۳	۳/۰	فوت	نوع راه = بزرگراه و استان = تهران بزرگ و جنسیت عابر = مرد	۱۵
۴/۲	۳/۸	۲/۸	فوت	نوع راه = بزرگراه	۱۶
۴/۱	۲/۵	۲/۸	فوت	استان = مازندران	۱۷
۴/۱	۱/۵	۲/۷	فوت	استان = مازندران و نوع راه = خیابان اصلی و کاربری محل = مسکونی	۱۸
۴/۱	۱/۴	۲/۷	فوت	استان = اصفهان و گروه سنی عابر = بالای ۶۵ سال	۱۹
۳/۹	۱/۸	۲/۶	فوت	استان = مازندران و نوع راه = خیابان اصلی	۲۰
۳/۹	۲/۹	۲/۶	فوت	نوع راه = بزرگراه و استان = تهران بزرگ	۲۱
۳/۸	۳/۶	۲/۵	فوت	بازه زمانی = ۵ تا ۸ صبح	۲۲
۳/۵	۲/۰	۲/۴	فوت	نوع خودرو = اتوبوس و جنسیت عابر = مرد	۲۳

این خوشه نسبت به سایر خوشه‌ها کمترین احتمال فوت را تجربه کرده‌اند (نمودار شکل ۱). باید توجه داشت که توانایی جسمی زنان در تحمل نیروی تصادف کمتر از مردان است. بنابراین کمتر بودن احتمال فوت عابرین این خوشه احتمالاً به دلیل رفتارهای کم‌ریسک‌تر عابرین مونث نسبت به عابرین مذکر است (Salehian et al., 2023). به عبارت دیگر، احتمالاً عابرین مونث کمتر خود را در موقعیت تصادفات شدید (فوتی) قرار می‌دهند. قواعد ارائه شده در جدول ۱۰، بیانگر شرایطی است که در صورت وقوع همزمان، موجب افزایش احتمال فوت عابر در تصادفات این خوشه می‌شود. همان‌طور که مشاهده می‌شود، اولین قاعده مربوط به تصادف عابرین (مونث) بالای ۶۵ سال در خیابان‌های اصلی در استان اصفهان است که این شرایط موجب افزایش احتمال فوت به میزان ۴/۲ برابر شده است (قاعده مشابهی نیز قبلاً در رابطه با تصادفات خوشه ۳ شناسایی شده بود). وقوع تصادف در بازه زمانی ۱۲ شب تا ۵ صبح نیز موجب افزایش احتمال فوت عابرین شده است (قاعده ۲). بر مبنای قاعده ۴ وقوع تصادف در بزرگراه احتمال فوت عابر در تصادفات این خوشه را ۳/۵ برابر می‌کند. به علاوه، تصادفات عابرین (مونث) با سن بالای ۶۵ سال، با رنگ لباس تیره و در تاریکی شب نیز از جمله شرایطی است که منجر به افزایش احتمال فوت عابر شده است (قاعده ۱۱). این یافته نیز در راستای نتایج مطالعات پیشین در رابطه با اثر تعاملی تاریکی هوا و تیره

قواعد کشف شده که در جدول ۹، نشان‌دهنده شرایطی است که احتمال مقصر شناخته شدن عابر در تصادفات خوشه ۳ افزایش می‌شود. برای مثال، قاعده اول در این جدول بیان می‌کند که در قطعاتی از بزرگراه‌ها که از مناطق غیرمسکونی عبور می‌کنند، احتمال مقصر شناخته شدن عابر ۷/۷ برابر است. احتمال مقصر بودن عابرین پیاده در تصادفات بزرگراه‌ها به خصوص در بازه زمانی ۷ تا ۱۲ شب بسیار بیشتر است (قاعده ۷). زیاد بودن احتمال مقصر شناخته شدن عابرین پیاده در تصادفات بزرگراهی استان تهران (به خصوص عابرین مذکر و در مناطق غیرمسکونی) نیز بسیار قابل ملاحظه است (قواعد ۸ و ۱۱). موضوع دیگری که با مقایسه میان قواعد شناسایی شده در جدول ۸ با جدول ۹ مشاهده می‌شود، عدم حضور موتورسیکلت در قواعد جدول ۸ و حضور آن در برخی از قواعد جدول ۹ است. به نظر می‌رسد احتمالاً به دلیل کم بودن اندازه حرکت (مومنتوم) موتورسیکلت، تصادفات موتورسیکلت-عابر اثر چندانی در افزایش احتمال فوت عابر ندارد. اما در بخش قابل توجهی از تصادفات موتورسیکلت-عابر (به خصوص در استان تهران)، عابرین پیاده به عنوان مقصر شناخته شده‌اند (قواعد ۲، ۴ و ۱۰).

قواعد پیوندی مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه چهارم «تصادفات درون‌شهری عابرین مونث با خودرو سواری» در جدول ۱۰ نمایش داده شده است. با توجه به شکل ۱، به طور کلی ۰/۸ درصد از تصادفات این خوشه منجر به فوت عابر شده است. بر این مبنای، عابرین پیاده در تصادفات

برای تصادفات این خوشه است. قواعد ۱، ۲ و ۳ به موضوع تردد غیرمجاز عابرین از بزرگراهها اشاره می‌کنند. همچنین، قواعد ۷ و ۸ بیان می‌کنند که در تصادفات استان قم به خصوص در بازه زمانی ۷ تا ۱۲ شب، احتمال مقصر بودن عابر در تصادف به میزان قابل توجهی بیشتر از حالت عادی در این خوشه است.

بودن لباس عابر بر افزایش احتمال فوت عابر در تصادف است (Hossain et al., 2024).

قواعد کشف شده که در جدول ۱۱ نشان‌دهنده شرایطی است که احتمال مقصر شناخته شدن عابر در تصادفات خوشه ۴ بیشتر می‌شود. برای مثال، قاعده اول در این جدول بیان می‌کند که در تصادفات به وقوع پیوسته در بزرگراه‌های استان تهران احتمال مقصر شناخته شدن عابر ۸/۹ برابر بیشتر از مقدار مورد انتظار

جدول ۹. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال مقصر بودن عابر در تصادفات خوشه ۳

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	کاربری محل = غیرمسکونی و نوع راه = بزرگراه	عابر مقصر	۷/۷	۱/۰	۳۶/۸
۲	نوع راه = بزرگراه و استان = تهران بزرگ و نوع خودرو = موتورسیکلت و جنسیت عابر = مرد	عابر مقصر	۷/۶	۱/۱	۳۶/۱
۳	نوع راه = بزرگراه و استان = تهران بزرگ و جنسیت عابر = مرد	عابر مقصر	۷/۳	۲/۳	۳۴/۶
۴	نوع راه = بزرگراه و استان = تهران بزرگ و نوع خودرو = موتورسیکلت	عابر مقصر	۷/۰	۱/۴	۳۳/۶
۵	کاربری محل = غیرمسکونی و استان = تهران بزرگ و جنسیت عابر = مرد	عابر مقصر	۷/۰	۱/۱	۳۳/۳
۶	نوع راه = بزرگراه و استان = تهران بزرگ	عابر مقصر	۶/۹	۲/۹	۳۲/۹
۷	نوع راه = بزرگراه و بازه زمانی = ۷ تا ۱۲ شب	عابر مقصر	۶/۸	۱/۰	۳۲/۳
۸	نوع راه = بزرگراه و استان = تهران بزرگ و جنسیت عابر = مرد و کاربری محل = مسکونی	عابر مقصر	۶/۸	۱/۴	۳۲/۲
۹	کاربری محل = غیرمسکونی و استان = تهران بزرگ	عابر مقصر	۶/۶	۱/۳	۳۱/۳
۱۰	نوع راه = بزرگراه و نوع خودرو = موتورسیکلت و جنسیت عابر = مرد	عابر مقصر	۶/۵	۱/۳	۳۰/۸
۱۱	نوع راه = بزرگراه و استان = تهران بزرگ و کاربری محل = مسکونی	عابر مقصر	۶/۲	۱/۸	۲۹/۷
۱۲	نوع راه = بزرگراه و جنسیت عابر = مرد	عابر مقصر	۵/۹	۳/۰	۲۸/۳
۱۳	نوع راه = بزرگراه	عابر مقصر	۵/۶	۳/۸	۲۶/۹
۱۴	نوع راه = بزرگراه و جنسیت عابر = مرد و کاربری محل = مسکونی	عابر مقصر	۵/۴	۱/۹	۲۵/۷

جدول ۱۰. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال فوت عابر در تصادفات خوشه ۴

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	نتیجه	پیشین			
۱	استان = اصفهان و گروه سنی عابر = بالای ۶۵ سال و نوع راه = خیابان اصلی	فوت	۴/۲	۱/۱	۳/۲
۲	بازه زمانی = ۱۲ شب تا ۵ صبح	فوت	۳/۷	۱/۷	۲/۹
۳	استان = اصفهان و گروه سنی عابر = بالای ۶۵ سال	فوت	۳/۷	۱/۴	۲/۸
۴	نوع راه = بزرگراه	فوت	۳/۵	۱/۹	۲/۷
۵	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و نوع راه = خیابان اصلی	فوت	۳/۴	۱/۵	۲/۶
۶	گروه سنی راننده = ۱۹ تا ۲۴ سال و گروه سنی عابر = بالای ۶۵ سال	فوت	۳/۴	۱/۳	۲/۶
۷	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و نوع راه = خیابان اصلی و کاربری محل = مسکونی	فوت	۳/۴	۱/۳	۲/۶
۸	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب	فوت	۳/۳	۱/۹	۲/۵
۹	استان = اصفهان و گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = روز	فوت	۳/۲	۱/۰	۲/۴

فصلنامه علمی پژوهشنامه حمل و نقل، سال بیست و دوم، دوره سوم، شماره ۸۴، پاییز ۱۴۰۴

۲/۴	۱/۷	۳/۱	فوت	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و کاربری محل = مسکونی	۱۰
۲/۴	۱/۰	۳/۱	فوت	گروه سنی عابر = بالای ۶۵ سال و بازه زمانی = ۷ تا ۱۲ شب و رنگ لباس عابر = تیره	۱۱
۲/۳	۲/۲	۳/۰	فوت	کاربری محل = غیرمسکونی	۱۲
۲/۲	۱/۲	۲/۹	فوت	بازه زمانی = ۱۲ شب تا ۵ صبح و نوع راه = خیابان اصلی	۱۳
۲/۲	۱/۱	۲/۹	فوت	استان = مازندران و رنگ لباس عابر = تیره و نوع راه = خیابان اصلی	۱۴
۲/۲	۲/۸	۲/۹	فوت	گروه سنی عابر = بالای ۶۵ سال و بازه زمانی = ۷ تا ۱۲ شب	۱۵
۲/۲	۱/۰	۲/۸	فوت	استان = مازندران و رنگ لباس عابر = تیره و کاربری محل = مسکونی و جنسیت راننده = مرد	۱۶
۲/۱	۱/۲	۲/۸	فوت	گروه سنی عابر = بالای ۶۵ سال و وضعیت روشنایی راه = شب و نوع راه = خیابان اصلی و کاربری محل = مسکونی و جنسیت راننده = مرد	۱۷
۲/۱	۲/۴	۲/۸	فوت	گروه سنی عابر = بالای ۶۵ سال و بازه زمانی = ۷ تا ۱۲ شب و کاربری محل = مسکونی	۱۸
۲/۱	۱/۱	۲/۷	فوت	استان = مازندران و رنگ لباس عابر = تیره و جنسیت راننده = مرد	۱۹
۲/۰	۱/۲	۲/۷	فوت	گروه سنی عابر = بالای ۶۵ سال و رنگ لباس عابر = تیره و گروه سنی راننده = ۲۵ تا ۳۴ سال و جنسیت راننده = مرد	۲۰

جدول ۱۱. برخی از قواعد کشف شده مربوط به شرایط افزایش احتمال مقصر بودن عابر در تصادفات خوشه ۴

شماره قاعده	قاعده پیوندی		لیفت	پشتیبان (%)	اطمینان (%)
	پیشین	نتیجه			
۱	نوع راه = بزرگراه و استان = تهران	عابر مقصر	۸/۹	۱/۱	۲۸/۸
۲	نوع راه = بزرگراه	عابر مقصر	۶/۶	۱/۹	۲۱/۳
۳	نوع راه = بزرگراه و کاربری محل = مسکونی	عابر مقصر	۵/۸	۱/۳	۱۸/۹
۴	کاربری محل = غیرمسکونی	عابر مقصر	۳/۲	۲/۲	۱۰/۳
۵	استان = قم و کاربری محل = مسکونی	عابر مقصر	۳/۰	۲/۳	۹/۸
۶	کاربری محل = غیرمسکونی و وضعیت روشنایی راه = روز و جنسیت راننده = مرد	عابر مقصر	۲/۸	۱/۲	۹/۱
۷	استان = قم و بازه زمانی = ۷ تا ۱۲ شب	عابر مقصر	۲/۸	۱/۰	۹/۰
۸	استان = قم	عابر مقصر	۲/۶	۳/۲	۸/۶
۹	استان = تهران بزرگ و گروه سنی عابر = ۳۵_۴۴	عابر مقصر	۲/۲	۱/۸	۷/۱

۵- نتیجه گیری

چهارگانه تصادفات حاکی از آن است که احتمال فوت عابر در تصادفات رخ داده در راه‌های برون‌شهری به میزان قابل توجهی بیشتر از راه‌های درون‌شهری است. علاوه بر این، از میان الگوهای شناسایی شده در پژوهش حاضر می‌توان به موارد زیر اشاره نمود.

-تاثیر تعاملی تاریکی هوا و تیره بودن رنگ لباس عابر موجب افزایش احتمال فوت عابر می‌شود. این موضوع را می‌توان به دلیل کاهش قابلیت رویت عابر توسط راننده، تاخیر در واکنش راننده در مواجهه با عابر و به تبع آن افزایش سرعت برخورد خودرو با عابر دانست؛ که موجب افزایش احتمال فوت عابر در اثر تصادف می‌گردد. در این راستا، آگاهی بخشی به مردم در رابطه

در مطالعه حاضر با استفاده از یک چارچوب ترکیبی داده‌کاوی به شناسایی الگوهای حاکم بر تصادفات ۷ ساله (۱۳۹۵ تا ۱۴۰۱) عابرین پیاده در ایران پرداخته شد. در مرحله اول از این رویکرد دومرحله‌ای، با بهره‌گیری از تحلیل خوشه‌بندی، داده‌های تصادفات در ۴ خوشه گروه‌بندی شد. این کار با هدف کاهش ناهمگنی پنهان میان داده‌های تصادفات انجام گردید. در مرحله دوم، از روش کشف قواعد پیوندی برای استخراج الگوهای حاکم بر هریک از خوشه‌های تصادف استفاده شد. در مطالعه حاضر، دو متغیر «فوت شدن عابر پیاده» و «مقصر بودن عابر در تصادف» به‌عنوان دو متغیر هدف در قواعد پیوندی در نظر گرفته شد. یافته‌های تحلیل‌های مقایسه‌ای انجام شده میان خوشه‌های

مربوط به ایمنی ترافیک را دارند. بنابراین، راه‌اندازی برنامه‌های آموزشی ویژه برای افزایش آگاهی این گروه نسبت به رفتارها و محل‌های پرخطر می‌تواند موجب کاهش تصادفات و تلفات این گروه از کاربران راه در تصادفات ترافیکی گردد. به‌علاوه، از این گروه می‌توان برای گسترش آموزش‌های ایمنی ترافیک به سایر اقشار جامعه با هزینه بسیار اندک (در محیط‌های خانوادگی) استفاده کرد.

برخی از استان‌ها (مانند تهران، البرز، اصفهان، مازندران و گیلان) الگوهای خاصی در تصادفات عابرین پیاده از خود نشان داده‌اند که بررسی دقیق‌تر دلایل وقوع این الگوها می‌تواند موضوع مطالعات آینده باشد.

۶- پی‌نوشت‌ها

1. Two-step Clustering
2. Silhouette Coefficient (SC)

۷- مراجع

-کی منش، محمود رضا و ذبیحیان، رسول (۱۴۰۲). ارزیابی عوامل موثر بر تصادفات عابرین پیاده در خیابان‌های شهری و ارائه مدلی برای پیش‌بینی این تصادفات (مطالعه موردی شهر زنجان). *پژوهشنامه حمل و نقل*، ۲۰(۳)، ۱۹۰-۱۷۱.

doi.org/10.22034/tri.2022.286782.2913

-مسلمی مهنی، صادق، نادران، علی و ناصر علوی، سید صابر (۱۴۰۳). مقایسه الگوریتم‌های یادگیری ماشین در انتخاب ویژگی و طبقه‌بندی شدت تصادفات عابر پیاده (مطالعه موردی: شهر قم). *پژوهشنامه حمل و نقل*. (پذیرفته شده برای چاپ).

doi.org/10.22034/tri.2024.443127.3227

-ندیمی، نوید، ناصرعلوی، سید صابر، خلیفه، وحید و امینیان، مهدی (۱۴۰۳). تحلیل و ارزیابی عوامل موثر بر تصادفات کاربران آسیب‌پذیر در معابر شهری با استفاده از استخراج قواعد انجمنی. *پژوهشنامه حمل و نقل*، ۲۲(۱)، ۸۲-۶۵.

doi.org/10.22034/tri.2024.430837.3216

-Alikhani, M., Nedaie, A., and Ahmadvand, A. (2013). Presentation of clustering-classification heuristic method for improvement accuracy in classification of severity of road accidents in Iran. *Safety science*, 60, 142-150.

با تردد در مسیرهای دارای روشنایی و استفاده از برجسب‌های شبرنگ بر روی لباس یا کفش‌ها می‌تواند به کاهش این نوع تصادفات کمک کند.

-به‌طور کلی در تصادفات به وقوع پیوسته در راه‌های برون‌شهری احتمال مقصر شناخته شدن عابر در تصادف بیشتر بوده است (احتمالاً به دلیل فراهم نبودن زیرساخت‌های موردنیاز برای تردد غیرهمسطح عابرین از عرض راه‌ها و یا عدم استفاده عابرین از تسهیلات عبور عرضی غیرهمسطح).

-همچنین، «تردد عرضی عابر از محل غیرمجاز» منجر به افزایش قابل توجه احتمال فوت عابر شده است. این موضوع به خصوص در راه‌های برون‌شهری نمایان‌تر بوده است. بنابراین، لازم است در قطعاتی از راه‌های برون‌شهری که از محدوده‌های مسکونی عبور می‌کند، محل‌های مجاز برای تردد همسطح عابرین ایجاد گردیده و با ایجاد تمهیداتی مانند روشنایی، نصب تابلو، کف نوشته و سایر موارد، رانندگان نیز نسبت به این موضوع آگاه شوند. در غیر این صورت، می‌بایست امکان تردد عرضی همسطح عابرین به کلی مسدود گردیده و زیرگذر/روگذرهایی برای عبور عرضی عابرین احداث گردد.

-در تصادفات به وقوع پیوسته در بزرگراه‌های عبوری از مناطق غیرمسکونی، هم احتمال فوت عابر و هم احتمال مقصر شناخته شدن عابر افزایش می‌یابد (احتمالاً عبور عرضی همسطح از سطح بزرگراه). با توجه به اینکه رانندگان در مناطق غیرمسکونی و به خصوص در سطح بزرگراه‌ها انتظار مواجهه با عابر را ندارند، سرعت برخورد نیز افزایش یافته و بالتبع، احتمال فوت عابر افزایش می‌یابد. بر این مبنا می‌توان گفت افزایش توجه به مسدود نمودن امکان تردد همسطح عابرین پیاده از سطح آزادراه‌ها و بزرگراه‌ها (درون‌شهری و برون‌شهری) و همزمان، فراهم نمودن امکانات کافی برای تردد غیرهمسطح عابر از عرض راه از اهمیت بسزایی برخوردار است.

-احتمال فوت عابرین با سن بالای ۶۵ سال بیشتر از سایر گروه‌های سنی بوده است. این احتمال در تعامل با برخی دیگر از متغیرها مانند رانندگان جوان (سن ۱۹ تا ۳۴ سال)، وقوع تصادف در بزرگراه یا خیابان اصلی، تاریکی هوا، تیره بودن رنگ لباس (به خصوص عابرین مونث)، مناطق غیرمسکونی و برون‌شهری، با افزایش قابل ملاحظه‌ای همراه بوده است. با توجه به سن بازنشستگی، این گروه از افراد به دلیل فراغت بال امکان شرکت در برنامه‌های آموزشی و افزایش آگاهی در مورد مسائل

- Lee, C., and Abdel-Aty, M. (2005). Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention*, 37(4), 775-786 .
- Ma, Z., Mei, G., and Cuomo, S. (2021). An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. *Accident Analysis and Prevention*, 160, 106322 .
- Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., and Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian-vehicle crashes in New York, US and Montreal, Canada. *Safety science*, 54, 27-37 .
- Montella, A., Aria, M., D'Ambrosio, A., and Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis and Prevention*, 49, 58-72 .
- Moudon, A. V., Lin, L., Jiao, J., Hurvitz, P., and Reeves, P. (2011). The risk of pedestrian injury and fatality in collisions with motor vehicles, a social ecological study of state routes and city streets in King County, Washington. *Accident Analysis and Prevention*, 43(1), 11-24 .
- Raihan, M. A., Hossain, M., and Hasan, T. (2018). Data mining in road crash analysis: the context of developing countries. *International Journal of Injury Control and Safety Promotion*, 25(1), 41-45.
- Salehian, A., Aghabayk, K., Seyfi, M., and Shiwakoti, N. (2023). Comparative analysis of pedestrian crash severity at United Kingdom rural road intersections and Non-Intersections using latent class clustering and ordered probit model. *Accident Analysis and Prevention*, 192, 107231 .
- Santos, K., Dias, J. P., and Amado, C. (2022). A literature review of machine learning algorithms for crash injury severity prediction. *Journal of safety research*, 80, 254-269 .
- Zajac, S. S., and Ivan, J. N. (2003). Factors influencing injury severity of motor vehicle-crossing pedestrian crashes in rural Connecticut. *Accident Analysis and Prevention*, 35(3), 369-379 .
- Ankunda, A., Ali, Y., and Mohanty, M. (2024). Pedestrian crash risk analysis using extreme value models: new insights and evidence. *Accident Analysis and Prevention*, 203, 107633 .
- Aziz, H., Ukkusuri, S. V., and Hasan, S. (2013). Exploring the determinants of pedestrian-vehicle crash severity in New York City. *Accident Analysis and Prevention*, 50, 1298-1309 .
- Dai, D. (2012). Identifying clusters and risk factors of injuries in pedestrian-vehicle crashes in a GIS environment. *Journal of transport geography*, 24, 206-214 .
- de Oña, J., López, G., Mujalli, R., and Calvo, F. J. (2013). Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. *Accident Analysis and Prevention*, 51, 1-10 .
- Depaire, B., Wets, G., and Vanhoof, K. (2008). Traffic accident segmentation by means of latent class clustering. *Accident Analysis and Prevention*, 40(4), 1257-1266 .
- Hossain, M. M., Zhou, H., Sun, X., Hossain, A., and Das, S. (2024). Crashes involving distracted pedestrians: Identifying risk factors and their relationships to pedestrian severity levels and distraction modes. *Accident Analysis and Prevention*, 194, 107359 .
- Iranitalab, A., and Khattak, A. (2017). Comparison of four statistical and machine learning methods for crash severity prediction. *Accident Analysis and Prevention*, 108, 27-37 .
- Kaufman, L., and Rousseeuw, P. J. (2009). *Finding Groups in Data: an Introduction to cluster Analysis*. John Wiley and Sons .
- Kim, J.-K., Ulfarsson, G. F., Shankar, V. N., and Kim, S. (2008). Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accident Analysis and Prevention*, 40(5), 1695-1702 .
- Kim, J.-K., Ulfarsson, G. F., Shankar, V. N., and Mannering, F. L. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis and Prevention*, 42(6), 1751-1758 .
- Kim, S. H. (2023). How heterogeneity has been examined in transportation safety analysis: A review of latent class modeling applications. *Analytic methods in accident research*, 100292 .

An Integrated Data Mining Model for Analyzing the Pedestrian Crash Patterns in Iran

Ali Tavakoli Kashani, Associate Professor, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran and Road Safety Research Center, Iran University of Science and Technology, Tehran, Iran.

Mohammad Mehdi Besharati, Assistant Professor, Department of Transportation Engineering, Isfahan University of Technology, Isfahan, Iran.

E-mail: besharati@iut.ac.ir

Received: May 2025- Accepted: August 2025

ABSTRACT

Pedestrians as one of the main groups of vulnerable road users account for a significant share of traffic crash fatalities. The purpose of this study is to investigate the pedestrian crash patterns and identify the set of factors that contribute to increase the fatality risk of pedestrians. To this end, pedestrian crash data in Iran during 2016 to 2022 was analyzed in a combined data mining framework. In the first step, crash data was divided into homogeneous groups using the clustering method. Then, the association rules discovery method was applied to identify the set of factors that contribute to increase the fatality risk of pedestrians as well as the probability of the pedestrians being at-fault in the crash. Results indicated that the interaction effects of factors such as darkness, pedestrian wearing dark clothes, being male and above 65-years old, crash being occurred on rural freeways and highways, collision with heavy vehicles, and crossing the road in unpermitted locations contribute to increase the fatality risk of pedestrians. Therefore, obstruction of the crossing passage of pedestrians and providing overpass/underpass might have a significant effect on reducing pedestrian fatal crashes. Moreover, providing illuminated crossings and encouraging the pedestrians to use these crossing might reduce the risk of occurrence and severity of such crashes.

Keywords: Fatal Crashes, Pedestrian, Clustering, Association Rules