

استاد: مشهدی فراهانی، حسین؛ (۱۳۹۷). «داده کاوی بر آمار تصادفات شهر اراک بر پایه نرم افزار وکا (weka)»، توسعه سازمانی پلیس، شماره ۶۵، صص ۱۱۷-۱۴۷.

داده کاوی بر آمار تصادفات شهر اراک بر پایه نرم افزار وکا (weka)

تاریخ دریافت مقاله: ۹۸/۰۴/۲۲

حسین مشهدی فراهانی^۱

تاریخ پذیرش مقاله: ۹۸/۱۰/۰۵



چکیده:

یکی از وظایف مهم پلیس، نظم بخشیدن و ایجاد انضباط ترافیکی در سطح جاده‌های کشور و شهرهای مختلف است. از آنجا که تعداد تصادفات در کشور ما نسبت به سایر کشورها بالاست، به منظور کاهش تعداد تصادفات اقدام‌های بسیاری صورت پذیرفته است؛ که می‌توان افزایش سطح بازدارندگی تخلفات جاده‌ای و اصلاح نقاط حادثه‌خیز و نیز افزایش حضور پلیس در جاده‌ها را برشمرد، اما با این همه میزان خسارت‌های جدی که به قتل و جرح نیز منجر می‌شود نشان‌دهنده ناکافی بودن و ناکارآمدی این اقدامات است. در این پژوهش، برای جلوگیری از این گونه تصادفات با استفاده از تکنیک‌های داده‌کاوی، علل تصادفات در سطح شهر اراک رتبه‌بندی شده است. سپس با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره و مدل تحلیل پوششی داده‌ها به ارائه راه‌کارهایی مناسب پرداخته شده و بهترین راه‌کار برای جلوگیری از تصادفات احتمالی انتخاب شده است. گفتنی است پژوهشگر تلاش کرده است راه‌کارهای مناسبی برای دستیابی به یک هدف‌گذاری مناسب در سال ۱۳۹۸ ارائه کند و در این راستا محدودیت‌های منابع انسانی، تجهیزات و بودجه اختصاص‌یافته را در نظر گرفته تا اهدافی که این پژوهش در پی آن است، به درستی محقق شود. نتیجه کلی به‌دست‌آمده از پژوهش حاکی از آن است که مهمترین عامل مؤثر بر تصادفات در شهر اراک، محل وقوع تصادفات است و پس از آن به ترتیب اولویت، کنترل نکردن وسیله نقلیه، تجاوز از سرعت مجاز و مطمئن، خستگی و خواب‌آلودگی، توجه نکردن به جلو و سیستم خودرو می‌باشد.

کلیدواژه‌ها:

تصادفات، داده‌کاوی، تکنیک‌های تصمیم‌گیری، بررسی عوامل، راه‌کار مناسب.

مقدمه

علوم و مهندسی نوین بر پایه به‌کارگیری مدل‌های بنیادی اولیه برای تحلیل نظام‌های فیزیکی، زیست‌شناسی و اجتماعی استوار هستند. چنین رویکردی با یک مدل علمی بنیادی آغاز می‌شود نظیر قوانین حرکت نیوتن یا معادلات ماکسول در الکترومغناطیس که براساس آنها کاربردهای گوناگونی در زمینه مهندسی مکانیک و مهندسی الکترونیک شکل می‌گیرد. در این روش، از داده‌های آزمایشی برای بازبینی مدل‌های بنیادی اولیه و برآورد برخی از مؤلفه‌ها که اندازه‌گیری آنها به‌طور مستقیم دشوار و گاهی اوقات غیرممکن است، استفاده می‌شود. با این حال در بسیاری از زمینه‌ها، اصول بنیادی اولیه ناشناخته بوده یا نظام‌های تحت بررسی و مطالعه برای قرار گرفتن در قالب‌های ریاضی بسیار پیچیده هستند. با افزایش کاربرد رایانه، بسیاری از داده‌ها از طریق چنین نظام‌هایی ایجاد می‌شوند. در صورت فقدان مدل‌های بنیادی اولیه، می‌توان از چنین داده‌هایی که به‌راحتی در دسترس هستند، برای برآورد و ایجاد مدل‌هایی با رابطه‌ای سودمند بین متغیرهای یک نظام استفاده کرد (یعنی وابستگی‌های ورودی-خروجی مجهول و ناشناخته). بنابراین، در حال حاضر یک تغییر الگو از مدل‌سازی و تحلیل کلاسیک بر پایه اصول اولیه به مدل‌های در حال پیشرفت و تحلیل‌های مربوط به‌طور مستقیم از داده‌ها وجود دارد.

ضرورت درک مجموعه داده‌های بزرگ، پیچیده و اطلاعات کامل و غنی در زمینه تجارت، علوم و مهندسی کمابیش رایج است. در دنیای تجارت، داده‌های شرکت‌ها و مشتریان، سرمایه‌های راهبردی هستند. توانایی استخراج دانش و اطلاعات مفید موجود در این داده‌ها و امکان استفاده از این دانش در جهان رقابتی امروز بیش از پیش حائز اهمیت است. به کل فرایند به‌کارگیری روش‌شناسی مبتنی بر رایانه از جمله روش‌های جدید برای دریافت دانش و اطلاعات داده‌ها، داده‌کاوی می‌گویند.

داده‌کاوی فرایندی تکرارپذیر است که پیشرفت در آن با کاوش از طریق روش‌های خودکار یا دستی صورت می‌گیرد. داده‌کاوی سودمندترین سناریوی تحلیلی-اکتشافی است که در آن تصور و برداشت از پیش تعیین‌شده‌ای درباره نتیجه چشمگیری که به دست می‌آید، وجود ندارد. در حقیقت داده‌کاوی، جست‌وجوی لازم برای یافتن اطلاعات کلی جدید، ارزشمند و غیربدیهی از میان حجم زیاد داده‌هاست. به تعبیر دیگر، داده‌کاوی تعامل همکاری بین انسان و رایانه است. بهترین نتایج از ایجاد تعادل بین دانش متخصصان در توصیف مسائل و اهداف با قابلیت‌های جست‌وجوی رایانه به دست می‌آید.

در عمل، دو هدف اصلی داده‌کاوی شامل پیشگویی و توصیف است. پیشگویی شامل به‌کارگیری برخی متغیرها یا گزینه‌ها در مجموعه داده‌ها برای پیشگویی و مقادیر ناشناخته یا دیگر متغیرهاست.

توصیف، از سوی دیگر بر یافتن الگوهای توصیف داده‌ها که توسط انسان‌ها قابل تفسیر هستند، تأکید دارد. بنابراین، می‌توان فعالیت‌های داده‌کاوی را در دو گروه زیر طبقه‌بندی کرد:

۱- داده‌کاوی پیشگویانه که الگویی از نظام را ارائه می‌دهد، توسط مجموعه داده‌های مشخصی توصیف می‌شود؛

۲- داده‌کاوی توصیفی که اطلاعات جدید و غیربدیهی را براساس مجموعه داده‌های موجود ارائه می‌دهد.

در طیف پیش‌بینی، پایان کار و هدف کلی داده‌کاوی، ایجاد مدلی است که به‌عنوان یک برنامه و شناسه اجرایی بتوان از آن برای طبقه‌بندی، پیش‌بینی، برآورد و دیگر اعمال مشابه استفاده کرد. از طرف دیگر در طیف توصیف، نهایت کار و هدف کلی به دست آوردن شناخت از نظام‌های تحلیل شده توسط الگوها و روابط آشکار در مجموعه داده‌های بزرگ است. اهمیت نسبی طیف «پیش‌بینی» و «توصیف» برای کاربردهای خاص داده‌کاوی می‌تواند به‌طور چشمگیری متفاوت باشد. اهداف «پیش‌بینی» و «توصیف» با استفاده از روش و تکنیک‌های داده‌کاوی محقق می‌شود (هان و کمبر، ۲۰۱۵).^۲

داده‌کاوی یکی از سریع‌ترین زمینه‌های رشد در صنعت رایانه است. زمانی که یک زمینه خاص و محدود در رشته رایانه یا آمار گشوده می‌شود، به سرعت پیشرفت و به سطح و جایگاه واقعی خود گسترش پیدا می‌کند. یکی از مهمترین توانایی‌های داده‌کاوی در دامنه وسیع «روش‌شناسی‌ها و تکنیک‌های آن» نمایان شده که می‌تواند برای میزبانی مجموعه مسائل مختلف نیز صادق باشد. داده‌کاوی یک فعالیت طبیعی است که بر مجموعه داده‌های حجیم و بزرگ اعمال می‌شود. داده‌کاوی می‌تواند رقابت و بازده بازاریابی را بهبود بخشیده و درآمدها، حمایت و رضایت مشتریان را جلب کند. همچنین از تکنیک‌های داده‌کاوی می‌توان برای حل مشکلات فرایند تجارت استفاده کرد که در حقیقت هدف، درک تقابل و ارتباط بین شیوه‌های کسب‌وکار و سازماندهی‌های لازم است.

بسیاری از مجریان قانون و واحدهای بازرسی ویژه که مأمور شناسایی فعالیت‌های کلاهبرداران و کشف روش‌های ارتکاب جرم هستند، از فرایند داده‌کاوی استفاده و از مزیت نسبی آن بهره‌مند می‌شوند. برای نمونه، این روش‌شناسی می‌تواند متخصصان و تحلیلگران را در تشخیص الگوهای بحران رفتاری در رابطه با مواد مخدر، معاملات و تراکنش‌های پولشویی، حرکات و فعالیت‌های گروه‌های آدم‌کشان زنجیره‌ای و شناسایی قاچاقچیان در نقاط مرزی یاری دهد. تکنیک‌های داده‌کاوی همچنین در مباحث

2. Han, J. and M. Kamber

مرتبط با مأموران اطلاعاتی توسط افرادی که بسیاری از منابع داده‌های بزرگ و حجیم را به عنوان بخشی از فعالیت‌های مربوط به امور امنیت ملی ذکر می‌کنند، مورد استفاده قرار می‌گیرد. از طرفی با توجه به حجم بالای تصادفات در ایران، تحلیل این تصادفات بسیار ضروری به نظر می‌رسد و با توجه به مطالب پیش‌گفته، استفاده از داده‌کاوی برای تحلیل این موضوع و بهره‌گیری از تکنیک‌های تصمیم‌گیری در محیط پیچیده کنونی نیز می‌تواند به ارائه راه‌کارهای مناسب برای کاهش تصادفات منجر شود.

بیان مسئله

در طول دهه گذشته با پیشرفت روزافزون کاربرد پایگاه داده‌ها، حجم داده‌های ثبت‌شده به‌طور متوسط هر ۵ سال ۲ برابر می‌شود. در این میان سازمان‌هایی موفق‌اند که بتوانند حداقل ۷ درصد داده‌هایشان را تحلیل کنند. تحقیقات نشان داده است که سازمان‌ها کمتر از یک درصد داده‌هایشان را برای تحلیل استفاده می‌کنند. به عبارت دیگر در حالی که غرق در داده‌ها هستند، تشنه دانش‌اند. بنابر اعلام دانشگاه MIT دانش نوین داده‌کاوی یکی از ده دانش در حال توسعه‌ای است که دهه آینده را با انقلاب فناوریانه مواجه می‌سازد. این فناوری امروزه دارای کاربرد بسیار وسیعی در حوزه‌های مختلف است، به‌گونه‌ای که امروزه حد و مرزی برای کاربرد این دانش در نظر نگرفته و زمینه‌های کاری این دانش را از ذرات کف اقیانوس‌ها تا اعماق فضا می‌دانند. امروزه، بیشترین کاربرد داده‌کاوی در بانک‌ها، مراکز صنعتی و کارخانجات بزرگ، مراکز درمانی و بیمارستان‌ها، مراکز تحقیقاتی، بازاریابی هوشمند و ... می‌باشد.

داده‌کاوی پل ارتباطی میان علم آمار، علم رایانه، هوش مصنوعی، الگوشناسی، فراگیری خودرو و بازنمایی بصری داده است. داده‌کاوی فرایندی پیچیده برای شناسایی الگوها و مدل‌های صحیح، جدید و به صورت بالقوه مفید، در حجم وسیعی از داده می‌باشد، به طریقی که این الگوها برای انسان‌ها درک‌پذیر باشند. داده‌کاوی نظیر یک محصول قابل خریداری نیست، بلکه یک رشته علمی و فرایندی است که بایستی به صورت طرح (پروژه) پیاده‌سازی شود. داده‌کاوی فرایندی تحلیلی است که برای کاوش داده‌ها (معمولاً حجم عظیمی از داده‌ها در زمینه‌های کسب‌وکار و بازار) صورت می‌گیرد و یافته‌ها با به‌کارگیری الگوهایی، احراز اعتبار می‌شوند. هدف اصلی داده‌کاوی، پیش‌بینی است.

یکی از وظایف بسیار مهم پلیس راهور، بررسی دلایل تصادفات در سطح شهر اراک و استان مرکزی است که این وظیفه با بررسی دلایل این تصادفات و پیدا کردن وزن هر یک از این عوامل و سپس ارائه راه‌حل برای برطرف کردن مشکلات در جاده‌ها حل می‌شود. به همین منظور برای رفع این مسئله با استفاده از تکنیک داده‌کاوی، ابتدا عوامل مؤثر بر تصادفات با استفاده از کروکی‌های تصادفات منجر به

فوت و جرح شناسایی نموده (مانند نوع خودرو، سیستم خودرو، ساعت تصادف، مکان تصادف و ...)، سپس با مراجعه به پایگاه داده، اطلاعات تصادفات گذشته را به چند دسته تقسیم کرده (مانند تصادفات قتل، جرح و خسارتی) و در ادامه به بررسی پایگاه داده می پردازیم و با به دست آوردن وزن هر یک از عوامل دخیل در تصادفات، الگویی ارائه کرده تا به وسیله آن نسبت به برطرف کردن آن عوامل در سطح استان با استفاده از تکنیک‌های جدید تصمیم‌گیری چندمعیاره و مدل تحلیل پوششی داده‌ها اقدام کرد.

الگوریتم‌های داده کاوی به چهار دسته اصلی تقسیم می‌شوند و در این پژوهش با استفاده از این الگوریتم‌ها به پرسش‌های اصلی این پژوهش پاسخ داده می‌شود. اولین دسته الگوریتم‌های داده کاوی، الگوریتم‌های انتخاب ویژگی است. با استفاده از این الگوریتم‌ها عوامل مؤثر بر تصادفات رتبه‌بندی می‌شود. گروه دوم الگوریتم‌ها، وظیفه دسته‌بندی بر عهده دارند که شامل الگوریتم‌های درخت تصمیم و الگوریتم‌های آماری می‌باشند، نظیر J48، Random Forest و Lib SVM. با استفاده از این الگوریتم‌ها نقاط تصادفات ارزیابی می‌شود. سومین دسته الگوریتم‌ها، الگوریتم‌های قانون رابطه می‌باشند و با استفاده از این الگوریتم‌ها قوانین پایگاه داده (نقاط دارای اولویت برای تصحیح و ارائه راه کار) مشخص می‌شوند. سپس با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره (MCDM) به بررسی ارائه راه کار مناسب برای جلوگیری از تصادفات پرداخته می‌شود که در این راستا با استفاده از تکنیک FAHP برای محاسبه وزن هر یک از عوامل دخیل در انتخاب بهترین راه کار با استفاده از نظر کارشناسان راهور استفاده می‌شود. در مرحله بعد راه کارهایی برای جلوگیری از حوادث با نظر کارشناسان راهور پیشنهاد می‌شود و با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره از جمله TOPSIS، راه کارها رتبه‌بندی می‌شود. همچنین با روزآمد کردن پایگاه داده می‌توان به یک بازخورد مناسب از اقدامات انجام شده و میزان موفقیت آنها دست پیدا کرد. نتیجه اینکه از سیستم داده کاوی می‌توان به عنوان نظامی برای پشتیبانی از تصمیمات مدیران راهور استفاده کرد.

اهمیت و ضرورت پژوهش

با توجه به نسبت بالای تصادفات در سطح شهرهای کشور نیاز به تحلیل این تصادفات و بررسی دلایل آن برای کاهش این نسبت به شدت احساس می‌شود که این بررسی‌ها باید به صورت میدانی انجام پذیرد. از طرفی با توجه به پیشرفت سریع علم و مقدار داده‌های خام، نیاز به تحلیل‌های جدیدی که در زمره تحلیل‌های سیستم‌های خبره هستند بیشتر احساس می‌شود که فرایند داده کاوی به خوبی می‌تواند به این تحلیل‌ها بپردازد. همچنین پس از انجام این تحلیل‌ها و استخراج مهمترین عوامل دخیل در این

تصادفات، نیاز به تعیین راه کارهای برای کم کردن این تصادفات است که در این ارتباط تکنیک‌های تصمیم‌گیری می‌تواند به کارشناسان کمک کند تا بهترین تصمیم را اتخاذ کنند و بازخوردهای راه کارها را دوباره ملاحظه نمایند.

اهداف پژوهش

- ۱- رتبه‌بندی عوامل مؤثر بر تصادفات شهری اراک (با استفاده از الگوریتم‌های انتخاب ویژگی، داده‌کاوی و پیدا کردن وزن هر یک از عوامل)؛
- ۲- دسته‌بندی تصادفات در دو دسته فوت، جرح و پیدا کردن نقاط حادثه‌خیز در سطح شهر اراک (با استفاده از الگوریتم‌های دسته‌بندی داده‌کاوی)؛
- ۳- پیدا کردن راه‌های جلوگیری از تصادفات با استفاده از نظرات کارشناسان اداره راهور و پلیس راه استان مرکزی؛
- ۴- اولویت‌بندی راه‌های جلوگیری از تصادفات و ارائه بهترین راه برای جلوگیری از تصادفات در هر یک از نقاط حادثه‌خیز با استفاده از تکنیک‌های تصمیم‌گیری چندمعیاره؛
- ۵- ارائه سیستم داده‌کاوی برای روزآمدسازی نقاط حادثه‌خیز و بازخورد فعالیت‌های مرتبط؛
- ۶- استفاده از مدل تحلیل پوششی داده‌ها برای تعیین اولویت نقاط حادثه‌خیز با توجه به محدودیت‌های اجرایی و بودجه‌ای نیروی انتظامی در شهر اراک.

فرضیه‌های پژوهش

به صورت فرض یا فرض‌های ویژه:

- ❖ مدل ابتکاری از چه سطوح و مؤلفه‌هایی تشکیل شده است؟
- ❖ وزن هر یک از سطوح و مؤلفه‌ها به چه میزان است؟
- ❖ اولویت هر یک از سطوح و مؤلفه‌ها چگونه است؟
- ❖ ارائه راه کار مناسب به چه شکل صورت می‌گیرد؟

اهمیت و ضرورت استفاده از داده‌کاوی

دلایل استفاده از داده‌کاوی به شرح زیر است:

- ۱- میزان و حجم داده‌های در دسترس: در دهه‌های اخیر قیمت سخت‌افزار به‌ویژه فضای دیسک سخت، بسیار کاهش یافته است، به طوری که در اواخر سال ۲۰۰۴ هزینه هر گیگابایت دیسک سخت به‌طور تقریبی می‌توان برابر یک دلار امریکا در نظر گرفت. به دنبال آن بنگاه‌های اقتصادی حجم بسیاری از داده‌ها را در ابعاد مختلف گردآوری کردند. با انفجار داده‌ها، بنگاه‌های اقتصادی و سایر سازمان‌ها در پی تلاش برای کشف الگوهای پنهان در این داده‌ها به‌منظور هدایت راهبردهای خود برآمدند.
- ۲- اغلب الگوریتم‌های داده‌کاوی نیازمند حجم زیادی از داده‌ها هستند تا مدل‌هایی را ترتیب دهند که در دسته‌بندی، پیش‌بینی، تخمین و سایر کارکردهای داده‌کاوی مورد استفاده قرار گیرند.
- ۳- افزایش رقابت: رقابت به علت وجود بازارهای مدرن و فضاهای توزیع نظیر اینترنت و ارتباطات از راه دور به‌طور فزاینده‌ای در حال افزایش است. بنگاه‌های اقتصادی با رقابت‌های فضای وب مواجه‌اند و کلید موفقیت در تجارت، حفظ مشتریان کنونی و به‌دست آوردن مشتریان جدید با شناخت دقیق مشتریان و تمام افرادی است که در سازمان‌ها ذی‌نفع هستند. داده‌کاوی، فناوری‌هایی دارد که اجازه می‌دهد بنگاه‌های تجاری مؤلفه‌هایی را برای مواجه‌شدن با این زمینه‌ها تحلیل کنند.
- ۴- آماده بودن فناوری: داده‌کاوی پیشتر در حوزه‌های آکادمیک قرار داشت، اما در حال حاضر بسیاری از فناوری‌ها کامل شده‌اند و می‌توانند به‌طور فزاینده‌ای داده‌های پیچیده را تحلیل و مدیریت کنند. افزون‌بر این رابطه برنامه‌نویسی با کاربردهای داده‌کاوی (APIها)، اکنون استاندارد شده است که به توسعه‌دهندگان این امکان را می‌دهند که کاربردهای داده‌کاوی بهتری بسازند.
- ۵- علاقه به مدیریت ارتباط با مشتریان: در طیف وسیعی از صنایع، سازمان‌ها به این بینش رسیده‌اند که مشتریان برای سازمان حیاتی هستند و اطلاعات مشتریان یکی از دارایی‌های اساسی سازمان‌هاست. اطلاعات سازمان‌ها از مشتریان خود، نه تنها برای خودشان بلکه برای دیگران هم ارزشمند است. یک شرکت کارت اعتباری چیزهایی می‌داند که شرکت‌های خطوط هوایی دوست دارند بدانند چه کسانی بلیط‌های پرواز متعددی می‌خرند. گوگل می‌داند افراد در وب دنبال چه چیزی هستند و از این شناخت با فروش پیوندهایی با پشتیبانی مالی بهره می‌برد و در واقع هر شرکتی که داده‌های باارزش‌تری جمع‌آوری کند در موقعیت یک واسطه اطلاعات قرار می‌گیرد (مشکانی و ناظمی، ۱۳۸۸).

رویکردهای مختلف فرایند داده‌کاوی

الگوهای مختلفی به‌عنوان رویکرد یا روش‌شناسی برای طول عمر پروژه‌های داده‌کاوی در منابع مطالعاتی معتبر مختلف دنیا یا شرکت‌های ارائه خدمات در زمینه‌های داده‌کاوی مستندسازی شده‌اند و مورد طراحی قرار گرفته‌اند، برای نمونه می‌توان به رویکرد شرکت SPSS با به‌کارگیری پنج مرحله

ارزیابی^۳، دسترسی^۴، تحلیل^۵، اقدام^۶ و خودکارسازی^۷ اشاره کرد. شرکت SAS نیز رویکرد ۵ مرحله‌ای نمونه‌برداری^۸، اکتشاف^۹، تغییر^{۱۰}، مدل‌سازی^{۱۱} و ارزیابی^{۱۲} است را ارائه کرده است. یک مجموعه از فروشندگان و کاربران شرکت‌های دایملر کرایسلر، SPSS و NCR نیز رویکرد دیگری را ارائه کرده‌اند که CRISP-DM^{۱۳} نامیده شده است.

فرایند داده‌کاوی

بدون تلاش برای پوشش همه روش‌های ممکن با دیدگاه‌های متفاوت درباره داده‌کاوی به عنوان یک نظم و آموزه، تعریفی نسبتاً جامع از داده‌کاوی ارائه می‌شود.

«داده‌کاوی به فرایند جست‌وجو و کشف الگوهای گوناگون، مختصرسازی‌ها و اخذ مجموعه‌ای از مقادیر از داده‌های معلوم اطلاق می‌شود». در اینجا کلمه و اصطلاح «فرایند» بسیار مهم است. حتی در بعضی محیط‌های حرفه‌ای این نظر وجود دارد که داده‌کاوی شامل انتخاب و به‌کارگیری ابزار مبتنی بر رایانه برای حل مسائل فعلی و به دست آوردن راه‌حل به طور خودکار است. این برداشت در حقیقت یک تصور اشتباه بر پایه آرمان‌گرایی تصنعی جهان است. البته دلایل زیادی برای این برداشت اشتباه وجود دارد. یکی از دلایل این است که داده‌کاوی فقط مجموعه‌ای از ابزارها نیست بلکه هر کدام کاملاً از یکدیگر متفاوت بوده و به دنبال حل مسائل هستند. دلیل دوم در انطباق مسئله با تکنیک نهفته است. خیلی به‌ندرت این حالت وجود دارد که یک پرسش تحقیقی با دقت کافی، کاربرد ساده و منحصر به فرد، یک روش و تکنیک را بیان کند. در حقیقت در عمل چه روی می‌دهد که داده‌کاوی به یک فرایند تکرارپذیر تبدیل می‌شود. به‌طور دقیق‌تر ما داده‌ها را مطالعه می‌کنیم، آنها را با استفاده از تکنیک‌های تحلیلی بررسی و بالاخره به آنها از دیدگاه‌ها و زوایای دیگر نگریسته می‌شود. شاید آنها را اصلاح کنیم و در این راستا

3. Assess
4. Access
5. Analyse
6. Act
7. Automate
8. Sample
9. Explore
10. Modify
11. Model
12. Assess
13. Cross- Industry Standard Process for Data Mining

لازم باشد کار را از ابتدا شروع و ابزار تحلیل داده دیگری را اعمال کنیم که حاصل همه اینها، شاید رسیدن به یک نتیجه جدید یا بهتر باشد. این فرایند می‌تواند بارها و بارها تکرار شود. هر تکنیک برای بررسی جنبه‌های نسبتاً متفاوت داده‌ها استفاده می‌شود. مطرح کردن یک پرسش نسبتاً متفاوت درباره داده‌ها به عنوان سفری اکتشافی توصیف می‌شود که داده‌کاوی جدید را هیجان‌انگیز می‌سازد. با وجود این، داده‌کاوی یک کاربرد تصادفی از روش‌های آماری، یادگیری خودرو و دیگر روش‌ها و ابزارها نیست. همچنین داده‌کاوی یک قدم‌زنی تصادفی در فضای روش‌های تحلیلی نمی‌باشد، بلکه این مبحث و پدیده جدید، یک فرایند تصمیم‌گیری با دقت برنامه‌ریزی شده می‌باشد که بسیار سودمند، امیدوارکننده و مناسب به نظر می‌رسد.

نرم‌افزارهای داده کاوی

از آنجا که داده‌کاوی با بانک‌های اطلاعاتی بزرگ ارتباط دارد، به‌گونه‌ای ایده‌آل با فناوری خدمت‌گیر- خدمت‌گر به کار می‌رود. کاربردهای عمومی داده‌کاوی بیشتر شامل تقسیم‌کردن داده‌ها در خوشه‌های مقتضی، شناسه‌گذاری‌های مناسب، کاوش برای الگوها و طراحی کردن با استفاده از فنون آماری و الگوریتم‌های ژنتیکی است. بسیاری از بسته‌های نرم‌افزاری در دسترس، دارای جنبه‌های داده‌کاوی با درجات متفاوتی از جامعیت هستند. برای نمونه بسته‌های نرم‌افزاری که صرفاً برای کاربردهای OLAP در دسترس هستند عبارت‌اند از: Oracle OLAP، DB2 OLAP Server، CleverPath OLAP. نرم‌افزارهای آماری عمومی نظیر Insightful، Enterprise Weka، Text Mining Software، Miner3 Poly Analyst 4.6 و Data Mining software مفید هستند.

همچنین نرم‌افزارهای تخصیص گران‌قیمت دیگری نظیر Enterprise Miner به‌وسیله شرکت Intelligence Miner، SAS، به‌وسیله IBM، CLEMENTINE، به‌وسیله شرکت SPSS به بازار ارائه شد. بسیاری از این نرم‌افزارها قابل دریافت از وب‌گاه‌های شرکت‌های سازنده‌شان هستند، بعضی از ابزارها مربوط به رشته‌های خاص بوده و از این رو پژوهشگران با شناخت از ویژگی ابزارها، قابل دسترس بودن و دامنه تحقیق خود اقدام به استفاده از آنها براساس نیازهایشان می‌کنند (مشکانی و ناظمی، ۱۳۸۸).

محصولات تخصصی تقریباً هر کاربرد تجاری سودآور ممکن را پوشش می‌دهند. یک منبع خوب برای دیدن محصولات فعلی سایت WWW.KDNUUGGESTS.COM است. منبع اطلاع دست‌گاه UCI در WWW.ici.edu/~mlearn/MLOther.html یک منبع خوب مجموعه داده‌های داده‌کاوی‌هاست.^{۱۴} این

14. C. J. Merz, P. M. Murphy. UCI Repository of Machine Learning Databases

وب‌گاه شامل منابعی برای درگاه‌های خوب داده‌کاوی دیگر نیز است. شرکت‌های IBM، SAS، مایکروسافت، Oracle و Brio Technology و ... عرضه‌کننده محصولات مربوط به داده‌کاوی هستند. برنامه کاوشگر هوشمند IBM مجموعه الگوریتم‌هایی برای داده‌کاوی به‌منظور تشخیص ارتباطات پنهانی، روندها و الگوها در اختیار دارد. سیستم SAS برای تحویل اطلاعات، نظام‌های اطلاعات اجرایی، نرم‌افزارهای آماری برای تحلیل داده‌ها و نرم‌افزارهای شبکه عصبی را با یکدیگر یکپارچه می‌سازد (دیوید ال السون و همکاران، ۱۳۹۶).

روش تحقیق

این پژوهش براساس هدف از نوع کاربردی و از منظر روش، تحقیقی توصیفی از نوع پیمایشی می‌باشد؛ توصیفی است از این جهت که تصویری از وضعیت موجود را ارائه می‌دهد. از طرفی برای دستیابی و توصیف ویژگی‌های جامعه آماری نیز از روش تحقیق پیمایشی بهره‌گیری شده است. داده‌های مورد استفاده در این پژوهش در مورد اطلاعات تصادفات از پایگاه داده ثبت تصادفات پلیس راه و راهنمایی و رانندگی استان استخراج شده است و اطلاعات مورد استفاده برای نقاط حادثه‌خیز در مورد جاده‌های استان مرکزی و وزن‌های معیارها و خودمعیارها در ارائه راه کار براساس نظرات کارشناسان راهور استان مرکزی و کارشناسان و فرمانده پلیس راه استان مرکزی و معاونت ترافیک شهرداری اراک از طریق پرسشنامه براساس اطلاعات پیمایشی است.

قلمرو مکانی و زمانی پژوهش

در تحقیق حاضر، قلمرو مکانی پژوهش، تصادفات فوتی در سطح شهر اراک و قلمرو زمانی آن، تصادفات سال ۱۳۹۷ بوده است.

حجم نمونه

در این پژوهش از گرفتن نمونه برای پایایی بیشتر تکنیک‌های داده‌کاوی و تصمیم‌گیری اجتناب شده است و تمامی تصادفات فوتی در شهر اراک تحلیل شده است.

روش گردآوری داده‌ها

اطلاعات این پژوهش مربوط به تصادفات شهر اراک (ثبت‌شده در سیستم ثبت اطلاعات تصادفات راهور) بوده است. در مورد نقاط حادثه‌خیز در شهر اراک از پرسشنامه استفاده شده است.

ابزار گردآوری اطلاعات

ابزار گردآوری اطلاعات، جداول و نیز پرسشنامه برای شناسایی نقاط حادثه خیز در شهر اراک است. همچنین برای امتیازدهی به معیارها از جدول متغیرهای فازی زبانی استفاده شده است.

اطلاعات توصیفی تصادفات شهری اراک

اطلاعات توصیفی تصادفات شهری اراک در سال ۱۳۹۷ در جداول ۱ تا ۷ نشان داده شده است.

جدول ۱. نوع وسیله مقصر در تصادفات

موتورسیکلت	سواری	وانت بار	مینی بوس	اتوبوس	کامیون	سایر
۲۰۰۱	۴۲۱۲	۵۴۲	۲۳۲	۱۴۶	۱۷۷	۸۶

جدول ۲. علت تصادف

عبور از محل ممنوع	عدم رعایت فاصله طولی	عبور از چراغ قرمز	عدم رعایت حق تقدم	سرعت غیرمطمئن	اتحراف به چپ یا راست	شتاب زدگی در راندگی	حرکت با دنده عقب	گردش به طرز غلط	سایر حالات
۵۴۸	۷۵۱	۷۳۳	۲۰۴۵	۱۴۵۰	۸۸۲	۲۸۸	۱۷۹	۱۸۳	۳۴۷

جدول ۳. محل تصادف

خیابان	میدان	تقاطع	بزرگراه	کمربندی	سایر
۳۲۷۰	۱۸۷۴	۱۲۶۰	۱۶۹	۴۱۱	۴۱۲

جدول ۴. استفاده از کمربند ایمنی

استفاده از کمربند ایمنی	
بلی	خیر
۲۷۷۰	۲۵۸۲

جدول ۵. استفاده از کلاه ایمنی

استفاده از کلاه ایمنی	
بلی	خیر
۱۱۰۹	۹۳۵

جدول ۶. زمان تصادفات شهری

زمان تصادف	
شب	روز
۲۲۱۹	۵۱۷۷

جدول ۷. وضعیت هوا

وضعیت هوا		
بارندگی	ابری	آفتابی
۲۳۲	۱۰۶۱	۵۲۶۹

رتبه‌بندی عوامل مؤثر بر تصادفات درون شهری اراک

در این مرحله با استفاده از پایگاه داده پلیس راهور و الگوریتم‌های داده‌کاوی، وزن عوامل مؤثر بر تصادفات محاسبه می‌شود که شامل انجام مراحل زیر است. پایگاه داده تصادفات به شکل زیر تشکیل می‌شود:

– رکوردها:

در پایگاه داده پلیس راهور، ۹۰۰ رکورد موجود بود که از آن تعداد رکوردهای حاوی اطلاعات هر یک از تصادفات منجر به فوت در سال ۱۳۹۷ شامل ۸۰ تصادف در سطح شهر اراک است. گفتنی است اطلاعات سایر تصادفات که شامل تصادفات جرحی و خسارتی می‌شوند در پژوهش حاضر استفاده نشده است.

– ویژگی‌ها:

با توجه به داده‌هایی در اختیار پژوهشگر، ۸ ویژگی مورد استفاده قرار گرفت که مشخصات و مقادیری آنها در این پایگاه داده، عبارت‌اند از:

۱. ماه تصادف (بین ۱ تا ۱۲ است).
۲. ساعت تصادف (به صورت طبقه‌بندی مورد استفاده قرار گرفته است و بر این اساس تصادفات به دسته‌های ۲ ساعته تقسیم‌بندی شده‌اند).

۳. نوع وسیله مقصر در تصادفات (شامل سواری، اتوبوس و مینی‌بوس، موتور، وانت، خودروهای سنگین و غیره است).
۴. سیستم خودرو (شامل انواع خودروهای سواری است).
۵. سن راننده (عددی است).
۶. محل فوت (شامل صحنه حادثه و بیمارستان است).
۷. جنسیت (شامل مرد و زن است).
۸. محل تصادف (شامل قرار گرفتن تصادف در حوزه استحفاظی هر یک از پاسگاه‌های راهنمایی و رانندگی ۴گانه شهرستان اراک شامل پاسگاه‌های سرداران، باغ ملی، میدان انقلاب و پارک آزادی است).

- برچسب (کلاس):

در این پژوهش به علت آنکه پایگاه مدنظر فقط پایگاه داده تصادفات فوتی در شهرستان اراک بود، از علت تامه تصادف به عنوان کلاس پایگاه داده استفاده شد که شامل موارد زیر است:

۱. توجه نکردن به جلو
۲. مهارت نداشتن در رانندگی
۳. انحراف ناگهانی در مسیر یا انحراف به چپ
۴. سرعت غیرمطمئن
۵. ناتوانی در کنترل وسیله نقلیه
۶. رعایت نکردن حق تقدم
۷. سایر موارد

اجرای طبقه‌بندی:

پس از تشکیل پایگاه داده در این مرحله برای اینکه پی ببریم آیا ویژگی‌های انتخاب‌شده مناسب هستند یا خیر، با استفاده از نرم‌افزار weka (وکا) از الگوریتم‌های طبقه‌بندی به تحلیل پایگاه داده می‌پردازیم که نتایج آن در جدول ۸ نشان داده شده است.

با توجه به میزان دقت به‌دست‌آمده از الگوریتم‌های مختلف طبقه‌بندی و به‌دست‌آوردن دقت ۵۰ درصد توسط تعدادی از الگوریتم‌ها و با توجه به تعداد بسیار کم رکوردها می‌توان نتیجه گرفت که دقت به‌دست‌آمده دقت خوبی است و ویژگی‌ها به‌درستی انتخاب شده‌اند.

جدول ۸. دقت هر یک از الگوریتم‌های طبقه‌بندی بر پایگاه داده راهور

الگوریتم	دقت طبقه‌بندی
Logistic	۵۰ درصد
Random forest	۷۰ درصد
SMO	۴۶ درصد
RBF	۵۰ درصد
J48	۷۲ درصد

تعیین بهترین روش برای رتبه‌بندی ویژگی‌ها:

در نرم‌افزار وکا از آنجایی که روش‌های متفاوتی برای تعیین وزن (اهمیت) ویژگی‌ها وجود دارد (مثل One R Attribute، Info gain و Gain Ratio) و در برخی از موارد رتبه‌بندی این عوامل با استفاده از روش‌های مختلف متفاوت است، بنابراین برای تعیین اینکه کدام الگوریتم نتیجه بهتری ارائه می‌دهد کافیسست که در هر رتبه‌بندی عواملی که دارای کمترین میزان اهمیت هستند را شناسایی و سپس آنها را از پایگاه داده حذف و دوباره نسبت به رتبه‌بندی پایگاه داده اقدام کرد. هر الگوریتم رتبه‌بندی ویژگی که با حذف کم‌اهمیت‌ترین عوامل، پایگاه داده با دقت بالاتری در طبقه‌بندی بگیرد، نتایج حاصل از آن دقیق‌تر است. در این پایگاه داده الگوریتم One R Attribute نتایج دقیق‌تری داشت.

رتبه‌بندی ویژگی‌ها:

در این مرحله با استفاده از الگوریتم‌های رتبه‌بندی ویژگی به رتبه‌بندی ویژگی‌های پیش‌گفته خواهیم پرداخت. از طرفی با توجه به تعداد پایین داده‌ها رتبه‌بندی‌ها توسط الگوریتم‌های مختلف رتبه‌بندی ویژگی، نتایج متفاوت اما نزدیک به هم به دست می‌آید. بر این اساس یکی از الگوریتم‌ها (One R Attribute) به عنوان پایه‌ای برای رتبه‌بندی عوامل استفاده شد و از ۲ الگوریتم دیگر رتبه‌بندی ویژگی (Latent, Reilf Attribute) به عنوان رتبه‌بندی کمکی استفاده می‌شود و در صورتی که ویژگی‌هایی در رتبه‌بندی پایه دارای رتبه مساوی باشند، با توجه به رتبه‌بندی‌های کمکی نسبت به رتبه‌بندی نهایی ویژگی‌ها اقدام می‌شود. بر این اساس رتبه‌بندی ویژگی‌های تصادفات به همراه اهمیت آنها، با استفاده از الگوریتم رتبه‌بندی ویژگی One R Attribute در جدول ۹ نشان داده شده است.

جدول ۹. رتبه‌بندی و تعیین اهمیت عوامل با استفاده از الگوریتم One R Attribute

عوامل	ماه تصادف	ساعت تصادف	نوع وسیله	سیستم خودرو	سن راننده	محل فوت	جنسیت	محل تصادف
اهمیت عوامل	۶۶	۶۶	۷۴	۶۶	۶۶	۶۶	۶۶	۶۶
رتبه عوامل	۲	۲	۱	۲	۲	۲	۲	۲

در جدول ۹ رتبه تعدادی از عوامل مشترک است، بنابراین برای تعیین رتبه‌بندی نهایی عوامل از دیگر الگوریتم‌های رتبه‌بندی ویژگی که پیشتر آمد استفاده می‌شود. رتبه نهایی عوامل در جدول ۱۰ نشان داده شده است.

جدول ۱۰. رتبه‌بندی نهایی عوامل

عوامل	ماه تصادف	ساعت تصادف	نوع خودرو	سیستم خودرو	سن راننده	محل فوت	جنسیت	محل تصادف
رتبه نهایی	۸	۷	۶	۵	۴	۳	۲	۱

همان‌گونه که در جدول ۱۰ مشاهده می‌شود محل تصادف به عنوان مهمترین عامل بر علت تصادفات شهرستان اراک است و پس از آن جنسیت در رده دوم قرار دارد.

تعیین اولویت نقاط حادثه‌ساز در شهر اراک

پس از تعیین اهمیت عوامل مؤثر بر تصادفات در شهر اراک در این مرحله باید نسبت به تعیین نقاط حادثه‌خیز اقدام کرد. از آنجایی که در سطح استان نمی‌توان نسبت به برطرف کردن برخی از عوامل تصادفات اقدام کرد (از جمله اینکه نمی‌توان برای نمونه نسبت به توقف خط تولید پراید اقدام کرد)، در این مرحله با توجه به خروجی‌های مرحله ۱ که نقاط حادثه‌خیز در شهر اراک نیز هست و همچنین با توجه به نظرات کارشناسان راهور این نقاط استخراج می‌شوند که مراحل انجام این کار به شرح زیر است:

۱. ابتدا پرسشنامه‌ای طراحی شد و در آن از کارشناسان راهور خواسته شد که با توجه به سابقه تجربی که دارند نقاط حادثه‌خیز شهر اراک را تعیین کنند و از کارشناسان خواسته شد که اهمیت اصلاح آن نقطه را نیز با متغیرهای زبانی فازی تعیین نمایند و همچنین بیان کنند که در هر یک از آن نقاط چه حوادثی (منجر به فوت، جرح و خسارت) روی داده است.

۲. پرسشنامه در مناطق چهارگانه راهنمایی و رانندگی اراک پخش شد و توسط کارشناسان محترم راهور تکمیل و ۵۵ نقطه حادثه‌خیز شناسایی شد که محل این نقاط در جدول ۱۱ نشان داده شده است.

جدول ۱۱. نقاط حادثه‌خیز اولیه با استفاده از پرسشنامه‌ها

شماره نقطه	محل نقطه حادثه‌خیز
۱	خ سیدالشهدا سه‌راه الزهرا
۲	سه‌راه شهرک قائم
۳	خ موزه- راهزان
۴	مهارت-راهزان
۵	میدان شاهد
۶	پل انتهایی کشتارگاه
۷	پل ابوذر
۸	ورودی داوران- شهدا از کمربندی
۹	شهرک حمید- استخر اتحاد
۱۰	آیت‌الله غفاری- راهزان
۱۱	چهرم- مهارت
۱۲	چهرم- غفاری
۱۳	میدان مقاومت
۱۴	ورودی کمربندی به امام حسین (ع)
۱۵	سیدالشهدا- راهزان
۱۶	ورودی شهدا از کمربندی
۱۷	ورودی بهشت زهرا
۱۸	تقاطع شهدا
۱۹	م ارگ به خ حصار
۲۰	م فراهان به سه‌راه زندان
۲۱	خ جنت به خ شریعتی
۲۲	انتهای قائم‌مقام شهرک جوادیه
۲۳	تقاطع فردوسی بعثت
۲۴	پل شهر صنعتی

شماره نقطه	محل نقطه حادثه خیز
۲۵	پل برق
۲۶	تقاطع الغدير- جوادیه
۲۷	خ امام پل زایشگاه
۲۸	سردشت
۲۹	پل فرنگی
۳۰	م ورزش
۳۱	م شریعی
۳۲	بریدگی های چهارم
۳۳	بریدگی لعلی و قائم مقام
۳۴	تقاطع اشکانی قائم مقام
۳۵	بریدگی مقابل پمپ بنزین کشتارگاه
۳۶	تقاطع غفاری الزهرا
۳۷	تقاطع الزهرا- لعلی
۳۸	طول خ مشهد تا میدان فراهان
۳۹	کمربندی جنوبی پیچ بالای شهرک ولیعصر
۴۰	کمربندی شمالی روبروی پیام نور
۴۱	م محیط زیست
۴۲	کمربندی شمالی پیچ طرمزد
۴۳	بلوار کربلا روبروی شهرک نبی
۴۴	خ سنجانی- تقاطع ناگهی
۴۵	بلوار سردشت بوتان گاز
۴۶	خ دانشگاه کوچه جمشیدی
۴۷	۲۰ متری فوتبال بنی هاشمی (کمربندی)
۴۸	شهرک گردو- م فلزی- سازمان آب
۴۹	تقاطع جوادیه- حمید
۵۰	م فرمانداری خ شریعی
۵۱	خ رجایی ۱۲ متری
۵۲	م انقلاب
۵۳	محور خمین روبروی دانشگاه آزاد

۳. در این مرحله نسبت به اولویت‌بندی اصلاح نقاط یادشده اقدام شد که این کار با تشکیل مدل تحلیل پوششی داده‌ها انجام می‌شود که شامل انجام مراحل زیر است:
- الف. ابتدا مدلی تشکلی می‌شود که واحدها، ورودی‌ها و خروجی‌های مدل عبارت‌اند از:
- واحدها: هر یک از نقاط یک واحد تصمیم‌گیری برای الگو هستند که نسبت به اولویت‌بندی اصلاح آنها اقدام می‌شود.
- ورودی: ورودی هر یک از نقاط، ثابت است (در این تحقیق عدد ثابت، عدد ۱ در نظر گرفته شده است)
- خروجی‌ها: برای هر واحد تصمیم‌گیری ۳ خروجی تعریف شده است که عبارت‌اند از:
۱. مرحله‌ای که کارشناسان، نقطه حادثه‌خیز موردنظر را در پرسشنامه‌ها ذکر کرده‌اند.
 ۲. درجه اهمیت فازی که کارشناسان در مورد اصلاح نقطه موردنظر در پرسشنامه ذکر کرده‌اند، در جدول ۱۲ نشان داده شده است.

جدول ۱۲. متغیرهای زبانی فازی

امتیاز	متغیر زبانی
۵	خیلی زیاد
۴	زیاد
۳	متوسط
۲	کم
۱	خیلی کم

۳. مجموع اهمیت اصلاح نقطه از نظر تصادفات انجام شده در مورد هر یک از نقاط بر اساس جدول ۱۳.

جدول ۱۳. امتیازات حوادث روی‌داده

امتیاز	حوادث روی‌داده
۶	فوتی
۳	جرحی
۱	خسارتی

برای نمونه در مورد نقطه شماره ۱ (خ سیدالشهدا سه راه الزهرا) که یک واحد تصمیم گیری است ورودی و خروجی ها به شرح زیر است:

ورودی: عدد ۱ است (که در مورد تمام نقاط ثابت است)

خروجی های سه گانه شامل اعداد زیر است:

خروجی اول: خروجی این نقطه عدد ۲ است، زیرا در ۲ پرسشنامه کارشناسان راهور این نقطه را به عنوان نقطه حادثه خیز ذکر کرده اند.

خروجی دوم: خروجی این نقطه عدد ۹ است، زیرا کارشناس اهمیت اصلاح نقطه را خیلی زیاد (۵) و دیگری زیاد (۴) ذکر کرده است.

خروجی سوم: خروجی این نقطه عدد ۱۴ است، زیرا یک کارشناس بیان کرده که در این نقطه فقط حوادث منجر به جرح و خسارتی روی می دهد ($3+1=4$) و دیگری بیان کرده که در این نقطه حوادث منجر به فوت، جرح و خسارتی داریم ($6+3+1=10$) و این نقطه از خروجی سوم امتیاز ۱۴ ($4+10$) دارد. و در مورد بقیه نقاط نیز به همین ترتیب خروجی ها به دست می آید که جدول ۱۴، امتیازات هر کدام از نقاط در هر یک از خروجی ها نشان داده شده است.

جدول ۱۴. امتیازات نقاط حادثه خیز

شماره	محل نقطه حادثه خیز	امتیاز خروجی اول	امتیاز خروجی دوم	امتیاز خروجی سوم
۱	خ سیدالشهدا سه راه الزهرا	۲	۹	۱۴
۲	سه راه شهرک قائم	۱	۵	۱۰
۳	خ موزه- راهزان	۳	۱۴	۱۲
۴	مهارت- راهزان	۲	۸	۸
۵	میدان شاهد	۱	۳	۴
۶	پل انتهای کشتارگاه	۲	۹	۸
۷	پل ابوذر	۱	۴	۴
۸	ورودی داوران- شهدا از کمربندی	۲	۹	۸
۹	شهرک حمید- استخر اتحاد	۲	۸	۸
۱۰	آیت الله غفاری- راهزان	۱	۳	۴
۱۱	چهرم- مهارت	۱	۱۰	۱۰
۱۲	چهرم- غفاری	۱	۳	۴

شماره	محل نقطه حادثه خیز	امتیاز خروجی اول	امتیاز خروجی دوم	امتیاز خروجی سوم
۱۳	میدان مقاومت	۱	۴	۴
۱۴	ورودی کمربندی به امام حسین	۱	۴	۱۰
۱۵	سیدالشهدا- راهزنان	۲	۹	۱۴
۱۶	ورودی شهدا از کمربندی	۱	۴	۴
۱۷	ورودی بهشت زهرا	۲	۸	۱۴
۱۸	تقاطع شهدا	۱	۳	۴
۱۹	م ارگ به خ حصار	۱	۳	۴
۲۰	م فراهان به سهراب زندان	۳	۱۵	۳۰
۲۱	خ جنت به خ شریعتی	۲	۹	۸
۲۲	انتهای قائم مقام شهرک جوادیه	۱	۵	۴
۲۳	تقاطع فردوسی بعثت	۱	۵	۱۰
۲۴	پل شهر صنعتی	۱	۴	۴
۲۵	پل برق	۲	۷	۱۱
۲۶	تقاطع النذیر- جوادیه	۲	۶	۲
۲۷	خ امام پل زایشگاه	۳	۱۲	۱۵
۲۸	سردشت	۱	۳	۱
۲۹	پل فرنگی	۲	۶	۲
۳۰	م ورزش	۲	۹	۵
۳۱	م شریعتی	۲	۸	۵
۳۲	بریدگی های چهارم	۱	۴	۴
۳۳	بریدگی لعلی و قائم مقام	۱	۴	۴
۳۴	تقاطع اشکانی قائم مقام	۱	۴	۴
۳۵	بریدگی مقابل پمپ بنزین کشتارگاه	۱	۴	۴
۳۶	تقاطع غفاری الزهرا	۲	۱۰	۸
۳۷	تقاطع الزهرا- لعلی	۱	۵	۴
۳۸	طول خ مشهد تا میدان فراهان	۱	۵	۴
۳۹	کمربندی جنوبی پیچ بالای شهرک ولیعصر	۲	۱۰	۱۴
۴۰	کمربندی شمالی روبروی پیام نور	۱	۵	۱۰
۴۱	م محیط زیست	۱	۴	۱۰

شماره	محل نقطه حادثه خیز	امتیاز خروجی اول	امتیاز خروجی دوم	امتیاز خروجی سوم
۴۲	کمربندی شمالی پیچ طرمزد	۲	۸	۱۴
۴۳	بلوار کربلا روبروی شهرک نبی	۲	۹	۱۴
۴۴	خ سنجانی- تقاطع ناگهی	۱	۴	۴
۴۵	بلوار سرودشت بوتان گاز	۱	۳	۱۰
۴۶	خ دانشگاه کوچه جمشیدی	۱	۵	۴
۴۷	۲۰متری فوتبال بنی هاشمی (کمربندی)	۱	۴	۱۰
۴۸	شهرک گردو- م فلزی- سازمان آب	۱	۴	۴
۴۹	تقاطع جوادیه- حمید	۱	۵	۴
۵۰	م فرمانداری خ شریعتی	۱	۳	۱
۵۱	خ رجایی ۱۲ متری	۱	۴	۱
۵۲	م انقلاب	۱	۴	۱۰
۵۳	محور خمین روبروی دانشگاه آزاد	۱	۵	۱۰

بدیهی است که هر چه نقطه‌ای امتیاز بالاتری به دست آورده باشد حادثه‌خیزی آن نقطه بیشتر است.

اولویت‌بندی نقاط براساس پرسشنامه‌ها

در این مرحله با استفاده از نرم‌افزار DEAP نسبت به اولویت‌بندی نقاط حادثه‌خیز براساس پرسشنامه‌ها اقدام می‌شود که نتیجه این رتبه‌بندی در جدول ۱۵ نشان داده شده است. این اولویت‌بندی براساس کارایی که هر یک از نقاط دارند انجام می‌شود. گفتنی است هر نقطه‌ای که کارایی بالاتری دارد، رتبه بالاتری دارد.

جدول ۱۵. رتبه‌بندی نهایی با استفاده از پرسشنامه‌ها

شماره نقطه	محل نقطه حادثه‌خیز	کارایی نقطه	رتبه نقطه
۱	خ سیدالشهدا سهره الزهرا	۰/۵۹۷	۵
۲	سهره شهرک قائم	۰/۳۶۱	۲۱
۳	خ موزه- راهزان	۰/۸۱۰	۳
۴	مهارت- راهزان	۰/۵۲۷	۱۵
۵	میدان شاهد	۰/۲۵۳	۴۳
۶	پل انتهای کشتارگاه	۰/۵۲۷	۱۲
۷	پل ابوذر	۰/۲۶۳	۳۶

شماره نقطه	محل نقطه حادثه خیز	کارایی نقطه	رتبه نقطه
۸	ورودی داوران - شهدا از کمربندی	۰/۵۳۷	۱۲
۹	شهرک حمید- استخر اتحاد	۰/۵۳۷	۱۵
۱۰	آیت‌الله غفاری - راهزان	۰/۲۵۳	۴۳
۱۱	چهرم - مهارت	۰/۳۰۳	۳۱
۱۲	چهرم - غفاری	۰/۲۵۳	۴۳
۱۳	میدان مقاومت	۰/۲۶۱	۳۶
۱۴	ورودی کمربندی به امام حسین	۰/۳۲۳	۲۴
۱۵	سیدالشهدا - راهزان	۰/۵۹۷	۵
۱۶	ورودی شهدا از کمربندی	۰/۲۶۳	۳۶
۱۷	ورودی بهشت زهرا	۰/۵۸۷	۸
۱۸	تقاطع شهدا	۰/۲۵۳	۴۳
۱۹	م ارگ به خ حصار	۰/۲۵۳	۴۳
۲۰	م فراهان به سهراب زندان	۱	۱
۲۱	خ جنت به خ شریعتی	۰/۵۳۷	۱۲
۲۲	انتهای قائم‌مقام شهرک جوادیه	۰/۳۲۳	۲۲
۲۳	تقاطع فردوسی بعثت	۰/۲۶۳	۳۶
۲۴	پل شهر صنعتی	۰/۵۴۷	۱۰
۲۵	پل برق	۰/۴۴۷	۱۹
۲۶	تقاطع الغدیر - جوادیه	۰/۲۲۳	۳۲
۲۷	خ امام پل زایشگاه	۰/۸۲۰	۲
۲۸	سردشت	۰/۲۲۳	۵۱
۲۹	پل فرنگی	۰/۴۴۷	۱۹
۳۰	م ورزش	۰/۵۰۶	۱۷
۳۱	م شریعتی	۰/۴۷۹	۱۸
۳۲	بریدگی‌های چهرم	۰/۲۶۳	۳۶
۳۳	بریدگی لعلی و قائم‌مقام	۰/۲۶۳	۳۶
۳۴	تقاطع اشکانی قائم‌مقام	۰/۲۶۳	۳۶
۳۵	بریدگی مقابل پمپ بنزین کشتارگاه	۰/۲۶۳	۳۶
۳۶	تقاطع غفاری الزهرا	۰/۵۴۷	۱۰

شماره نقطه	محل نقطه حادثه خیز	کارایی نقطه	رتبه نقطه
۳۷	تقاطع الزهرا- لعلی	۰/۲۷۳	۳۲
۳۸	طول خ مشهد تا میدان فراهان	۰/۲۷۳	۳۲
۳۹	کمربندی جنوبی پیچ بالای شهرک ولیعصر	۰/۶۰۷	۴
۴۰	کمربندی شمالی روبروی پیام نور	۰/۳۳۳	۲۲
۴۱	م محیط زیست	۰/۳۲۳	۲۴
۴۲	کمربندی شمالی پیچ طرمزد	۰/۵۸۷	۸
۴۳	بلوار کربلا روبروی شهرک نبی	۰/۵۹۷	۵
۴۴	خ سنجانی- تقاطع ناگهی	۰/۲۶۳	۳۶
۴۵	بلوار سرودشت بوتان گاز	۰/۳۱۳	۲۹
۴۶	خ دانشگاه کوچه جمشیدی	۰/۲۷۳	۳۲
۴۷	۲۰متری فوتبال بنی هاشمی (کمربندی)	۰/۳۲۳	۲۴
۴۸	شهرک گردو- م فلزی- سازمان آب	۰/۲۶۳	۳۶
۴۹	تقاطع جوادیه- حمید	۰/۱۳۱	۲۹
۵۰	م فرمانداری خ شریعتی	۰/۲۲۳	۵۱
۵۱	خ رجایی ۱۲ متری	۰/۳۲۳	۵۱
۵۲	م انقلاب	۰/۳۲۳	۳۴
۵۳	محور خمین روبروی دانشگاه آزاد	۰/۳۳۳	۲۴

رتبه‌بندی نهایی نقاط و هدف‌گذاری اصلاح نقاط در سال ۱۳۹۷

در این مرحله پس از اولویت‌بندی نقاط براساس پرسشنامه‌ها، نسبت به اولویت‌بندی نهایی نقاط اقدام شد که شامل مراحل زیر است:

۱. از ۵۳ نقطه به‌دست‌آمده در مرحله ۱ تعدادی از این نقاط در حال اصلاح هستند که از تحلیل و هدف‌گذاری حذف شدند و در نشست با حضور فرمانده عملیات راهور استان مرکزی و رئیس راهور شهرستان اراک، نقاط تعیین شده‌اند که عبارت‌اند از: نقاط شماره ۱۱، ۱۶، ۱۷، ۲۰، ۲۱، ۲۵، ۲۷ و ۳۰.
۲. در این مرحله نیز با نظر فرمانده عملیات راهور استان مرکزی و رئیس راهور اراک نقاطی که در پرسشنامه‌ها لحاظ نشده بود به نقاط حادثه‌خیز اضافه شد که این نقاط عبارت‌اند از:

۱. محور اراک قم بریدگی یگان ویژه

۲. میدان امام حسن (تره‌بار جدید)

۳. بریدگی بلوار دانشگاه با خانه‌سازی اراک

۴. اواسط خ جنت پارک جنت

۳. در این مرحله با برگزاری جلسه طوفان فکری با حضور رئیس پلیس راهور اراک و سرکلانترهای بخش‌های ۴گانه راهور اراک نسبت به رتبه‌بندی نهایی نقاط براساس رتبه‌بندی پرسشنامه‌ها اقدام شد. سپس براساس هدف‌گذاری سال ۱۳۹۸، ۲۹ نقطه هدف‌گذاری شد که رتبه‌بندی نهایی نقاط در جدول ۴-۴۲ نشان داده شده است.

جدول ۱۶. اولویت‌بندی نهایی نقاط برای اصلاح در سال ۱۳۹۸

اولویت نقطه	محل نقطه حادثه‌خیز
۱	بلوار کرپلا روبروی شهرک نبی
۲	میدان امام حسن (تره‌بار جدید)
۳	بریدگی خانه‌سازی با بلوار دانشگاه
۴	خ موزه راهزان
۵	خ سیدالشهدا سهره الزهرا
۶	کمربندی جنوبی پیچ بالای شهرک ولیعصر
۷	خ سیدالشهدا راهزان
۸	پل شهر صنعتی
۹	ورودی داوران شهدا از کمربندی
۱۰	کمربندی شمالی پیچ طرمزد
۱۱	پل انتهای کشتارگاه
۱۲	تقاطع غفاری الزهرا
۱۳	مهارت راهزان
۱۴	شهرک حمید استخر اتحاد
۱۵	م شریعتی
۱۶	سهره شهرک قائم
۱۷	کمربندی شمالی روبروی پیام نور
۱۸	انتهای قائم‌مقام شهرک جوادیه
۱۹	م محیط زیست
۲۰	ورودی کمربندی به امام حسین

اولویت نقطه	محل نقطه حادثه خیز
۲۱	م انقلاب
۲۲	تقاطع جوادیه حمید
۲۳	بلوار سردشت بوتان گاز
۲۴	بریدگی مقابل پمپ بنزین کشتارگاه
۲۵	تقاطع اشکانی قائم مقام
۲۶	تقاطع العدیر جوادیه
۲۷	بریدگی لعلی و قائم مقام
۲۸	اواسط خ جنت پارک جنت
۲۹	بریدگی مقابل یگان ویژه

پس از تعیین نقاط اصلاحی در سال ۱۳۹۸ و هدف گذاری برای این سال در مرحله آخر راه کارهای متناسب با توجه به بودجه و محدودیت های مالی شهرداری اراک ارائه شد که این کار در نشست با حضور کارشناسان راهور شهرستان اراک و معاونت ترافیک شهرداری اراک انجام شد.

اولویت نقطه	محل نقطه حادثه خیز	راه کار
۱	میدان امام حسن (تره بار جدید)	اصلاح تقاطع غیرهم سطح (اجرا شده)
۲	بریدگی خانه سازی با بلوار دانشگاه	در حال مسدودسازی است
۳	خ سیدالشهدا سه راه الزهرا	انسداد تقاطع و اجرای دور برگردان
۴	خ سیدالشهدا راهزان	انسداد تقاطع و اجرای دور برگردان
۵	پل شهر صنعتی	اصلاح هندسی میدان
۶	ورودی داوران شهدا از کمربندی	یک طرفه کردن مسیر
۷	کمربندی شمالی پیچ طرمزد	ایمن سازی در جهت کاهش سرعت وسایل نقلیه با نصب علائم و ایجاد سرعت گیر
۸	مهارت راهزان	اصلاح هندسی و اجرای راستگرد و نصب علائم اخباری
۹	م شریعتی	تقاطع غیرهم سطح در حال اجرا
۱۰	سه راه شهرک قائم	پس از اجرای کامل خیابان مهارت مشکل مرتفع می شود
۱۱	کمربندی شمالی روبروی پیام نور	اجرای سرعت گیر برای کاهش سرعت وسایل نقلیه
۱۲	انتهای قائم مقام شهرک جوادیه	اصلاح هندسی
۱۳	م محیط زیست	اجرای تقاطع غیرهم سطح در حال اجراست

اولویت نقطه	محل نقطه حادثه خیز	راه کار
۱۴	ورودی کمربندی به امام حسین	کاهش سرعت وسایل نقلیه از طریق اجرای سرعت گیر و نصب علائم
۱۵	م انقلاب	اصلاح تقاطع غیرهم سطح
۱۶	تقاطع جوادیه حمید	شکل گیری جدید از طریق اصلاح هندسی
۱۷	بلوار سردشت بوتان گاز	در حال تعریض است
۱۸	بریدگی مقابل پمپ بنزین کشتارگاه	بسته شدن بریدگی
۱۹	تقاطع اشکانی قائم مقام	مسدود کردن تقاطع و اجرای دوربرگردان
۲۰	تقاطع الغدیر جوادیه	اجرای طرح هندسی
۲۱	بریدگی لمی و قائم مقام	بسته شدن بریدگی در دستورکار قرار خواهد گرفت
۲۲	اواسط خ جنت پارک جنت	اجرای طرح ۵۵ متری امیرکبیر

نتیجه گیری

همانگونه که اشاره شد، مهمترین عامل مؤثر بر نوع تصادفات در شهر اراک، محل وقوع تصادفات است. بنابراین می توان نتیجه گرفت که در نقاط مختلف شهر تصادفات های متفاوتی روی می دهد، درحالی که مهمترین عامل تصادف در نقاط اطراف شهر (به ویژه کمربندی های شمالی و جنوبی)، کنترل نکردن وسیله نقلیه، تجاوز از سرعت مجاز، خستگی و خواب آلودگی است. اما بررسی ها نشان داده است که در مکان های دیگر شهر توجه نکردن به جلو عامل اساسی در نوع تصادف به وقوع پیوسته است. همچنین از اهمیت بالای محل وقوع تصادفات فوتی می توان نتیجه گرفت که پراکندگی نقاط تصادف در اراک بسیار بالاست که باید این نقاط اصلاح شوند یا مورد پوشش راهنمایی و رانندگی قرار گیرند و در این پژوهش به این موارد پرداخته شده است.

با توجه به اینکه جنسیت در درجه دوم اهمیت قرار دارد، می توان نتیجه گرفت که جنسیت نیز در وقوع تصادفات اهمیت زیادی دارد. همچنان که با تحلیل داده های فوتی مشاهده می شود بیشتر خودروهایی که علت تصادف آنها کنترل نکردن وسیله نقلیه بوده، راننده خودرو مقصر، «مؤنث» بوده است. سومین عامل مؤثر بر تصادفات در پژوهش حاضر، محل فوت (بیمارستان یا صحنه تصادف) است که با توجه به درصد بالای فوتی ها در بیمارستان (حدود ۸۰ درصد) نسبت به فوت در صحنه تصادف، می توان به رسیدگی نکردن مناسب (طولانی شدن مدت زمان رساندن مجروح توسط آمبولانس به بیمارستان) و حتی وجود نداشتن اورژانس ها در مکان های حادثه خیز پی برد.

چهارمین عامل مؤثر در رتبه‌بندی عوامل مؤثر بر تصادفات، «سن راننده» است. این در حالی است که با توجه به ۲ بازه زمانی مهم در وقوع تصادفات (بیش از ۶۰ سال و کمتر از ۳۰ سال)، به نظر می‌رسد باید قوانینی وضع شود تا مهارت رانندگی کسانی که در بازه‌های سنی یادشده هستند، به درستی سنجیده و ارزیابی شود یا قوانینی وضع شود تا این گونه افراد کمتر مجال حادثه‌آفرینی را داشته باشند (همچون اضافه‌شدن جریمه‌ها یا دادن امتیاز منفی به متخلفان حادثه‌ساز).

در بحث ارائه راه کار، راه کارهایی که به منظور اصلاح نقاط حادثه‌خیز اشاره شد، بیشتر اصلاح هندسی بودند، اما در بحث ابزارهای الکترونیکی متأسفانه شهرداری اراک، بودجه‌ای به این مهم اختصاص نداد. با توجه به تخطی از سرعت مجاز که یکی از مهمترین علل وقوع تصادفات به‌ویژه در کمربندی‌های اطراف شهر اراک است، پیشنهاد می‌شود شهرداری اراک نسبت به نصب دوربین در محل‌های حادثه‌خیز اقدام کند.

پیشنهاد‌های پژوهش

- با توجه به نتایج به‌دست آمده برای کاهش تصادفات در شهر اراک پیشنهادهایی به شرح زیر است:
- ۱- اجرای دقیق هدف‌گذاری اصلاح نقاط در سال ۱۳۹۸ و ادامه اصلاح نقاط حادثه‌خیز کنونی؛
 - ۲- برنامه‌ریزی هر ساله راهور و معاونت ترافیک شهرداری اراک برای تعیین نقاط حادثه‌خیز، رتبه‌بندی و اصلاح آنها با استفاده از راه کارهای کوتاه‌مدت و بلندمدت؛
 ۳. فعال‌سازی بیشتر گشت‌های راهور در کمربندی‌های شمالی و جنوبی در جهت نظم‌دهی بیشتر به این مناطق؛
 ۴. اختصاص دادن نیروهای بیشتر (افسر و وظیفه) برای افزایش حضور فیزیکی پلیس در خیابان‌های شهر اراک؛
 ۵. استفاده از لوازم الکترونیکی نظیر دوربین کنترل سرعت در شهر اراک به‌ویژه کمربندی‌های شمالی و جنوبی؛
 ۶. نبود پارکینگ در مرکز شهر اراک و ضرورت ایجاد پارکینگ‌های طبقاتی در مرکز شهر؛
 ۷. اجرای شدید طرح توقیف موتورسیکلت‌های خطرآفرین در زمان‌های مختلف و جمع‌آوری موتورسیکلت‌های فاقد تجهیزات ایمنی یا حادثه‌آفرین؛
 ۸. تهیه هرچه بیشتر برنامه‌های آموزشی در رسانه‌های عمومی در راستای ایجاد فرهنگ عمومی استفاده از وسایل نقلیه به‌ویژه موتورسیکلت و پخش این برنامه‌ها به‌طور مداوم از شبکه استانی؛

۹. آموزش و بسترسازی مناسب به منظور تشویق ساکنان استان مرکزی و شهر اراک برای همراهی و همدمی بیشتر با پلیس راهور؛
۱۰. اهمیت بالایی محل فوت در حوادث منجر به فوت در شهر اراک و استقرار هر چه بیشتر واحدهای اورژانس در محل‌های حادثه‌خیز.

منابع:

منابع فارسی

ابراهیمی، علیرضا و صابر ساعتی (۱۳۹۰)، *ارزیابی عملکرد پژوهشی اساتید با استفاده از روش تحلیل پوششی داده‌ها*، مجله ریاضیات کاربردی واحد لاهیجان، سال هشتم، شماره ۲، صفحه ۸۰-۷۱.

ال. السون. د، دلن. د (۱۳۹۶)، *داده کاوی پیشرفته*. مترجم جندقی. غ، هاشمی. ا، شادکام. ا. کرمان: جهاد دانشگاهی استان کرمان.

اصغریور د. م (۱۳۹۵)، *تصمیم‌گیری‌های چندمعیاره*. چاپ هفتم. تهران: انتشارات دانشگاه.

بزله، آرزو، پیمان غلامی، نیما فقیه و محمد احمدی (۱۳۹۴)، *روشی نوین جهت طبقه‌بندی داده‌ها به کمک TOPSIS*، اولین کنفرانس ملی برق و رایانه ایران. دانشگاه آزاد نیریز. تهران: علوم رایانه.

بزله، آرزو، پیمان غلامی، نیما فقیه و محمد احمدی (۱۳۹۴)، *رویکردی به منظور رتبه‌بندی الگوریتم‌های طبقه‌بندی با کمک DEA*، دومین کنفرانس مباحث پیشرفته در علوم رایانه. دانشگاه تحصیلات تکمیلی علوم پایه زنجان. تهران: علوم رایانه.

سروش، عباس، احمد بحرینی نژاد (۱۳۸۸)، *هوشمندی کسب‌وکار و داده کاوی*. تهران: ناقوس.

سمیع‌زاده، رحیم (۱۳۸۷)، *داده کاوی و مدیریت ارتباط با مشتری*. تهران: رشد اندیشه.

شهرابی، ج (۱۳۹۶)، *داده کاوی*. جهاد دانشگاهی (دانشگاه صنعتی امیرکبیر).

شهرابی، جعفر، عباس ذوالقدر شجاعی (۱۳۹۶)، *داده کاوی پیشرفته مفاهیم و الگوریتم‌ها*. تهران: جهاد دانشگاهی.

شهرابی، جلال، وحید شکور نیاز (۱۳۹۶)، *مفاهیم داده کاوی در اوراکل*.

شهرابی، جواد، علی شجاعی (۱۳۹۶)، *داده کاوی پیشرفته مفاهیم و الگوریتم‌ها*. جهاد دانشگاهی (دانشگاه صنعتی امیرکبیر).

عسکرزاده، حسین (۱۳۸۸)، *مجموعه مقالات داده کاوی*. تهران: رویش جوانه‌های فردا.

علیزاده، سهراب، بابک تیمورپور و محمد غضنفری (۱۳۸۷)، *داده کاوی و کشف دانش*. دانشگاه علم و صنعت ایران.

کارتاردزیک، م (۱۳۹۶)، *داده کاوی (Data mining)*، مترجم: علیخانزاده، علوم رایانه
مشکانی، علی، علی ناظمی (۱۳۹۲)، *مقدمه‌ای بر داده کاوی*. نیشابور: دانشگاه آزاد نیشابور.

منابع لاتین

- A. Charnes, W. W. Cooper, and E. Rhodes. (1978). *Measuring the efficiency of decision making units*. European J. Oper. Res. 2 , 429-444
- Berry, J. A., & Linoff, G. (2016). *Data Mining Techniques for Marketing, Sales, and Customer Support*. New York: John Wiley & Sons.
- Chapman, P., Clinton, J. Kerber, R., Khabaz, Th. (2016). *www.crisp_dm.org*. Retrieved from *Crisp _ DM Step by Step Data Mining Guid*: http://www.crisp_dm.org/
- D. L. Olson, Wu, D. (2017). *Decision making with uncertainty and data mining. Lecture notes in artificial intelligence* .
- Dan Sun, Daoqiang Zhang. (2018). *Bagging Constraint Score for Feature Selection with Pairwise Constraint*. Pattern Recognition , 2106-2118.
- Gupta, G. K. (2006). *Introduction to Data Mining with Case Studies*. PHI Learning Pvt. Ltd.
- Han, J. and M. Kamber. (2015). *Chapter 1: Introduction and Advanced Topics*. In *Data Mining Concepts and Techniques*. San Francisco: Morgan Kaufmann.
- Hwang, C.L., and Yoon, K. (1985). *Multiple Attribute Decision Making*. Springer Verlag.
- I. H. Witten, Frank. (2016). *Data Mining-Practical Machine Learning Tools and Techniques*, second ed. San Francisco, CA: Morgan Kaufman..
- Journal: *Procedia Economics and Finance* - Volume 27, 2015, Pages 77–84.
- Westphal, C. and Blaxton. (2016). *Data Mining Solutions: Methods and Tools for Solving Real World Problems*. New York: John Wiley.