



نسخه از دور

GIS ایران



سنجش از دور و GIS ایران
Iranian Remote Sensing & GIS

سال هشتم، شماره اول، بهار ۱۳۹۵
Vol.8, No. 1, Spring 2016

۳۷-۵۴

توسعه یک مدل تصمیم‌گیری مبتنی بر محاسبات نرم جهت پیش‌بینی شدت تصادفات در راه‌های برون‌شهری

میثم عفتی*

۱. استادیار، گروه مهندسی عمران، دانشکده فنی دانشگاه گیلان، رشت

تاریخ پذیرش مقاله: ۱۳۹۵/۳/۲۳

تاریخ دریافت مقاله: ۱۳۹۴/۲/۱۹

چکیده

توسعه روش‌های جدید و هوشمند برای جلوگیری از وقوع تصادف یا کاهش شدت تصادفات در راه‌های برون‌شهری یکی از اهداف اصلی مطالعات ایمنی راه است. هدف این تحقیق تلفیق قابلیت‌های سیستم‌های اطلاعات مکانی (GIS) با روش‌های مبتنی بر محاسبات نرم، جهت برآورد شدت تصادفات و تعیین فاکتورهای مؤثر بر آن در راه‌های دوخطه برون‌شهری است. روش پیشنهادی با ارائه مدل درخت دسته‌بندی و رگرسیون فازی (FCART) و ایجاد پایگاه داده مکانمند متشکل از داده‌های تصادفات و اطلاعات راه و محیط مجاور آن در محور قزوین-رشت (ایران) بررسی می‌شود. نتایج با استفاده روش اعتبارسنجی ده‌قسمتی بر رویدادهایی که شدت تصادفات آنها معلوم است، ارزیابی و با مدل درخت دسته‌بندی و رگرسیون (CART) مقایسه می‌شود. نتایج نشان می‌دهد که مدل درخت دسته‌بندی و رگرسیون فازی در مقایسه با درخت تصمیم CART فرایند استنتاج قوی‌تری دارد و شدت تصادفات را با صحت بیشتری پیش‌بینی می‌کند. تحلیل حساسیت روش پیشنهادی ضمن کشف تأثیرات مکانی طرح هندسی و عوارض و کاربری‌های مجاور راه بر شدت تصادفات، نقص فنی خودرو، کمربند ایمنی و شرایط آب‌وهوایی را نیز مهم‌ترین فاکتورهای تأثیرگذار در شدت تصادف می‌شمارد. این مطالعه به متخصصان ایمنی راه کمک می‌کند تا عوامل مکانی تأثیرگذار در سطوح متفاوت شدت تصادفات را شناسایی کنند و اقدامات پیشگیرانه لازم را برای کاهش شدت یا جلوگیری از وقوع تصادفات انجام دهند.

کلیدواژه‌ها: تحلیل‌های مکانی، محاسبات نرم، شدت تصادف، راه دوخطه برون‌شهری

* نویسنده عهده‌دار مکاتبات: رشت، کیلومتر ۶ جاده رشت-تهران، دانشگاه گیلان، دانشکده فنی، گروه مهندسی عمران.

۱- مقدمه

با استناد به کمیته جهانی کنترل ایمنی راه‌ها^۱، هر ساله نزدیک به ۱/۲ میلیون نفر بر اثر تصادفات در سرتاسر جهان کشته می‌شوند و ۲۵-۲۰ برابر این تعداد آسیب‌های جدی می‌بینند. با افزایش استفاده از وسایل نقلیه در کشورهای در حال توسعه، به‌نظر می‌رسد آسیب‌های ناشی از تصادفات جاده‌ای سومین علت اصلی مرگ‌ومیر و معلولیت در سال ۲۰۲۰ شمرده شود. بنا بر پیش‌بینی‌های سازمان بهداشت جهانی^۲، تصادفات جاده‌ای از رتبه نهم علل مرگ‌ومیر در جهان در سال ۲۰۰۴، به رتبه پنجم در سال ۲۰۳۰ خواهد رسید (گزارش سازمان بهداشت جهانی، ۲۰۱۳). تحقیقات اخیر نشان می‌دهد که میزان تصادفات در ایران حدود بیست‌برابر (UNICEF, 2013)^۳ و میزان مرگ‌ومیر بیش از دوبرابر میانگین جهانی، ۲۵ نفر به‌ازای هر 100,000 نفر (Bhall et al., 2009)، است. از سویی، برآوردهای آماری نشان می‌دهد که هزینه‌های اقتصادی مرگ‌ومیر و صدمات جانی ناشی از تصادفات هفت درصد تولید ناخالص داخلی کشور است که این رقم، درمقایسه با کشورهای دیگر با درآمد متوسط که هزینه تصادفات در آنها حدود ۱/۵ درصد از تولید ناخالص داخلی را تشکیل می‌دهد، بالاست (Ayati, 2012).

کاهش میزان و شدت تصادفات یکی از اهداف اصلی متولیان حمل‌ونقلی در سراسر جهان شمرده می‌شود. برای افزایش ایمنی راه و کاهش شدت تصادفات جاده‌ای لازم است که اطلاعات کاملی از فاکتورهای تأثیرگذار در وقوع تصادفات، در بخش‌های گوناگون راه در اختیار باشد. صدیقی و همکاران^۴ (۲۰۱۲) در تحقیقات خود بیان می‌کنند یکی از کاستی‌های مطالعات صورت‌گرفته در تحلیل تصادفات جاده‌ای بی‌توجهی به تأثیرات مکانی راه است. با توجه به تحقیقات یامادا و تیل^۵ (2004) عوامل مکانی همراه با عوامل انسانی، خودرو، محیط و عوامل ترافیکی می‌توانند نقش مؤثری در تمرکز و افزایش شدت تصادفات در راه‌های برون‌شهری داشته باشند. بسیاری

از محققان نیز از تحلیل‌های مکانی صرفاً برای تحلیل و نمایش داده‌های تصادفات در راه‌های برون‌شهری استفاده کرده‌اند (Jha and McCall, 2001; Flahaut et al., 2002; Steenberghen et al., 2004; Erdogan et al., 2008; Gundogdu, 2010; Wang and Wang, 2011; Ha and Thill, 2011; Effati et al., 2012). این تحقیق‌ها از تحلیل‌های مکانی برای نمایش موقعیت‌های تصادفات روی نقشه‌های رقومی، خوشه‌بندی تصادفات و تشخیص نواحی حادثه‌خیز راه استفاده شده است. این درحالی است که نقش فاکتورهای مکانی و تأثیرات مکانی در شدت تصادفات به‌ندرت مورد مطالعه قرار گرفته است. کشف این تأثیرات چه‌بسا در برنامه‌ریزی‌های آتی ایمنی راه از سوی متولیان حمل‌ونقل جاده‌ای استفاده شود.

از دیدگاه روش مورد استفاده، بسیاری از تحقیق‌های پیشین تحلیل رگرسیون را برای برآورد شدت تصادفات و نیز، تعیین پارامترهای تأثیرگذار در وقوع تصادفات به‌کار برده‌اند. مدل‌های خطی کمترین مربعات، رگرسیون دوجمله‌ای منفی، پواسون، و مدل‌های پیچیده‌تری همچون رگرسیون به‌ظاهر نامربوط از رایج‌ترین این روش‌ها به‌شمار می‌آیند (Joshua and Garber, 1990; Martin, 2002; Miaou and Lum, 1993; Shanker et al., 1995; Theofilatos et al., 2012; Wang and Kockelman, 2007; Yannis et al., 2010). این تحقیق‌ها بیان می‌کنند که مدل‌های رگرسیون برای تحلیل تصادفات به فرضیات و روابط از پیش تعریف‌شده بین متغیرهای ورودی نیاز دارند. در صورت تغییر یا وجود ابهام در این روابط، نتایج و خروجی تحلیل‌های تصادفات نیز تحت تأثیر قرار خواهند گرفت.

به‌دلیل اهمیت استخراج دانش از داده‌های تصادفات، روش‌های مبتنی بر داده‌کاوی همچون خوشه‌بندی و طبقه‌بندی این امکان را در اختیار

1. Permanent International Association of Road Congresses (PIARC)
2. World Health Organization (WHO)
3. http://www.unicef.org/iran/media_4783.html, Accessed October 08, 2013
4. Siddiqui et al.
5. Yamada and Thill

دسته‌بندی و رگرسیون فازی، تاکنون برای تحلیل شدت تصادفات در راه‌های برون‌شهری به کار نرفته‌اند. هدف اصلی این مطالعه برآورد شدت تصادفات و شناسایی مهم‌ترین عوامل تأثیرگذار در شدت تصادفات در راه‌های برون‌شهری، با استفاده از یادگیری ماشین و تحلیل‌های مکانی است. روش‌های پیشنهادی تحقیق، برخلاف روش‌های آماری، بدون نیاز به فرضیات و روابط از پیش تعریف‌شده بین متغیرهای ورودی با کشف الگوهای پنهان در داده‌ها می‌توانند حجم بسیاری از داده‌های مکانی و غیرمکانی مرتبط با رویدادهای تصادف را تحلیل کنند. بدین منظور روشی بر پایه درخت دسته‌بندی و رگرسیون فازی^۱ ارائه و بر پایگاه داده تصادفات اعمال می‌شود. شناسایی حوزه مکانی و چگونگی تأثیر فاکتورهای مکانی در قطعات راه از دیگر اهدافی است که در این مطالعه بررسی می‌شود. نتایج این مطالعه در شناسایی فاکتورهای مؤثر در سطوح گوناگون شدت تصادف چه‌بسا کارشناسان پلیس و متخصصان ایمنی راه را در گرفتن تصمیمات لازم برای کاهش تصادفات جاده‌ای یاری کند.

۲- مواد و روش‌ها

در این بخش، نخست منطقه مورد مطالعه و داده‌های مورد استفاده بیان می‌شود و سپس، روش به کار برده‌شده جهت دستیابی به اهداف تحقیق تشریح می‌شود.

۲-۱- داده و منطقه مورد مطالعه

محور مورد مطالعه محور قدیم قزوین- رشت است که تهران را به شمال ایران متصل می‌کند. محدوده ارتفاعی این منطقه کوهستانی بین ۳۰۰ تا ۲۳۹۴ متر است.

1. artificial neural network

2. decision tree

3. Pakgohar et al.

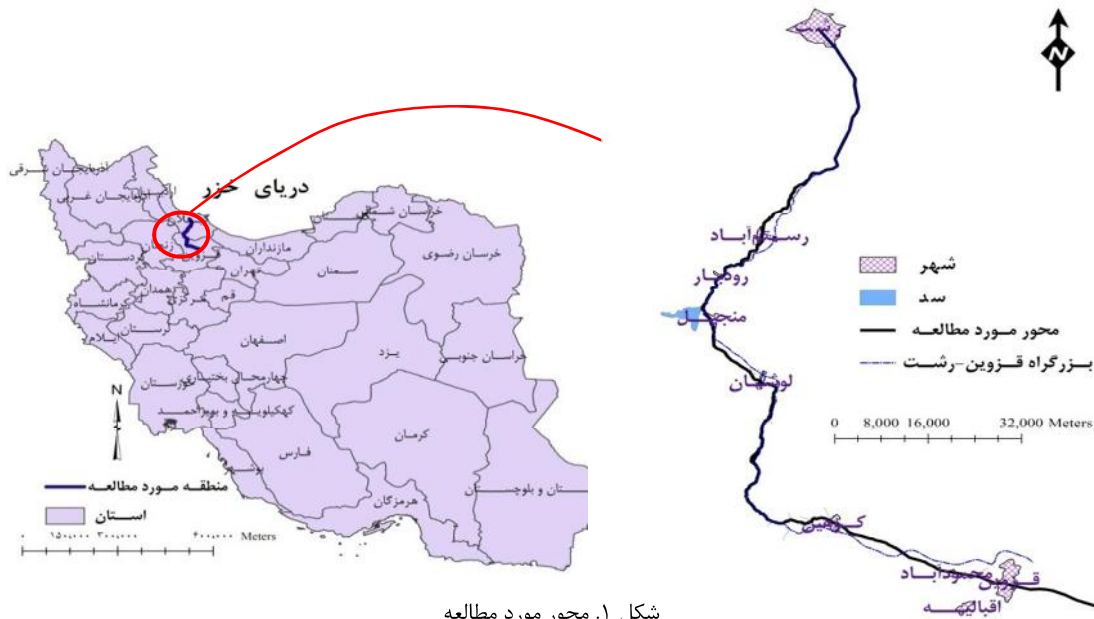
۴. Tavakoli Kashani and Shariat Mohaymany: این رفرنس بین‌المللی است به همین علت، به‌رغم نام ایرانی نویسنده، به انگلیسی ارجاع شده است

5. FCART

متخصصان قرار می‌دهند که حجم بسیاری از داده‌های مرتبط با تصادفات را تحلیل کنند. در روش‌های آماری در بیشتر موارد، فرض بر این است که توزیع داده‌ها نرمال است و درستی یا نادرستی نتایج نهایی به درست بودن فرض اولیه وابسته است. در حالی که روش‌های داده‌کاوی از هیچ فرضی در مورد داده‌ها تبعیت نمی‌کنند. در تحقیقات پیشین، تکنیک‌هایی از داده‌کاوی مانند شبکه‌های عصبی^۱ (Sohn and Hyungwon, 2001; Delen et al., 2006) و درخت تصمیم^۲ (Chang and Chen, 2005; Clarke et al., 1998) برای تعیین فاکتورهای مؤثر در شدت تصادفات استفاده شده‌اند. درخت‌های تصمیم‌گیری ابزاری قدرتمند در تعیین مهم‌ترین متغیرهای مستقل و حل مسائل دسته‌بندی و پیش‌بینی به‌شمار می‌روند. از مطالعات این حوزه می‌شود به تحقیقات پاک‌گهر و همکاران^۳ (۲۰۱۰) اشاره کرد که نقش فاکتورهای انسانی را در وقوع و شدت تصادفات در راه‌های ایران، با استفاده از مدل CART و روش رگرسیون لجستیک بررسی کردند. توکلی کاشانی و شریعت‌مهمینی^۴ (۲۰۱۱) نیز از مدل CART برای تحلیل شدت مصدومیت ناشی از تصادفات در راه‌های دوخطه برون‌شهری استفاده کردند. نتایج مطالعات آنها نشان می‌دهد که روش CART روشی کارآ با قابلیت تفسیرپذیری بالا و قادر به مدلسازی روابط پیچیده بین متغیرها است. این در حالی است که درخت‌های تصمیم به تغییر کوچکی در داده‌های ورودی بسیار حساس‌اند و قادر به مدلسازی عدم قطعیت در فرایند برآورد شدت تصادفات نیستند. برای رفع این کاستی‌ها، برخی از مطالعات پیشین تلفیق مفاهیم فازی و درخت تصمیم را جهت کلاسه‌بندی و پیش‌بینی پیشنهاد می‌کنند (Janikow, 2004; Umanol et al., 1994; Yuan and Shaw, 1995). طبق این تحقیق‌ها، درخت‌های تصمیم فازی در فرایند کلاسه‌بندی خطای کمتری دارند و در مقایسه با درخت‌های تصمیم و بسیاری از روش‌های یادگیری ماشین در مدلسازی الگوهای پیچیده کارآترند. درخت‌های تصمیم فازی و به‌ویژه درخت

بسیار مستعد تصادفات است. شکل ۱ منطقه مورد مطالعه را نمایش می‌دهد. این مطالعه از چندین منبع داده و در فرمت‌های گوناگون استفاده می‌کند. جدول ۱ داده‌های مورد استفاده تحقیق و ویژگی‌های آنها را توصیف می‌کند.

محور مورد مطالعه از مختصات جغرافیایی $19^{\circ} 45' N$ و $36^{\circ} 37' 57'' E$ در شهر قزوین آغاز و به مختصات جغرافیایی $30^{\circ} 29' E$ و $36^{\circ} 37' 34'' N$ در شهر رشت منتهی می‌شود. این مسیر از راه‌های اصلی برون‌شهری پرخطر کشور به‌شمار می‌آید که



شکل ۱. محور مورد مطالعه

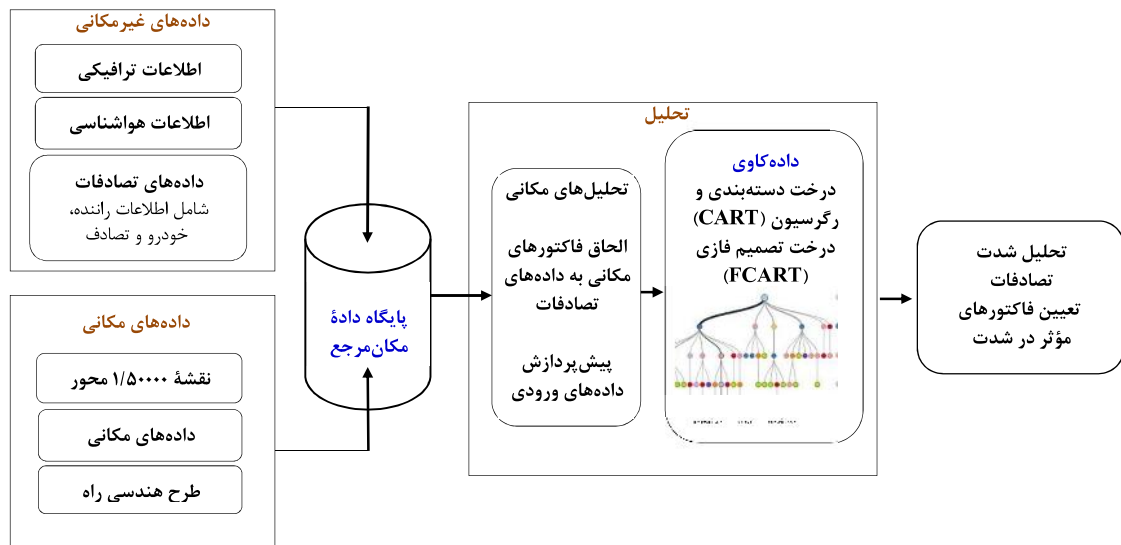
جدول ۱. توصیف داده‌های مورد استفاده

داده	منبع	مقیاس/قدرت تفکیک	توضیح
نقشه توپوگرافی	سازمان نقشه‌برداری کشور	۱/۵۰۰۰۰	رقومی
مدل رقومی ارتفاعی	سازمان نقشه‌برداری کشور	۱۰ متر	رقومی
اطلاعات هواشناسی	سازمان هواشناسی کشور	-	ایستگاه‌های هواشناسی
داده‌های تصادفات	پلیس راه کشور	-	توصیفی
عوارض راه و کاربری مجاور	سازمان راه‌داری و حمل‌ونقل جاده‌ای	-	مکانی و توصیفی
تصویر ماهواره‌ای	سازمان جغرافیای نیروهای مسلح	۵/۸ متر	IRS-LISS III-P6

۲-۲- روش تحقیق

دوخطه برون‌شهری می‌پردازد. بدین منظور، با ارجاع‌دهی خطی تصادفات بر محور و افزودن فاکتورهای مکانی به رکوردهای تصادفات، پیش‌پردازش‌های لازم روی داده‌های ورودی صورت می‌گیرد و عوامل مؤثر در تصادفات خسارتی، جرحی و فوتی تحلیل می‌شود.

روش پیشنهادی تحقیق براساس تلفیق تحلیل‌های مکانی و مدل درخت دسته‌بندی و رگرسیون فازی، جهت تحلیل شدت تصادفات و تعیین فاکتورهای مؤثر در آن استوار است. شکل ۲ روش کلی تحقیق را نشان می‌دهد. این مطالعه با استراتژی‌ای تلفیقی از مفاهیم فازی و مدل CART به تحلیل مکانی شدت تصادفات در راه‌های



شکل ۲. روش پیشنهادی تحقیق

در گره ایجاد و شاخص جینی صفر می‌شود. برعکس، اگر داده‌ها در گره به یک نسبت از هر دسته وجود داشته باشند، بیشترین مقدار شاخص جینی حاصل می‌شود. در واقع در هر گره، شاخص جینی برای همه متغیرها محاسبه می‌شود و متغیری به‌منزله متغیر جداکننده انتخاب می‌شود که کمترین مقدار شاخص جینی را داشته باشد. شاخص جینی با رابطه (۱) تعریف می‌شود.

$$G(m) = 1 - \sum_{j=1}^J p^2(j|m) \quad \text{رابطه (۱)}$$

که در این رابطه، مقادیر $p(m)$ و $p(j,m)$ از روابط زیر تعیین می‌شوند:

$$p(j,m) = \frac{\pi(j) N_j(m)}{N_j} \quad \text{رابطه (۲)}$$

$$p(m) = \sum_{j=1}^J P(j,m)$$

$$p(j|m) = \frac{p(j,m)}{p(m)}$$

1. splitter
2. purity
3. leaf or terminal node
4. Gini coefficient

۲-۱- مدل درخت دسته‌بندی و رگرسیون فازی پیشنهادی

ساختار مدل درخت دسته‌بندی و رگرسیون فازی پیشنهادی در این تحقیق به‌گونه‌ای است که نخست، همه داده‌ها در اولین گره، که گره ریشه نامیده می‌شود، قرار می‌گیرند و سپس براساس متغیر جداکننده^۱ که بیشترین همگنی و خلوص^۲ را در هر شاخه ایجاد می‌کند، در ریشه انشعاب ایجاد می‌شود. عمل تقسیم‌بندی متغیرها در شاخه‌های درخت آن قدر ادامه می‌یابد که داده‌های موجود در هر گره بیشترین همگنی را برای تعلق به دسته‌ای خاص داشته باشند. گره‌هایی که در انتهای درخت قرار می‌گیرند، گره نهایی یا برگ^۳ نامیده می‌شوند. به گره‌هایی که میان گره ریشه و گره‌های نهایی قرار دارند، گره میانی گفته می‌شود. شاخص‌های گوناگونی برای انتخاب متغیر جداکننده و نیز، جداسازی هر گره به دو زیرگره وجود دارد که شناخته‌شده‌ترین آنها برای داده‌های اسمی شاخص جینی^۴ است. چنانچه همه داده‌های موجود در یک گره متعلق به یک دسته باشند، بیشترین خلوص

در این رابطه، S_{sum} مجموع مقادیر عضویت محاسبه شده با استفاده از k مقدار میانگین M_j است. بیشترین مقدار شاخص جینی فازی $Gini(\omega_k)$ بیانگر بهترین دسته‌بندی در نود t است. رشد درخت نیز تا جایی ادامه می‌یابد که در گره‌ها، مشاهدات یکسانی از یک متغیر وجود داشته باشد. در چنین حالتی درخت بر داده‌های آموزشی برآزش بیش از حد^۱ دارد و بزرگ‌ترین درخت^۲ ایجاد می‌شود.

در مدلسازی متغیرهای پیوسته، مدل CART محدوده متغیرهای پیوسته را به بازه‌های کریسپ تقسیم می‌کند. این فرایند ممکن است به ایجاد بازه‌هایی غیرمعمول منجر شود که قدرت تفسیرپذیری الگوهای استخراجی از مدل را کاهش می‌دهد. یکی دیگر از مشکلات استفاده از متغیرهای پیوسته در مدل CART این است که تعداد بازه‌های مورد نظر جهت تقسیم‌بندی متغیرهای پیوسته به وسیله مدل قابل دریافت و اجرا نیست، هر چند کارشناسان این بازه‌ها را از پیش تعریف و مشخص کرده باشند. این درحالی است که مدل درخت دسته‌بندی و رگرسیون فازی پیشنهادی می‌تواند با استفاده از مجموعه‌های فازی و تعریف مقادیر زبانی، نظر متخصصان را در تعریف بازه‌ها و تقسیم‌بندی‌ها لحاظ کند. چنانچه این اطلاعات موجود نباشد، روش‌های اتوماتیکی برای ایجاد بازه‌های مقادیر زبانی قابل استفاده‌اند که نقاط تقسیم منطقی و مناسبی را با توجه به تغییرات مقادیر متغیر مورد نظر در داده‌های آموزشی ایجاد می‌کنند.

تکرار در استفاده از یک متغیر در الگو و وجود چندین تقسیم‌بندی از بازه مقادیر ممکن، تفسیرپذیری الگو را در مدل‌های درخت تصمیم کاهش می‌دهد. در مدل پیشنهادی، درخت تصمیم با متغیرهای فازی شده ایجاد می‌شود بنابراین، متغیرها بیش از یکبار در هر الگو به کار نمی‌روند. از سوی دیگر، در درخت‌های

در این روابط J بیانگر دسته‌های شدت تصادفات است، $\pi(j)$ احتمال اولیه دسته z است که از سوی متخصصان مورد نظر تعیین می‌شود. N_j تعداد کل مشاهدات مربوط به کلاس z در گره ریشه درخت و $N_j(m)$ تعداد مشاهدات مربوط به دسته z در گره m است. $p(j, m)$ نیز احتمال قرارگیری داده‌های مربوط به دسته z در گره m است.

در مدل پیشنهادی تحقیق یک شاخص جینی فازی برای دسته‌بندی و استنتاج بر مبنای داده‌های آموزشی ارائه شد. فرض کنید A_1, A_2, \dots, A_N داده‌های آموزشی باشند که باید در یکی از کلاس‌های شدت تصادفات C_1, C_2, C_3 و C_4 دسته‌بندی شوند و مقادیر عضویت داده‌های ورودی، μ_i برای داده ورودی i ام، از کوچک به بزرگ مرتب شوند. برای محاسبه متغیر جداکننده، میانگین دو تابع عضویت متوالی با استفاده از رابطه (۳) محاسبه می‌شود.

$$\text{رابطه (۳)} \quad M_j = (\mu_i + \mu_{i+1})/2 \quad j=1,2,3,\dots, N-1$$

در این رابطه، هر مقدار M_j مقادیر عضویت را که از کوچک به بزرگ مرتب شده‌اند، به بخش‌های چپ و راست تقسیم می‌کند. مقادیر توابع عضویت در سمت چپ، کمتر از میانگین و در سمت راست، بیشتر از مقدار میانگین است. چنانچه ω_j^L مجموع مقادیر توابع عضویت کلاس زام در سمت چپ و ω_j^R مجموع مقادیر توابع عضویت کلاس زام در سمت راست k امین مقدار میانگین باشد، معیار جداکننده در نود t درخت از رابطه (۴) محاسبه می‌شود.

$$\text{رابطه (۴)} \quad Gini(\omega_k) = \sum_{j=1}^4 \frac{\omega_j^L}{S_{sum}} \times \left\{ 1 - \sum_{j=1}^4 \left(\frac{\omega_j^L}{\sum_{j=1}^4 \omega_j^L} \right)^2 \right\} + \sum_{j=1}^4 \frac{\omega_j^R}{S_{sum}} \times \left\{ 1 - \sum_{j=1}^4 \left(\frac{\omega_j^R}{\sum_{j=1}^4 \omega_j^R} \right)^2 \right\}$$

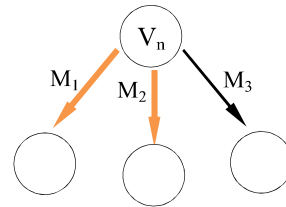
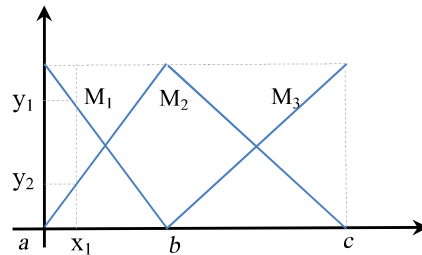
1. Overfitting
2. Maximal tree

۲-۲- متغیرهای مورد استفاده در تحقیق

داده‌های تصادفات پیشین محور مورد مطالعه شامل بیش از شصت ویژگی‌اند که خصوصیات هر رویداد تصادف را بیان می‌کنند. این داده‌ها از سوی افسر پلیس برای هر رویداد تصادف تکمیل می‌شوند و در پایان هر ماه، با توجه به وضعیت مصدومان و افزایش تعداد متوفیان تصادف در روزهای پس از حادثه، اطلاعات به‌روز می‌شود. مطالعات پیشین نشان می‌دهد که داده‌های تصادفات از سالی به سال دیگر متفاوت است و تغییرپذیری آماری چشمگیری دارد. بدین منظور، ۴۹۵۷ رکورد متشکل از رویدادهای تصادف و بدون تصادف در فرمت متن، عدد، تاریخ و زمان در دوره زمانی چهارساله، در محور مورد مطالعه بررسی شد. با توجه به متغیرهای مورد استفاده در تحقیقات پیشین برای برآورد شدت تصادفات و نیز در دسترس بودن داده‌ها، برخی از ویژگی‌های تصادف انتخاب و در محیط GIS با داده‌های مکانی تلفیق شدند (شکل ۴). رفرنس‌دهی خطی تصادفات روی محور، الحاق اطلاعات مکانی راه، محیط و کاربری‌های مجاور را به داده‌های تصادفات امکان‌پذیر می‌کند. با استفاده از روابط توپولوژیکی چون فاصله، هم‌پوشانی^۱ و مجاورت^۲، معیارهای مکانی استخراج می‌شوند و پس از الحاق به رکوردهای تصادفات در تحلیل شدت تصادفات مورد بررسی قرار می‌گیرند.

از آنجاکه ورودی شامل فاکتورهای مکانی راه و عوارض مجاور آن، ترافیک و نیز بسیاری از متغیرهای ثبت‌شده در رکوردهای تصادف است، این حجم از داده ورودی فرایند پیش‌بینی شدت تصادف را پیچیده می‌کند. از سویی، چنانچه متغیرهای ورودی مدل وابسته باشند و یا ارتباطی با خروجی نداشته باشند، صحت و مطمئن بودن پیش‌بینی یا دسته‌بندی تحت تأثیر قرار خواهند گرفت (Mitchell, 1997). استفاده از

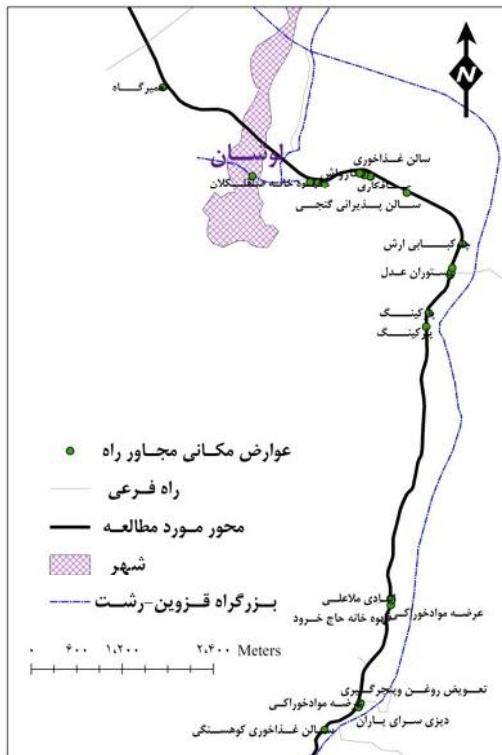
تصمیم معمولاً مجموعه‌ای از الگوهای مجزا از هم وجود دارند که فقط یک الگو برای کلاسه‌بندی و پیش‌بینی اجرا می‌شود. در حالی که در مدل پیشنهادی معمولاً دو شاخه از درخت همزمان اجرا و درجهت سازگاری با ورودی مدل چک می‌شود. شکل ۳ این ویژگی درخت تصمیم فازی را تشریح می‌کند.



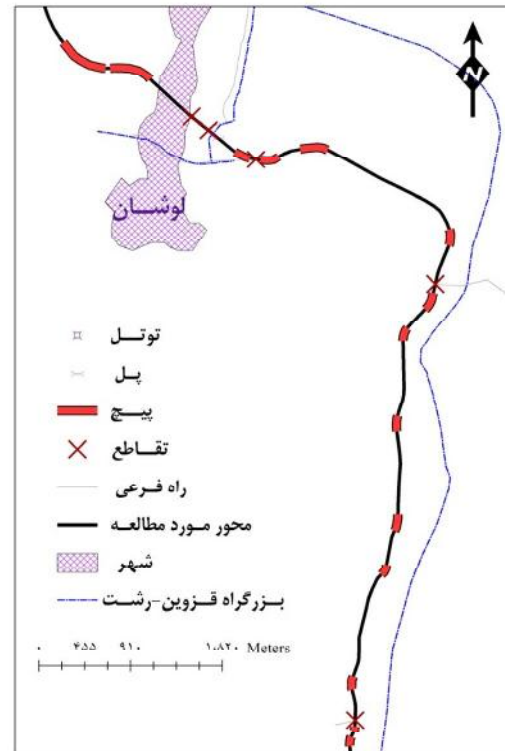
شکل ۲. استنتاج به روش FCART

در این شکل، ورودی x_1 به ترتیب، در نقاط y_1 و y_2 با توابع عضویت M_1 و M_2 تقاطع می‌یابد. بنابراین، بخش‌های M_1 و M_2 از درخت تصمیم فازی که با فلش‌های جهتی نارنجی پرنج در شکل ۲ نمایش داده شده‌اند، اجرا خواهند شد. برای هر مقدار ورودی در بازه a تا b ، بخش‌های M_1 و M_2 و برای هر مقدار ورودی در بازه b تا c بخش‌های M_2 و M_3 از درخت تصمیم فازی اجرا خواهند شد. از آنجاکه در مدل پیشنهادی با استفاده از توابع عضویت تعریف شده برای هر متغیر، چندین الگو از درخت تصمیم اجرا می‌شود، رویداد در کلاس الگویی قرار می‌گیرد که بیشترین سازگاری را با داده ورودی داشته باشد. بدین ترتیب، مدل FCART پایدار و صحت بیشتری در فرایند کلاسه‌بندی و پیش‌بینی شدت تصادفات خواهد داشت. در ادامه، ضمن بیان متغیرهای مورد استفاده در پژوهش، فرایند پیاده‌سازی روش پیشنهادی و نتایج حاصل تشریح می‌شوند.

1. Overlay
2. proximity



د) کاربری‌ها و تسهیلات مجاور راه



ج) طرح هندسی راه

شکل ۴. عوارض مکانی در امتداد قطعه‌ای از راه

الگوهای نامناسب جلوگیری شود. این پیش‌پردازش‌ها شامل نرمال‌سازی متغیرهای ورودی، هم‌مقیاس کردن و حذف داده‌های تکراری و ناقص است. افزون بر این، برخی از متغیرهای ورودی با هم ترکیب شدند تا حافظه و حجم پردازش‌های لازم برای تحلیل شدت تصادفات کاهش یابد.

۲-۳- نتایج و بحث

نتایج به‌کارگیری روش پیشنهادی در محور مورد مطالعه نشان داد چنانچه درخت‌های تصمیم‌گیری با مفاهیم فازی تلفیق شوند، در کلاس‌بندی مقادیر ورودی در بخش‌های مرزی متغیرهای پیوسته، استنتاج داده‌های مبهم همراه با عدم قطعیت و همچنین، مدل‌سازی الگوهای پیچیده مکانی کارآتر خواهند بود.

۱. بدون تصادف

متغیرهای وابسته، ضمن افزایش مدت زمان لازم جهت استنتاج مدل، کارایی فرایند پیش‌بینی را کاهش می‌دهد و تفسیر نتایج به‌درستی صورت نخواهد گرفت. بنابراین، پیش از فرایند پیش‌بینی با استفاده از تست χ^2 وابستگی متغیرها بررسی شد و با حفظ صحت، تعداد متغیرهای ورودی به کمترین میزان ممکن کاهش یافت. جدول ۲ متغیرهای ورودی و ویژگی‌های آن‌ها را که در نهایت، برای برآورد شدت تصادف در درخت دسته‌بندی و رگرسیون فازی به کار می‌روند، نمایش می‌دهد. خروجی مدل نیز برآورد میزان شدت تصادف در یکی از دسته‌های خسارتی، جرحی، فوتی و ایمن^۱ است.

همچنین شایان ذکر است که پیش‌پردازش‌های لازم روی داده‌های تصادفات و داده‌های مکانی پیش از فرایند داده‌کاوی انجام شد تا از ایجاد خوشه‌ها و

جدول ۲. توصیف متغیرهای ورودی

کلاس	متغیر	توضیح
راننده	کمر بند ایمنی	وضعیت بستن کمر بند ایمنی راننده
	سن	سن راننده
	جنسیت	جنسیت راننده
	تحصیلات	میزان تحصیلات راننده
ترافیکی و محیطی	عوامل انسانی تصادف	خواب‌آلودگی، خستگی، عجله، بی‌توجهی به قوانین، مصرف مشروبات الکلی، هیچ‌یک از موارد صبح‌زود (۰۹:۰۰-۰۷:۰۱)، روز (۰۹:۰۰-۱۶:۰۰)، عصر (۱۶:۰۰-۱۹:۰۰)، شب (۱۹:۰۰-۲۳:۰۰)، نیمه‌شب (۰۷:۰۰-۲۳:۰۰)
	روز هفته	اوایل هفته، میان هفته، اواخر هفته، تعطیلی آخر هفته
	شرایط نوری	روز یا شب
	آب‌وهوا	ابری، مه‌آلود، آفتابی، بارانی، برفی، طوفانی
وسيلة نقلیه	نقص فنی خودرو	لاستیک، سیستم نوردهی، ترمز، فرمان، برف‌پاک‌کن، زنجیر چرخ، سرعت، هیچ‌یک از موارد
	نوع خودرو	سواری، کامیونک و مینی‌بوس، اتوبوس، کامیون، دوچرخه یا موتورسیکلت
	مانع دید	درخت، ساختمان، نور خورشید، کوه، خودروهای دیگر، هیچ‌یک از موارد
	شرایط سطح جاده	ماسه‌ای، خشک، مرطوب، یخی
راه	نقص راه	اختلاف ارتفاع آسفالت و شانه راه، نبود حفاظ، مانع، علائم افقی، علائم عمودی، شانه راه و پارکینگ، هیچ‌یک از موارد
	فاصله از کاربری‌های مجاور راه	فاصله از اماکن تجاری (رستوران، تعمیرگاه، مسجد و ...) و مسکونی؛ عددی بین ۵۰ تا ۶۰۰۰ متر
	عرض راه	حداقل ۳ و حداکثر ۴۰ متر در امتداد راه
	شیب	شیب جاده در موقعیت مکانی خودرو بین نیم تا سی درصد
	پیچ	وجود یا وجود نداشتن
	تقاطع	وجود یا وجود نداشتن
	تونل	وجود یا وجود نداشتن

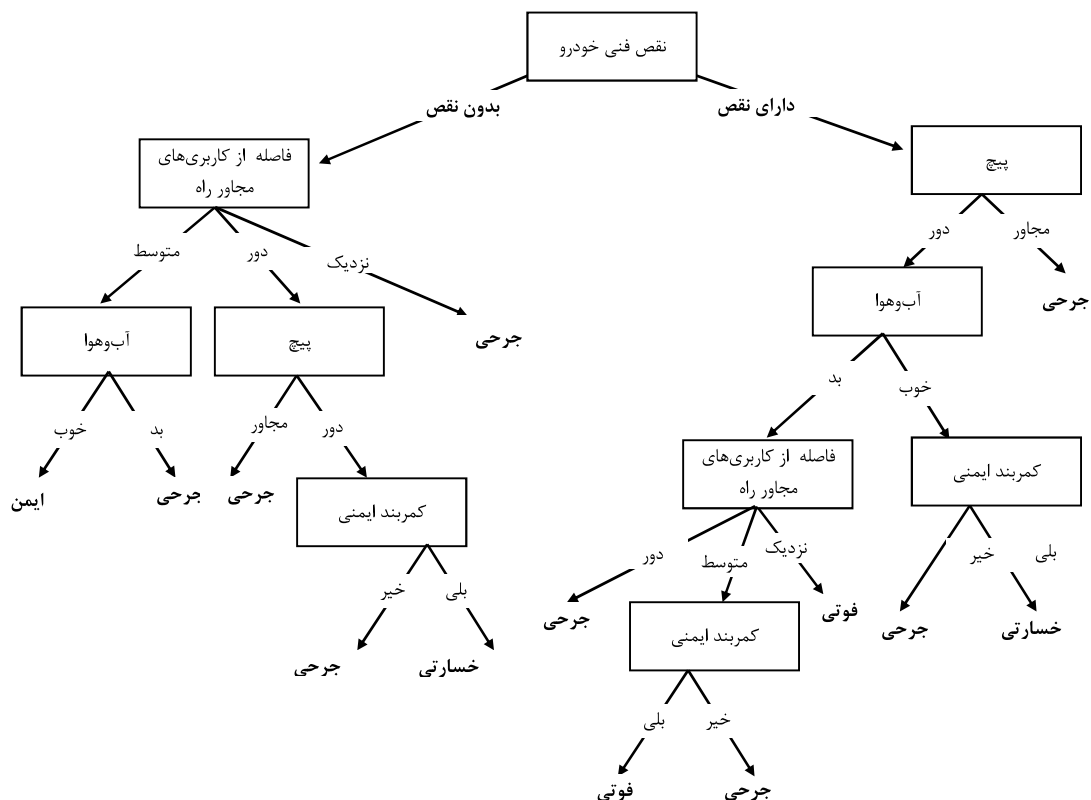
در محور مورد مطالعه است. گره ۲ در سمت راست درخت داده‌های مربوط به خودروهایی را نشان می‌دهد که یکی از نقص‌های فنی لاستیک، سیستم نوردهی، ترمز، فرمان، برف‌پاک‌کن، زنجیر چرخ و سرعت را دارند. اگر خودرو دارای نقص فنی و در مجاورت پیچ باشد، تصادف جرحی محتمل‌تر است. در غیر این صورت، شدت تصادف وابسته به شرایط آب‌وهوایی بستگی دارد. در شرایط آب‌وهوایی آفتابی و ابری، در صورت بسته بودن کمر بند ایمنی، تصادفات خسارتی محتمل‌تر از تصادفات جرحی و فوتی خواهند بود، در حالی که در شرایط آب‌وهوایی بارانی و برفی، شدت تصادف افزون بر کمر بند ایمنی، به فاصله از کاربری‌ها و تسهیلات

۱. یعنی بیشتر مشاهدات آن گره در یک دسته باشند.

شکل ۶ درخت دسته‌بندی و رگرسیون ایجاد شده به روش FCART را پس از فرایند هرس که با توجه به نظر متخصصان ایمنی راه انجام گرفته، نمایش می‌دهد. در این ساختار درختی، متغیرهای ورودی پیش از فرایند استنتاج با استفاده از مجموعه‌های فازی، فازی می‌شوند. سپس داده‌های ورودی در گره ریشه قرار می‌گیرند و براساس متغیری که بیشترین همگنی را برای هر شاخه ایجاد می‌کند، در ریشه انشعاب ایجاد می‌شود. بنابراین، در این درخت متغیری به‌منزله متغیر جداکننده انتخاب شده که کمترین مقدار جینی برای آن به‌دست آمده است^۱. مطابق شکل ۶، گره ریشه به‌وسیله متغیر «نقص فنی خودرو» به دو شاخه تقسیم شده است. این الگو نشان می‌دهد که «نقص فنی خودرو» مهم‌ترین متغیر تأثیرگذار در شدت تصادفات

«پیچ» و «شرایط آب و هوایی» و گره برگ تصادفات با شدت جرحی تقسیم می‌شود. گره برگ نشان می‌دهد که در بخش‌هایی از راه که در مجاورت کاربری‌ها و تسهیلات تجاری قرار دارد، احتمال تصادفات جرحی زیاد است. اگر خودرو از کاربری‌ها و تسهیلات مجاور راه دور باشد اما در مجاورت پیچ قرار گرفته باشد، شدت تصادفات باز هم از نوع جرحی خواهد بود. اگر خودرو در مجاورت پیچ، کاربری‌ها و تسهیلات مجاور راه نباشد، این کمربند ایمنی است که تأثیر بسیاری در شدت تصادفات خواهد داشت، درحالی‌که در فواصل متوسط از کاربری‌های مجاور راه، شرایط آب‌وهوایی بر وقوع تصادفات و شدت آن تأثیرگذار است.

مجاور راه وابسته است. نگاهی دقیق به درخت تصمیم فازی شکل ۵ نشان می‌دهد که در شش گره از هفت گره برگ سمت راست درخت، شدت تصادفات جرحی یا فوتی است. این موضوع دلالت بر آن دارد که اگر سرعت وسیله نقلیه زیاد باشد و یا خودرو دارای یکی از نقص‌های لاستیک، سیستم نوردهی، ترمز، فرمان، برف‌پاک‌کن، و زنجیرچرخ باشد، صرف‌نظر از تأثیر سایر عوامل، شدت تصادفات جرحی یا فوتی خواهد بود. آخرین گره در سمت راست درخت نیز نشان می‌دهد که اگر سرنشینان خودرو از کمربند ایمنی استفاده کنند، از میزان تصادفات فوتی کاسته خواهد شد. در سمت چپ درخت نیز گره ۱ به وسیله متغیر «فاصله از کاربری‌های مجاور راه»، به گره‌های میانی



شکل ۵. درخت دسته‌بندی و رگرسیون فازی برای تحلیل شدت تصادفات

جدول ۳. بررسی کارایی مدل FCART در پیش‌بینی کلاس‌های گوناگون شدت تصادفات

روش	کلاس	بازخوانی	دقت	میزان TP
FCART	فوتی	۰/۶۷	۰/۸۱	۰/۶۷
	جرحی	۰/۸۴	۰/۷۷	۰/۸۴
	خسارتی	۰/۷۵	۰/۷۲	۰/۷۵
	بدون تصادف	۰/۷۲	۰/۷۸	۰/۷۲
CART	فوتی	۰/۵۳	۰/۷۳	۰/۵۳
	جرحی	۰/۸۲	۰/۷۱	۰/۸۲
	خسارتی	۰/۷۰	۰/۶۹	۰/۷۰
	بدون تصادف	۰/۶۴	۰/۷۳	۰/۶۴

نتایج گویای آن است که مدل درخت دسته‌بندی و رگرسیون فازی، درمقایسه با مدل CART، شدت تصادفات را با درستی بیشتری دسته‌بندی می‌کند. معیارهای بازیابی و دقت در جدول ۳ نشان می‌دهند که پیش‌بینی تصادفات فوتی با استفاده از مدل درخت دسته‌بندی و رگرسیون، درمقایسه با مدل CART، بهبود چشمگیری داشته است.

همان‌طور که در بخش‌های پیشین بیان شد، یکی از اهداف این تحقیق تعیین فاکتورهای مؤثر در شدت تصادفات در محور مورد مطالعه است. مدل‌های درخت دسته‌بندی و رگرسیون این قابلیت را دارند که پس از ایجاد درخت بهینه، اهمیت متغیرهای ورودی را با استفاده از شاخص اهمیت متغیر در اختیار قرار می‌دهند. شاخص اهمیت متغیر از رابطه (۵) تعیین می‌شود.

$$\text{VIM}(v_j) = \sum_{m=1}^M \frac{n_m}{N} \Delta G(S(V_j, m)) \quad (5)$$

در این رابطه M تعداد گره‌ها، N تعداد داده‌ها و $S(v_j, m)$ جداکننده متغیر v_j در محل گره m است. $\frac{n_m}{N}$ نیز نسبت داده‌ها در گره m است که میانگین وزنی داده‌ها را در فرایند اهمیت‌دهی متغیرها وارد می‌کند. در رابطه (۴)، $\Delta G(S(V_j, m))$ تغییرات

درنهایت، باید یادآور شد که استفاده از درخت دسته‌بندی و رگرسیون فازی پیشنهادی معایبی نیز دارد. برای نمونه، درخت اولیه‌ای که به‌وسیله مدل FCART از داده‌های آموزشی ایجاد می‌شود، دارای ساختار بهینه نیست و نیاز به اصلاح یا هرس^۱ دارد. افزون بر این، تغییرات کمی در داده‌های آموزشی ممکن است به تغییر ساختار درخت تصمیم فازی منجر شود. تعریف توابع عضویت مناسب برای هر متغیر نیز از دیگر پیچیدگی‌های این روش است. بنابراین، در این مطالعه توابع عضویت مناسب برای هر متغیر با در نظر گرفتن بازه‌های استاندارد تعریف‌شده در دستورالعمل‌های مصوب موجود، نتایج تحقیقات پیشین و نیز، نظر متخصصان تعریف شد.

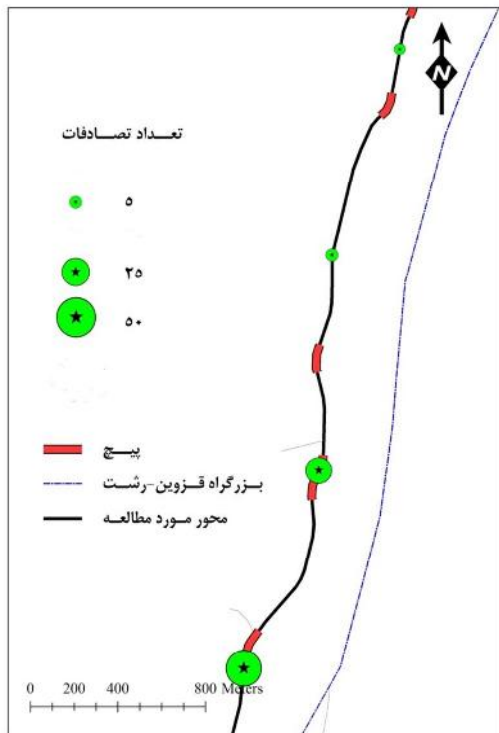
۳- ارزیابی و تعیین فاکتورهای مؤثر

برای بررسی درستی روش پیشنهادی در برآورد شدت تصادفات، با به‌کارگیری فرایند اعتبارسنجی ده‌قسمتی^۲، مدل ایجادشده روی داده‌هایی که شدت تصادفات آنها مشخص است، آزموده شد. طی این فرایند، داده‌های ورودی به ده زیرمجموعه افراز شد. در هر بار تکرار، یک زیرمجموعه برای اعتبارسنجی و نه زیرمجموعه دیگر برای آموزش مدل به‌کار برده شدند. این فرایند ده بار تکرار شد، به‌طوری‌که همه داده‌ها دقیقاً یکبار برای آموزش و یکبار برای اعتبارسنجی به‌کار رفتند. درنهایت، میانگین نتایج ده فرایند اعتبارسنجی به‌منزله تخمینی از کارایی و درستی مدل پیشنهادی تفسیر شد. نتایج ارزیابی صحت کلی کلاسه‌بندی را ۷۶/۸۲ و ضریب کاپا را ۰/۶۷ برآورد کرده است. این درحالی است که با به‌کارگیری مدل CART روی داده‌های ورودی، صحت کلی کلاسه‌بندی ۷۱/۵۸ و ضریب کاپا ۰/۵۹ برآورد شده است. جدول ۳ با استفاده از شاخص‌های گوناگون ارزیابی، کارایی مدل FCART را در پیش‌بینی کلاس‌های متفاوت شدت تصادفات با مدل CART مقایسه می‌کند.

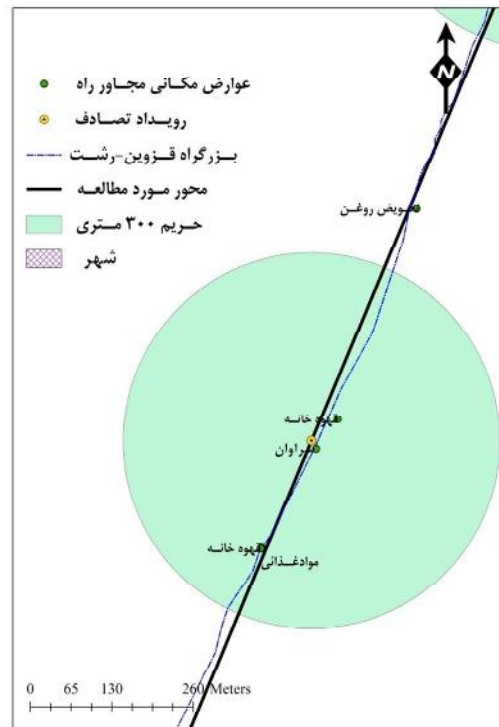
1. pruning
2. 10-fold cross validation

تحلیل حساسیت مدل پیشنهادی نشان می‌دهد که افزون بر نقص فنی خودرو، فاکتورهای مکانی‌ای مانند پیچ و فاصله از کاربری‌ها و تسهیلات مجاور راه، تأثیر بسزایی در شدت تصادفات در راه‌های دوخطه برون‌شهری دارند. تحلیل مکانی شدت تصادمیت تصادفاتی که در مجاورت کاربری‌ها و تسهیلات اطراف راه رخ داده‌اند، نشان می‌دهد در بخش‌هایی از راه که در فاصله کمتر از ۳۰۰ متری کاربری‌ها و تسهیلات رفاهی قرار دارند (شکل ۶-الف)، تصادفات معمولاً از نوع جرحی یا فوتی بوده است. افزون بر این، بیشتر تصادفات رخ داده در این بخش‌های راه از نوع وسیله‌نقلیه-عابر، وسیله‌نقلیه-وسیله‌نقلیه، وسیله‌نقلیه-موتورسیکلت بوده است. بررسی تصادفاتی که در مجاورت پیچ‌ها رخ داده‌اند، نشان می‌دهد که وجود پیچ‌های با شعاع کم، افزون بر افزایش شدت تصادفات، در میزان تصادفات نیز تأثیرگذار است. مطابق شکل ۶-ب، تعداد تصادفات در مجاورت پیچ‌ها بیشتر از بخش‌های مستقیم راه است.

شاخص جینی در گره m براساس متغیر V_j است. شاخص اهمیت را می‌شود در مدل FCART برای همه متغیرهای ورودی که وابسته نبودن آنها پیش‌تر بررسی شد، محاسبه کرد، طوری که متغیری که بیشترین سهم را در فاکتور خروجی دارد، بزرگ‌ترین عدد را در مقایسه با بقیه اختیار می‌کند. شکل ۸ اهمیت نسبی فاکتورهای مؤثر در شدت تصادفات را در مدل پیشنهادی نمایش می‌دهد. مطابق این شکل، نقص فنی خودرو مهم‌ترین فاکتور تأثیرگذار در شدت تصادفات در محور مورد مطالعه است. بی‌توجهی به استانداردهای ایمنی در طراحی خودرو و همچنین، استفاده از خودروهای فرسوده در بازه زمانی‌ای که داده‌های تصادفات جمع‌آوری شدند، از مهم‌ترین دلایلی به‌شمار می‌روند که فاکتور نقص فنی خودرو، به‌منزله مهم‌ترین فاکتور تأثیرگذار در شدت تصادفات، در مدل پیشنهادی شناسایی شده‌اند.



ب: مجاورت به پیچ‌ها

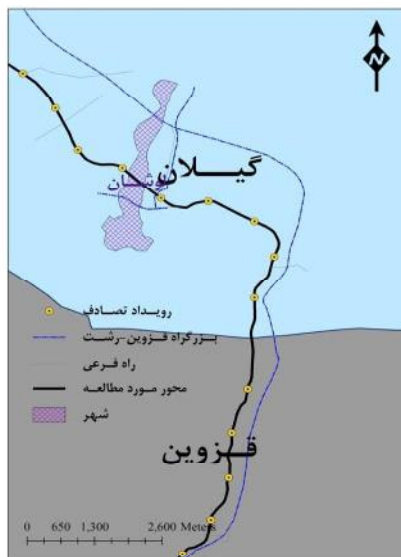


الف: مجاورت به کاربری‌های اطراف راه

شکل ۶. تحلیل مکانی تصادفات در مجاورت پیچ‌ها و کاربری‌ها و تسهیلات مجاور راه

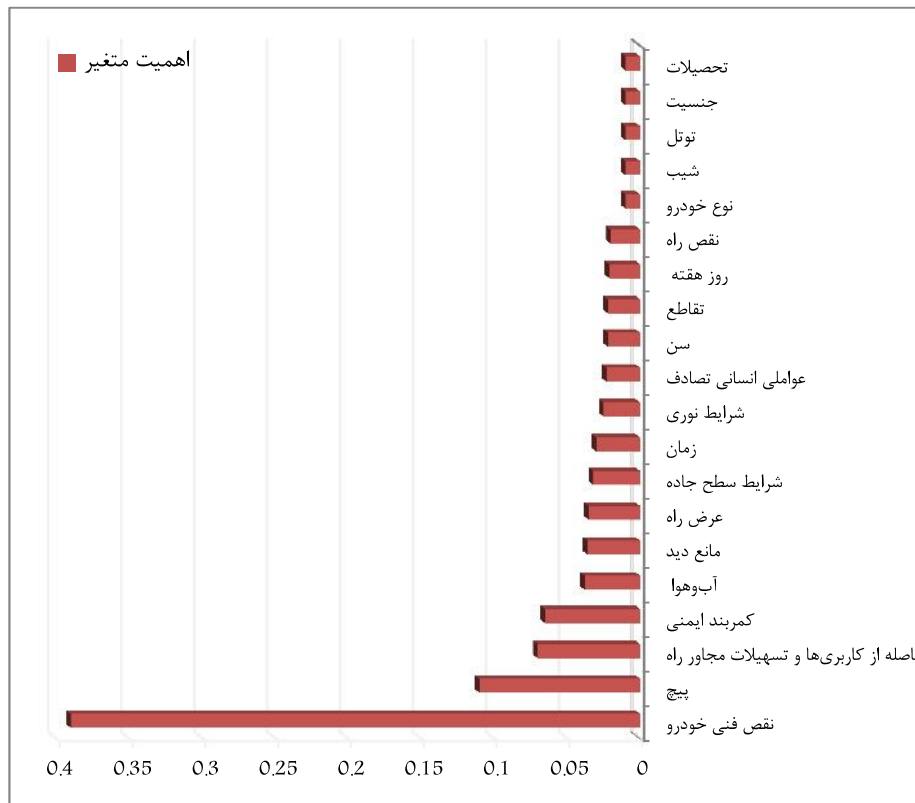
کردن تصادفات براساس استان محل وقوع در محیط GIS (شکل ۷) و تحلیل‌های مکانی و آماری روی هر دسته از تصادفات نشان می‌دهد که ۵۶ درصد از رکوردهای تصادفاتی که در دوره چهارساله این مطالعه، در محدوده مکانی و مرزهای جغرافیای استان گیلان، در محور مورد مطالعه رخ داده‌اند، از نوع جرحی یا فوتی بودند و این عدد برای استان قزوین ۴۲ درصد است. از دیگر فاکتورهای مؤثرگذار در شدت تصادفات در محور مورد مطالعه مانع دید و عرض راه است. موجالی و داونیا^۱ (2011) در مطالعات خود که به روش شبکه‌های بیزین روی داده‌های تصادفات یکی از راه‌های دوخطه برون شهری در جنوب اسپانیا انجام شد، عرض راه را یکی از فاکتورهای مؤثر در شدت تصادفات شناسایی کردند. از آنجاکه بخش اعظمی از محور مورد مطالعه از مناطق کوهستانی عبور می‌کند، عرض کمی دارد و صخره‌های مجاور راه یکی از موانع اصلی دید رانندگان است. رکوردهای تصادفات پیشین محور نشان می‌دهد که افزون بر صخره‌های مجاور، درخت و ساخت‌وسازهای اطراف راه از دیگر عوامل مؤثر در دید راننده و شدت تصادفات به‌شمار می‌روند.

کمربند ایمنی یکی دیگر از متغیرهای مؤثر در تعیین شدت تصادفات است. به این معنا که در صورت استفاده نکردن از کمربند ایمنی، احتمال تصادفات جرحی و فوتی بیشتر خواهد شد. برخی تحقیقات پیشین (AWoremi et al., 2010; Tavakoli Kashani and Shariat Mohaymany, 2011; Morgan and Mannering, 2011) نیز در مطالعات خود به این نتیجه رسیده‌اند که بستن کمربند ایمنی از عوامل مؤثر در کاهش شدت تصادفات است. پنجمین فاکتور مؤثر در شدت تصادفات در محور مورد مطالعه، شرایط آب‌وهوایی شناسایی شده است. مارتین^۱ (2002)، گرتس و همکاران^۲ (۲۰۰۵)، کورتسه و ریتفیلد^۳ (۲۰۰۹)، تئوفیلاتوس و همکاران^۴ (۲۰۱۲)، و چیو و همکاران^۵ (۲۰۱۳) نیز در مطالعات خود بر شدت تصادفات، که بیشتر به روش‌های آماری و رگرسیونی صورت پذیرفته است، شرایط آب‌وهوایی را یکی از عوامل مؤثر در شدت تصادفات بیان کرده‌اند. از آنجاکه بخشی از محور مورد مطالعه از شمال کشور عبور می‌کند، شرایط آب‌وهوایی ممکن است تأثیر بسزایی بر شدت تصادفات در این بخش‌های راه داشته باشد. مجزا



شکل ۷. بررسی شدت تصادفات براساس مرزهای جغرافیایی استان‌ها

- | | |
|------------------------|-----------------------|
| 1. Martin | 2. Geurts et al. |
| 3. Koetse and Rietveld | 4. Theofilatos et al. |
| 5. Chiou et al. | 6. Mujalli and De Oña |



شکل ۸. اهمیت نسبی متغیرها در تحلیل شدت تصادفات

۴- نتیجه‌گیری

مدل‌های درخت دسته‌بندی و رگرسیون فازی که از تلفیق مفاهیم فازی و مدل CART ایجاد می‌شوند، از مزایای استنتاج فازی و درخت تصمیم بهره می‌برند و با کمترین حجم محاسبات، مدل‌هایی ساده، دقیق و تفسیرشدنی را برای کلاسه‌بندی و پیش‌بینی ارائه می‌دهند. در این مطالعه، با تلفیق تحلیل‌های مکانی و داده‌کاوی به روش درخت دسته‌بندی و رگرسیون فازی، شدت تصادفات تحلیل و فاکتورهای مکانی مؤثر در آن، در محور قزوین- رشت تعیین شد. نتایج این مطالعه نشان داد که مدل درخت دسته‌بندی و رگرسیون فازی درمقایسه با مدل CART، شدت تصادفات به‌ویژه تصادفات فوتی را با صحت بیشتری پیش‌بینی می‌کند. در مدل درخت تصمیم فازی پیشنهادی، داده‌های آموزشی پیش از فرایند استنتاج، با تعریف توابع عضویت مناسب فازی می‌شوند. بنابراین، برای

در پایان، شایان ذکر است که خروجی روش درخت تصمیم فازی پیشنهادی در تعیین فاکتورهای مؤثر در شدت تصادفات تحت تأثیر ویژگی‌های محور مورد مطالعه و داده‌های ورودی است. نتایج شکل ۸ در بیان اهمیت نسبی متغیرهای ورودی، حاصل پیاده‌سازی روش پیشنهادی روی داده‌های مکانی و غیرمکانی مرتبط با تصادفات در محور قزوین- رشت است. بنابراین، با در اختیار داشتن رکوردهای تصادفات و داده‌های مکانی دیگر راه‌های برون‌شهری کشور، روش پیشنهادی برای تحلیل شدت تصادفات و تعیین فاکتورهای مؤثر در آن راه‌ها قابل استفاده است. این نتایج به متخصصان ایمنی راه کمک می‌کند تا عوامل تأثیرگذار از سطوح گوناگون شدت تصادف را شناسایی کنند و اقدامات پیشگیرانه لازم را برای کاهش شدت یا جلوگیری از وقوع تصادفات در بخش‌های گوناگون راه انجام دهند.

۶- منابع

گزارش سازمان بهداشت جهانی، ۲۰۱۳، بازیابی:
۱۹ تیر ۱۳۹۴ از وبسایت:

http://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en/

Ayati, E., 2012, **The Cost of Interurban and Intercity Accidents**, Transportation Research Institute report, the University of Science and Technology.

Aworemi, J.R., Segun, A.I. & Oluwaseun, O., 2010, **Analytical Study of the Causal Factors of Road Traffic Crashes in Southwestern Nigeria**, Educational Research, Vol.1, No.4, pp. 118-124.

Bhall, K., Naghavi, M., Shahraz, S., Bartels, D. & Murray, C.J.L., 2009, **Building National Estimates of the Burden of Road Traffic Injuries in Developing Countries from All Available Data Sources: Iran**, Injury Prevention, Vol.15, No. 3, 150-156.

Clarke, R., Forsyth, R. & Wright, R., 1998, **Machine Learning in Road Accident Research: Decision Trees Describing Road-Accidents During Cross-Flow Turns**, Ergonomics, Vol.41, No.7, pp.1060-1079.

Chang, L. & Chen W., 2005, **Data Mining of Tree-Based Models to Analyze Freeway Accident Frequency**, Safety Research, Vol.36, No.4, pp.365-375.

Chiou, Y.C., Lan, L.W. & Chen, W.P., 2013, **A Two-Stage Mining Framework to Explore Key Risk Conditions on One-Vehicle Crash Severity**, Accident Analysis & Prevention, Vol.50, pp. 405-415.

Delen, D., Sharda, R. & Bessonov, M., 2006, **Identifying Significant Predictors of Injury Severity in Traffic Accidents Using a Series of Artificial Neural Networks**, Accident Analysis & Prevention, Vol.38, No. 3, pp. 434-444.

پیش‌بینی شدت تصادف، چندین الگو از درخت تصمیم اجرایی‌اند و رویداد تصادف در کلاس الگویی از شدت تصادف قرار می‌گیرد که بیشترین سازگاری را با داده ورودی داشته باشد. افزون بر این، مدل پیشنهادی با استفاده از مجموعه‌های فازی و تعریف مقادیر زبانی، به نظر متخصصان در تعریف بازه‌ها و تقسیم‌بندی‌های معیارهای ورودی توجه می‌کند و صحت فرایند پیش‌بینی را افزایش می‌دهد.

در این مطالعه تحلیل حساسیت برای تعیین اهمیت نسبی متغیرهای ورودی در شدت تصادفات صورت پذیرفت. نتایج تحلیل‌ها بیانگر این است که پیچ و فاصله از کاربری‌ها و تسهیلات اطراف راه مهم‌ترین فاکتورهای مکانی تأثیرگذار در افزایش شدت تصادفات به‌شمار می‌روند. افزون بر فاکتورهای مکانی، نقص فنی خودرو، کمربند ایمنی، آب‌وهوا و مانع دید از دیگر فاکتورهای مؤثر در شدت تصادفات در محور مورد مطالعه‌اند. این مطالعه به متخصصان ایمنی راه کمک می‌کند تا عوامل تأثیرگذار در سطوح مختلف شدت تصادف را شناسایی کنند و اقدامات پیشگیرانه لازم را برای کاهش شدت یا جلوگیری از وقوع تصادفات در بخش‌های گوناگون راه انجام دهند.

در این مقاله، فاکتورهای مکانی مؤثر در شدت تصادفات در راه‌های برون‌شهری بررسی شدند. در تحقیقات آتی می‌شود روش پیشنهادی را می‌توان از تحلیل فاکتورهای مکانی مؤثر در شدت تصادفات راه‌های درون‌شهری نیز استفاده کرد.

۵- سپاسگزاری

بدین‌وسیله نویسندگان مقاله مراتب قدردانی خود را از سازمان نقشه‌برداری کشور، پلیس راه، سازمان راهداری و حمل‌ونقل جاده‌ای و پژوهشکده حمل‌ونقل دانشگاه علم‌وصنعت، برای در اختیار قرار دادن داده‌های مورد استفاده در این تحقیق اعلام می‌دارند.

- Effati, M., Rajabi, M.A., Samadzadegan, F. & Blais, J. A., 2012, **Developing a Novel Method for Road Hazardous Segment Identification Based on Fuzzy Reasoning and GIS**, Journal of Transportation Technologies, Vol.2, pp.32-40.
- Erdogan, S., Yilmaz, I., Baybura, T. & Gullu, M., 2008, **Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar**, Accident Analysis and Prevention, Vol.40, No.1, pp.174-181.
- Engelbrecht, A.P., 2007, **Computational Intelligence: An Introduction**, Second Edition, John Wiley & Sons, Ltd.
- Flahaut, B., Mouchart, M., Martin, E.S. & Thomas, I., 2002, **The Local Spatial Autocorrelation and Kernel Method for Identifying Black Zones a Comparative Approach**, Accident Analysis and Prevention, Vol.35, pp. 991-1004.
- Gundogdu, I.B., 2010, **Applying Linear Analysis Methods to GIS-Supported Procedures for Preventing Traffic Accidents: Case Study of Konya**, Safety Science, Vol. 48, pp.763-769.
- Geurts, K., Thomas I. & Wets, G., 2005, **Understanding Spatial Concentrations of Road Accidents Using Frequent Item Sets**, Accident Analysis and Prevention, 37(4), pp. 787-799.
- Ha, H.H. & Thill, J.C., 2011, **Analysis of Traffic Hazard Intensity: A Spatial Epidemiology Case Study of Urban Pedestrians**, Computers. Environment and Urban Systems, Vol.35, pp. 230-240.
- Janikow, C.Z., 2004, **FID4**, 1: An Overview, IEEE 2, PP. 877-881.
- Jha, M., McCall, C. & Schonfeld, P., 2001, **Using GIS, Genetic Algorithms and Visualization in Highway Development**, Journal of Computer - Aided Civil and Infrastructure Engineering, Vol. 16, No.6, pp. 399 – 414.
- Joshua, S.C. & Garber, N., 1990, **Estimating Truck Accident Rate and Involvements Using Linear and Poisson Regression Models**, Transportation Planning and Technology, Vol .15, pp. 41-58.
- Koetse, M.J. & Rietveld P., 2009, **The Impact of Climate Change and Weather on Transport: An Overview of Empirical Findings**, Transportation Research Part D, Vol. 14, No.3, pp. 205-221.
- Martin, J.L., 2002, **Relationship between Crash Rrate and Hourly Traffic Flow on Interurban Motorways**, Accident Analysis and Prevention, Vol. 34, No. 5, pp. 619 – 629.
- Miaou, S. & Lum, H., 1993, **Modeling Vehicle, Accidents and Highway Geometric Design Relationships**, Accident Analysis and Prevention, Vol. 25, No. 6, pp. 689-709.
- Mitchell, T., 1997, **Machine Learning**, New York: McGraw-Hill.
- Morgan, A. & Mannering, F., 2011, **The Effects of Road-Surface Conditions, Age, and Gender on Driver-Injury Severities**, Accident Analysis and Prevention, Vol. 43(5), pp.1852–1863.
- Mujalli, R.O. & De Oña, J., 2011, **A Method for Simplifying the Analysis of Traffic Accidents Injury Severity on Two-Lane Highways Using Bayesian Networks**, Journal of Safety Research, Vol. 42, pp. 317–326.
- Pakgozar, A., Tabrizi, R.S., Khalilli, M. & Esmacili, A., 2010, **The Role of Human Factor in Incidence and Severity of Road Crashes Based on the CART and LR regression: A Data Mining Approach**, Procedia Computer Science, Vol. 3, pp. 764–769.

- Shanker, V., Mannering, F. & Barfield, W., 1995, **Effect of Roadway Geometric and Environment Factors on Rural Freeway Accident Frequencies**, Accident Analysis and Prevention, Vol. 27, pp. 11-23.
- Siddiqui, C., Abdel-Aty, M. & Huang, H., 2012, **Aggregate Nonparametric Safety Analysis of Traffic Zones**, Accident Analysis and Prevention, Vol. 45, pp. 317-325.
- Sohn, S. & Hyungwon, S., 2001, **Pattern Recognition for a Road Traffic Accident Severity in Korea**, Ergonomics, Vol. 44, No. 1, pp. 101-117.
- Steenberghen, T., Dufays, T., Thomas, I. & Flahaut, B., 2004, **Intra-Urban Location and Clustering of Road Accidents Using GIS: A Belgian Case**, International Journal of Geographic Information Science, Vol. 18, No. 2, pp.169–181.
- Tavakoli Kashani, A. & Shariat Mohaymany, A., 2011, **Analysis of the Traffic Injury Severity on Two Lane, Two-Way Rural Roads Based on Classification Tree Models**, Safety Science, Vol. 49, pp.1314–1320.
- Theofilatos, A., Graham, D. & Yannis, G., 2012, **Factors Affecting Accident Severity Inside and Outside Urban Areas in Greece**, Traffic Injury Prevention, Vol. 13(5), pp. 458-467.
- Umanol, M., Okamoto, H., Hatono, I., Tamura, H., Kawachi, F., Umedzu & Kinoshita, J.S., 1994, **Fuzzy Decision Trees by Fuzzy ID3 Algorithm and Its Application to Diagnosis Systems, In Fuzzy Systems**, IEEE world Congress on Computational Intelligence, Proceedings of the third IEEE conference : pp. 2113–2118.
- Wang J. & Wang, X., 2011, **An Ontology-Based Traffic Accident Risk Mapping Framework**, Advances in Spatial and Temporal Databases, Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg 2011, Vol. 6849/2011, pp. 21-38.
- Wang, X. & Kockelman, K.M., 2007, **Specification and Estimation of a Spatially and Temporally Autocorrelated Seemingly Unrelated Regression Model: Application to Crash Rates in China**, Transportation, Vol. 34(3), pp. 281-300.
- Yamada, I. & Thill, J.C., 2004, **Comparison of Planar and Network K-Functions in Traffic Accident Analysis**, Journal of Transport Geography, Vol. 12, pp. 149-158.
- Yannis, G., Papadimitriou, E., Dupont, E. & Martensen, H., 2010, **Estimation of Fatality and Injury Risk by Means of in-Depth Fatal Accident Investigation Data**, Traffic Injury Prevention, Vol. 11(5), pp. 492-502.
- Yuan, Y. & Shaw, M.J., 1995, **Induction of Fuzzy Decision Trees**, Fuzzy Sets and Systems, Vol. 69, pp. 125–139.