تحلیل و تخمین شدت سوانح ترافیکی شهری با بهره‌گیری از الگوریتم‌های هوشمند

مقاله علمی – پژوهشی

علیرضا طاهریان، دانش آموخته کارشناسی ارشد، دانشکده مهندسی عمران و محیط زیست، گروه مهندسی و برنامه ریزی حمل و نقل، دانشگاه تربیت مدرس، تهران، ایران
سعید شرافتی پور*، استادیار، دانشکده مهندسی عمران و محیط زیست، گروه مهندسی و برنامه ریزی حمل و نقل، دانشگاه تربیت مدرس، تهران، ایران
امیرحسین قیامی، دانش آموخته کارشناسی ارشد، دانشکده مهندسی عمران و محیط زیست، گروه مهندسی و برنامه ریزی حمل و نقل، دانشگاه تربیت مدرس، تهران، ایران
محمود صفارزاده، استاد، دانشکده مهندسی عمران و محیط زیست،گروه مهندسی و برنامه ریزی حمل و نقل، دانشگاه تربیت مدرس، تهران، ایران
*پست الکترونیکی نویسنده مسئول:sherafatipour@modares.ac.ir

دریافت: ۱۴۰۴/۰۴/۰۵– پذیرش: ۱۴۰۴/۰۸/۲۰
صفحه ۱۰۰-۹۱

چکیده

تصادفات درون‌شهری یکی از چالش‌های جدی در ایران به شمار می‌رود که به دلایلی مانند ضعف زیرساخت‌ها، رعایت‌نشدن قوانین، و نبود فرهنگ مناسب رانندگی، تلفات و خسارات زیادی به همراه داشته است. در مقابل، بسیاری از کشورهای پیشرفته با استفاده از فناوری، آموزش، و برنامه‌ریزی هوشمند شهری توانسته‌اند میزان این تصادفات را به‌طور چشم‌گیری کاهش دهند. تجربه موفق این کشورها نشان می‌دهد که مدیریت مؤثر ترافیک نقش مهمی در افزایش ایمنی شهری دارد. پژوهش حاضر با هدف شناسایی و پیش‌بینی شدت تصادفات معابر درون‌شهری استان البرز، از الگوریتم درخت تصمیم بهره گرفته است. برای این منظور، داده‌های مربوط به ۵۰۰۰ مورد تصادف ثبت‌شده در بازه زمانی سال‌های ۱۳۹۷ تا ۱۴۰۱ گردآوری و پس از پاک‌سازی، در محیط برنامه‌نویسی پایتون مورد تحلیل قرار گرفت. نتایج مدل نشان داد که سه متغیر «نحوه برخورد وسایل نقلیه»، «ساعت وقوع تصادف» و «روزهای غیرکاری هفته» به‌ترتیب بیشترین تأثیر را در پیش‌بینی شدت تصادف دارند. مدل نهایی با دقت ۰,۷۹ توانسته است تفکیک نسبتاً دقیقی میان تصادفات جرحی و فوتی ارائه دهد. یافته‌ها نشان می‌دهد که علاوه بر عوامل انسانی، برخی مؤلفه‌های محیطی و ساختاری نظیر نوع برخورد بین وسایل نقلیه نقش مؤثری در بالا رفتن احتمال وقوع تصادف‌های منجر به فوت ایفا می‌کنند. این پژوهش با بهره‌گیری از داده‌های واقعی و مدل‌سازی تفسیرپذیر، می‌تواند چشم‌اندازی کاربردی برای سیاست‌گذاری‌های ایمنی راه و طراحی مداخلات مؤثر فراهم آورد.

واژه‌های کلیدی: شدت تصادف، درخت تصمیم، تصادفات درون‌شهری، ایمنی ترافیک

۱-مقدمه

وضعیت تصادفات درون‌شهری یکی از مسائل مهم در حوزه حمل‌ونقل و ایمنی شهری در بسیاری از کشورهای جهان، از جمله ایران، به شمار می‌رود. در ایران، آمار تصادفات درون‌شهری همچنان بالاست و به دلایلی مانند نبود زیرساخت های مناسب، رعایت‌نکردن قوانین راهنمایی و رانندگی، ضعف در فرهنگ ترافیکی و کمبود نظارت مؤثر، خسارات جانی و مالی قابل توجهی به شهروندان وارد می شود. در مقایسه، بسیاری از کشورهای توسعه‌یافته با به‌کارگیری

فناوری‌های نوین، آموزش‌های عمومی، اصلاح قوانین، و بهبود طراحی شهری توانسته‌اند میزان تصادفات درون‌شهری را کاهش دهند. کشورهایی مانند سوئد، هلند و ژاپن با اجرای طرح‌هایی چون دیدگاه صفر[1] توانسته‌اند به طرز چشم‌گیری آمار تصادفات مرگ‌بار را کاهش دهند. این تفاوت‌ها نشان می‌دهد که مدیریت هوشمند و همه‌جانبه در حوزه ترافیک شهری می‌تواند نقش کلیدی در افزایش ایمنی و کاهش سوانح ایفا کند.

در پژوهش‌های پیشین، مدل‌سازی شدت تصادفات عمدتاً در دو چارچوب کلی انجام شده است: مدل‌های آماری سنتی و روش‌های یادگیری ماشین. در رویکرد آماری، معمولاً از انواع مدل‌های رگرسیونی (مثلاً لجستیک چندجمله‌ای یا ترتیبی) برای شناسایی عوامل تعیین‌کننده شدت تصادف استفاده می‌شود. مطالعات متعددی مانند ابراری واجاری و همکاران در استرالیا با مدل لجستیک چندجمله‌ای شدت تصادفات موتور سیکلت در تقاطع‌های شهری را بررسی کرده‌اند (Abrari Vajari et al. 2020)، کاپلان و پرتّو عوامل مؤثر بر شدت تصادفات اتوبوس در آمریکا را با مدل لجستیک ترتیبی تحلیل نموده‌اند (Kaplan and Prato 2012)، و یوان و همکاران به بررسی تأثیر متغیرهای مرتبط با تصادفات کشنده کامیون‌ها برای گروه‌های مختلف رانندگان پرداختند(Yuan et al. 2021). اگرچه این مدل‌های آماری تفسیرپذیری خوبی دارند، اما نیاز به فرضیات مشخصی در مورد توزیع داده‌ها و ساختار روابط دارند که گاهی در عمل برقرار نیست.

رویکردهای مبتنی بر یادگیری ماشین با پدید آمدن حجم بالای داده‌ها[2] و پیشرفت تکنیک‌های داده‌کاوی مورد توجه قرار گرفته‌اند. نتایج یک بررسی مروری نشان می‌دهد که الگوریتم‌های جنگل تصادفی[3]، ماشین بردار پشتیبان[4] و درخت تصمیم[5] معمولاً دقت بالایی در پیش‌بینی شدت تصادفات دارند. به همین جهت، بسیاری از پژوهش‌ها از این روش‌ها برای مدلسازی شدت تصادفات استفاده کرده‌اند (Santos, Dias, and Amado 2022). برای نمونه، برخی مطالعات عناصر محیطی-ترافیکی را نیز وارد مدل سازی کرده‌اند: آرون و همکاران با بررسی داده‌های یک بزرگراه شهری در فرانسه رابطه بین بارندگی و شاخص‌های ترافیکی را در وقوع حوادث سنجیدند(Aron et al. 2015)، و هیودو و تادوروکی با تحلیل آماری تصادفات برخورد از عقب در یک آزادراه شهری نشان دادند که وضعیت ترافیک و شدت تصادف به هم وابسته‌اند(Hyodo and Todoroki 2018). علاوه بر آن، پژوهش‌هایی متنوع روی تصادفات شهری نیز انجام شده است. گارسیا و همکاران با تمرکز بر شهر مادرید نقش خدمات حمل‌ونقل اشتراکی (Uber و Cabify) را در بروز تصادفات شهری بررسی کردند و نشان دادند که این خدمات با کاهش تصادفات شهری مرتبط بوده است(García et al. 2021). کولاگراندیا نیز روشی چندجانبه بر مبنای تحلیل فضایی داده‌های تصادف و داده‌کاوی ارائه کرده است که ضمن شناسایی نقاط حادثه‌خیز (نقاط سیاه[6])، تنظیم مداخلات اصلاحی و اولویت‌بندی آنها را بر اساس محدودیت بودجه ممکن می‌سازد (Colagrande 2022). الیان و آنجلیکا نیز با به‌کارگیری مدل‌های پواسون تو در تو و مدل‌های دوجمله‌ای منفی در دو محدوده صنعتی شهری نشان دادند که در یک منطقه افزایش تردد اتوبوس‌های تندرو و اتوبوس‌های بزرگ با افزایش تصادفات همراه است و در منطقه دیگر بارش باران تاثیر مثبت و معنی‌داری بر شدت حوادث داشته است (Eliane and Angelica 2024).

ونگ و همکاران در مطالعه‌ای شدت تصادفات جاده‌ای در چین را با استفاده از داده‌های سرشماری بررسی کرده و نشان می‌دهد که با وجود کاهش کلی تلفات، شدت حوادث تا سال ۲۰۱۵ افزایش داشته است. تصادفات در مناطق روستایی و برخورد با عابران و موتورسواران شدیدتر هستند. عوامل محیطی، نوع وسیله نقلیه و محل حادثه تأثیر قابل‌توجهی دارند (Wang et al. 2019). یون با استفاده از مدل XGBoost[7] و روش تفسیرپذیر LIME به شناسایی مناطق پرخطر برای تصادفات شدید عابران پیاده در سئول پرداخته است. داده‌های سال‌های ۲۰۱۶ تا ۲۰۱۸ تحلیل شده و مناطقی مانند ایستگاه‌های Gangnam و Cheongnyangni به‌عنوان نقاط بحرانی شناسایی شده‌اند. خروجی مدل به شکل نقشه‌های حرارتی برای تصمیم‌گیری سیاستی ارائه شده است(Yoon 2025). اسماعیل‌زاده و دلاور با استفاده از روش چگالی کرنل شبکه‌ای (NKDE) و آمار مکانی

به شناسایی و اولویت‌بندی آماری نقاط پرتصادف در شبکه خیابانی تهران پرداخته است. نتایج نشان می‌دهد که تصادفات دارای الگوی خوشه‌ای معنادار هستند و استفاده ترکیبی از این روش‌ها، شناسایی دقیق‌تری از مناطق پرخطر ارائه می‌دهد. این تحلیل به ارتقاء ایمنی و پایداری شهری کمک می‌کند(EslamiNezhad and Delavar 2019). رضاپور و همکاران با بهره‌گیری از الگوریتم‌هایی چون جنگل تصادفی، ماشین بردار پشتیبان و رگرسیون تطبیقی چندمتغیره به پیش‌بینی شدت جراحات در تصادفات موتورسیکلت در ایالت وایومینگ آمریکا پرداخته است. مدل جنگل تصادفی بهترین عملکرد را با کمترین نرخ خطا و بالاترین AUC داشته است. عواملی مانند سرعت، ترافیک سنگین، مصرف الکل و سن راکب از مهم‌ترین متغیرهای مؤثر شناسایی شدند(Rezapour et al. 2021).

۲- پیشینه تحقیق

در جمع‌بندی مطالعات پیشین، روشن می‌شود که گرایش به کاربرد مدل‌های یادگیری ماشین در مدلسازی شدت تصادفات در حال گسترش است؛ این مدل‌ها به دلیل توانایی در مدیریت حجم بالای متغیرها و رابطه‌های پیچیده، معمولاً در موارد بسیاری دقت بالاتری نسبت به مدل‌های آماری سنتی داشته‌اند. همچنین در مواردی با بهره‌گیری از ترکیبی از روش‌های آماری و یادگیری ماشین به پیش‌بینی شدت تصادفات پرداخته‌اند که در مجموع نشان‌دهنده اهمیت انواع متغیرهای محیطی، جاده‌ای و رفتاری در بروز تصادفات با شدت بالا است. تاکنون تعداد کمی از این تحقیقات به طور مشخص تصادفات درون‌شهری پرداخته‌اند. لذا خلاء عمده‌ای در زمینه تحلیل داده‌های تصادفات معابر شهری وجود دارد.

۳- داده

در این مطالعه از حدود ۵ هزار داده‌ی مربوط به تصادفات درون شهری استان کرج طی سال‌های ۱۳۹۷ تا ۱٤۰۱ استفاده شد. پس از پاکسازی داده‌ها و حذف رکوردهای با مقادیر نامشخص، نزدیک به ٤ هزار داده باقی ماند. خالص‌های از متغیرها و دسته‌بندی‌های هریک در جدول ۱ آورده شده‌است. این داده‌ها مربوط به شرایط محیطی، وسیله نقلیه و ویژگی‌های راننده‌ی دخیل در وقوع تصادف می‌باشد. متغیر تصمیم به دو دسته‌ی تصادفات فوتی و تصادفات جرحی تقسیم شده‌است.

٤- روش‌شناسی

برای دستیابی به اهداف این پژوهش، از الگوریتم درخت تصمیم به عنوان یکی از تکنیک‌های یادگیری ماشین استفاده شد. فرایند مدل‌سازی در محیط برنامه‌نویسی پایتون پیاده‌سازی گردید. درخت تصمیم یکی از روش‌های پرکاربرد در یادگیری نظارت‌شده محسوب می‌شود که به‌ویژه برای مسائل طبقه‌بندی و شناسایی الگوهای پنهان در داده‌ها عملکرد مناسبی دارد. مزیت اصلی این روش در قابلیت بالای تفسیر بصری و امکان به‌کارگیری آن بر مجموعه داده‌های بزرگ نهفته است. پژوهش‌های پیشین نیز تأیید کرده‌اند که درخت تصمیم از جمله دقیق‌ترین الگوریتم‌ها در زمینه پیش‌بینی شدت تصادفات می‌باشد (Santos, Dias, and Amado 2022).

جدول ۱. توضیح و دسته‌بندی متغیرها

ردیف	نام متغیر	توضیح متغیر
۱	شدت تصادف(متغیر هدف)	۱) جرحی ۲) فوتی
۲	ماه وقوع تصادف	۱) فروردین ۲) اردیبهشت ۳)خرداد ٤)تیر ۵)مرداد٦)شهریور ۷)مهر ۸) آبان ۹)آذر ۱۰)دی ۱۱)بهمن ۱۲)اسفند
۳	ساعت وقوع تصادف	متغیر پیوسته
٤	روز وقوع تصادف	۱)روزهای کاری(شنبه، یکشنبه، دوشنبه، سه شنبه، چهارشنبه) ۲) روزهای غیرکاری(پنج شنبه و جمعه)
۵	روشنایی معبر	۱) روز ۲) شب ۳) طلوع ٤) غروب
٦	وضعیت معبر	۱) بلوار۲) جاده۳) جاده ای٤) خیابان فرعی ۵) دو طرفه٦) غیرمسکونی۷) مسکونی۸) میدان۹) یک طرفه
۷	نحوه برخورد	۱) جلو به جلو۲) جلو به عقب۳) جلو به پهلو٤) جلو به پهلو ۵) سایر٦) عقب به پهلو ۷) پهلو به پهلو
۸	شرایط راه	۱) فقدان خط کشی۲) بی توجهی در حرکت با دنده عقب۳) در حال تعمیر٤) سایر ۵) نقص روشنایی راه٦) نقض ماده۲۱۲ ۷) هموار و سالم ۸) کم بودن عرض راه
۹	شرایط جوی	۱) ابری۲) بارانی۳) برفی و کولاک٤) صاف ۵) مه آلود
۱۰	شرایط معبر	۱) بارانی۲) خشک۳) سایر ٤) لغزنده ۵) یخبندان و برفی

مکانیزم عملکرد مدل بدین صورت است که ابتدا تمامی داده‌ها در گره ابتدایی (گره ریشه) قرار می‌گیرند. سپس بر اساس معیاری نظیر ضریب جینی یا آنتروپی، متغیر بهینه جهت تفکیک داده‌ها انتخاب می‌شود. متغیر منتخب، داده‌ها را به زیرگروه‌هایی با توجه به مقادیر آن تقسیم می‌کند. چنانچه گره‌های نهایی حاصل، دارای خلوص کافی از منظر شاخص‌های ارزیابی باشند، به عنوان برگ‌های نهایی در نظر گرفته می‌شوند؛ در غیر این صورت، روند تقسیم‌بندی بر اساس سایر متغیرهای موجود ادامه می‌یابد.

از دیگر ویژگی‌های قابل توجه این روش، قابلیت هرس‌کردن درخت برای جلوگیری از بیش‌برازش[۸] مدل است. ضریب جینی که در الگوریتم CART نیز به‌کار گرفته می‌شود، معیاری برای اندازه‌گیری ناخالصی در تقسیم‌بندی گره‌ها است و برای تمامی تقسیمات ممکن محاسبه می‌شود؛ سپس شکستی انتخاب می‌گردد که کمترین مقدار Gini را داشته باشد. در این میان، P نشان‌دهنده احتمال تعلق داده به هر یک از دسته‌های طبقه‌بندی‌شده است.

$$Gini(t) = 1 - \sum_{i=1}^{c} p_i^2 \qquad (۱)$$

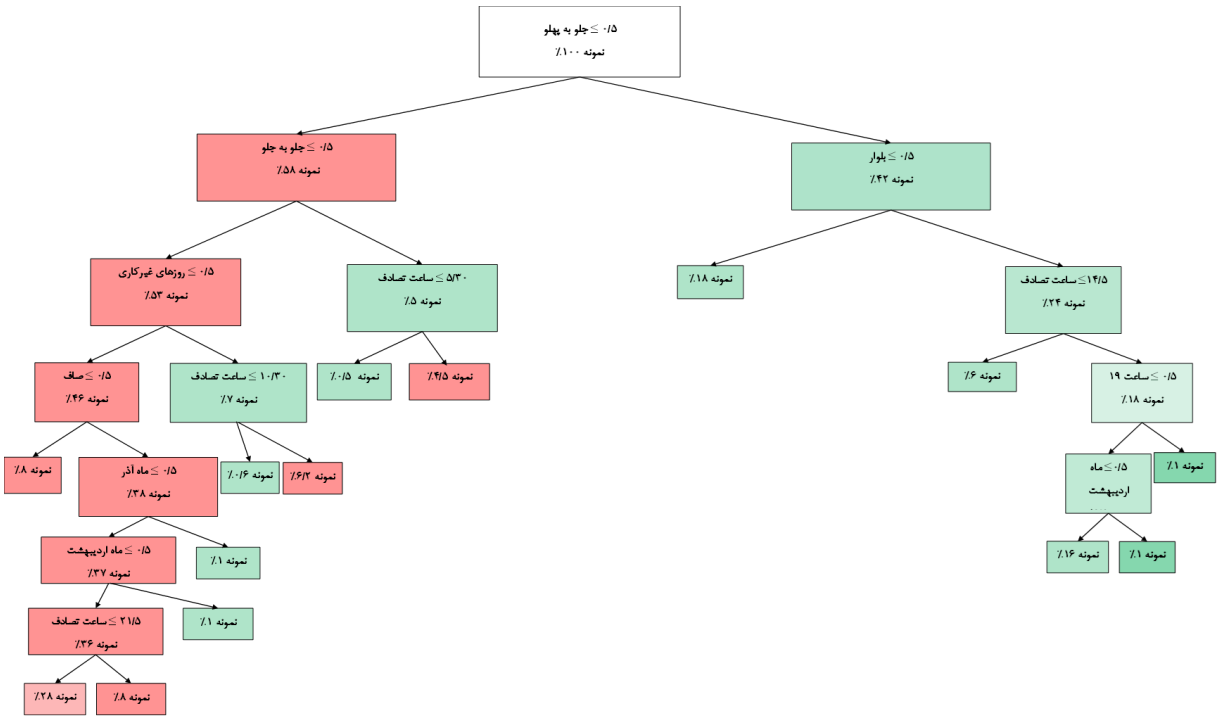
مدل درخت تصمیم ارائه‌شده در این پژوهش، با هدف پیش‌بینی شدت تصادفات درون‌شهری استان البرز، ساختاری سلسله‌مراتبی از تصمیمات را بر اساس متغیرهای مؤثر بر شدت حادثه ترسیم کرده است. این مدل با انتخاب متغیر «نحوه برخورد» در جایگاه گره ریشه، به‌درستی بر اهمیت نوع تصادف در تعیین شدت آن تأکید دارد. از میان تمامی برخوردها، «برخورد جلو به پهلو» به عنوان نخستین فیلتر دسته‌بندی در نظر گرفته شده که نشان‌دهنده نقش بحرانی این نوع برخورد در بروز تصادفات فوتی است. در گام بعدی، در مواردی که برخورد جلو به پهلو رخ نداده است، مسیر تصمیم‌گیری با متغیر «برخورد جلو به جلو» ادامه یافته که آن نیز با افزایش احتمال تصادف فوتی همراه بوده است. مدل در ادامه به سراغ متغیر «ساعت وقوع تصادف» می‌رود، که نقش تعیین‌کننده‌ای در شدت حادثه دارد؛ به‌گونه‌ای که تصادفات رخ‌داده در ساعات ابتدایی صبح، به‌ویژه قبل از ساعت ۵:۳۰، احتمالاً با شدت بیشتری همراه هستند. این یافته می‌تواند به شرایط خاص رانندگی در ساعات شب، کاهش دید، یا خواب‌آلودگی رانندگان نسبت داده شود. در مسیر دیگر درخت، زمانی که برخورد جلو به پهلو وجود دارد، مدل متغیر «روزهای غیرکاری» را بررسی می‌کند

و نشان می‌دهد که تصادف در روزهای پنج‌شنبه و جمعه (روزهای تعطیل) بیشتر به تصادف‌های فوتی ختم می‌شود. این موضوع احتمالاً ناشی از افزایش سفرهای برون‌شهری یا افزایش سرعت در مسیرهای خلوت‌تر در این روزها است. در گام‌های بعدی، مدل به سراغ عواملی چون «شرایط جوی» رفته و تصریح می‌کند که هوای صاف در ترکیب با روز کاری یا غیرکاری می‌تواند الگوی پیش‌بینی شدت تصادف را تغییر دهد. اگر هوا صاف نباشد، احتمال فوتی بودن حادثه بالا می‌رود؛ در حالی‌که در آب و هوای صاف، بسته به ماه تصادف (مثلاً آذرماه) یا شرایط روشنایی معبر، مدل تصمیم‌گیری را در مسیرهای متفاوتی هدایت می‌کند. اهمیت ماه وقوع تصادف، مانند آذر، نیز حاکی از نقش فصلی و اقلیمی در بروز تصادفات شدیدتر است، که ممکن است به علت لغزندگی، بارش یا کاهش دید در فصول سرد باشد.

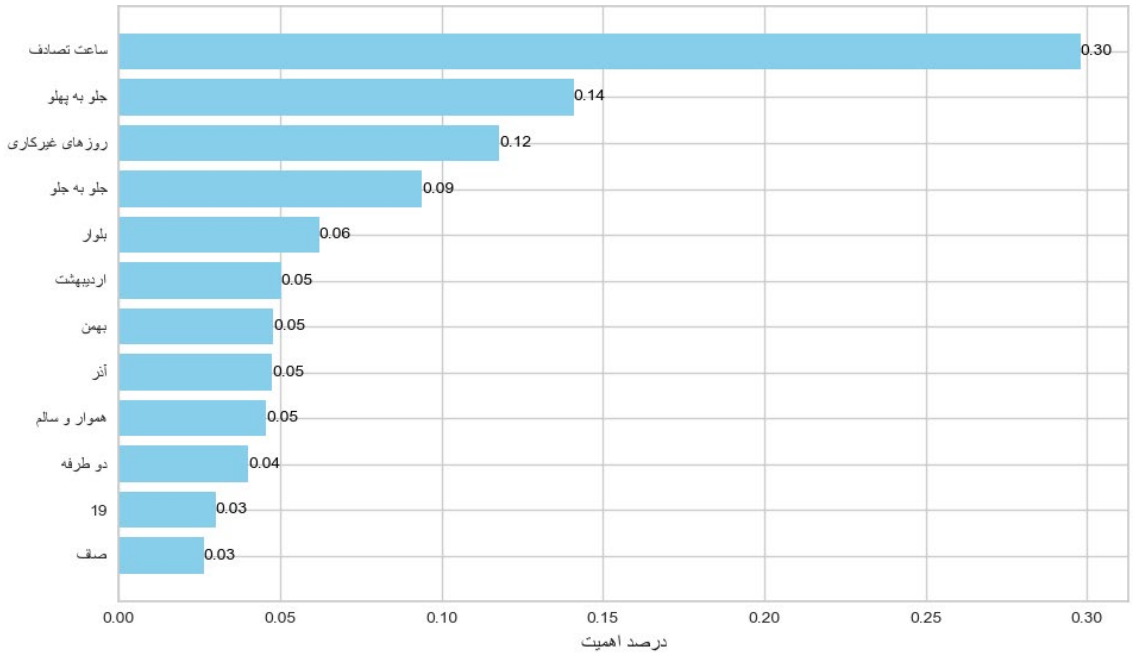
در سطوح پایین‌تر درخت، متغیرهایی همچون «روشنایی معبر»، «میزان دریافت نور»، و «ساعت‌های خاصی از روز» نیز وارد فرآیند تصمیم‌گیری شده‌اند، که نشان می‌دهد شدت تصادف به‌طور پیچیده‌ای تابع تعامل بین شرایط محیطی و زمانی است. به‌عنوان مثال، تغییر شدت تصادف در بازه‌های زمانی ۱۰:۰۰ تا ۱۸:۰۰ عصر در روزهای کاری یا غیرکاری، همراه با تغییرات شرایط نوری و جوی، نشان می‌دهد که ریسک‌پذیری در ساعات مختلف روز، بسته به الگوهای ترافیکی و میزان رفت‌وآمد نیز دگرگون می‌شود.

مدل همچنین نشان داده که در برخی موارد، در ترکیب خاصی از متغیرها (مثلاً برخورد جلو به جلو، در ساعات قبل از طلوع)، دقت پیش‌بینی وقوع تصادف‌های فوتی بسیار بالا بوده است؛ به‌گونه‌ای که احتمال فوتی بودن به بیش از ۹۰ درصد می‌رسد. در مقابل، در شرایطی مثل تصادف در ماه آذر با هوای صاف و نور روز، احتمال وقوع تصادف‌های جرحی بیشتر است.

به‌طور کلی، این درخت تصمیم به‌خوبی توانسته مجموعه‌ای از متغیرهای تأثیرگذار شامل نوع برخورد، ساعت وقوع، روز هفته، شرایط جوی، ماه وقوع، و روشنایی را با منطقی سلسله‌مراتبی در کنار هم قرار دهد و از میان هزاران ترکیب ممکن، مسیرهایی با بیشترین خلوص دسته‌بندی را استخراج کند. نتیجه کلی تحلیل نشان می‌دهد که شدت تصادف عمدتاً تابعی از دو محور اصلی است: نخست نوع برخورد وسایل نقلیه (برخورد مستقیم با سمت حساس خودرو)، و دوم زمان‌بندی وقوع تصادف (اعم از ساعت و روز). ترکیب این عوامل با شرایط محیطی و فصلی، امکان پیش‌بینی دقیق‌تر و طراحی مداخلات ایمنی هدفمند را فراهم می‌سازد. این مدل می‌تواند مبنای تحلیل‌های سیاست‌گذارانه برای بهینه‌سازی زیرساخت‌های شهری، برنامه‌ریزی آموزش‌های ایمنی در ایام خاص، یا حتی پیاده‌سازی سیستم‌های هشدار هوشمند برای رانندگان در شرایط پرریسک باشد.



شکل ۱. مدل درخت تصمیم



شکل ۲. ضرایب اهمیت متغیرها در مدل درخت تصمیم

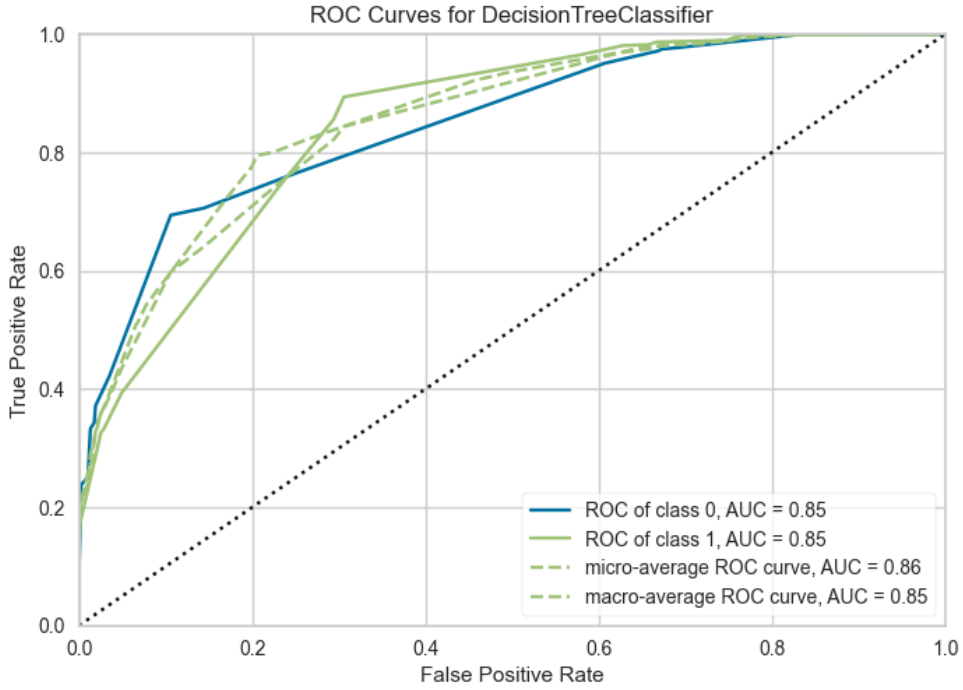
با توجه به شکل۲ ساعت تصادف مهمترین متغیر در مدل درخت تصمیم بوده است که نشان دهنده این است که ساعت وقوع تصادف اهمیت بسیار بالایی برای نوع تصادف است که آیا منجر به فوتی یا جرحی خواهد شد در ادامه متغیر نحوه برخورد از نوع تصادف جلو به پهلو اهمیت بیشتری دارد به این صورت که اگر تصادف به این صورت باشد که وسیله نقلیه از جلو به پهلو وسیله نقلیه اصابت کند احتمال زیاد منجر به فوتی خواهد شد. همچنین اهمیت متغیر روزهای غیرکاری در جایگاه سوم به این صورت است که نشان میدهد اکثر تصادف‌های منجر به فوتی در روزهای غیرکاری یعنی آخر هفته‌ها اتفاق خواهد افتاد یعنی زمانی که افراد در تعطیلات هستند و چهارمین متغیر مورد اهمیت نحوه برخورد از نوع جلو به جلو است به این صورت که در تمامی نوع‌های نحوه برخورد وسیله نقلیه با یکدیگر بعد از اصابت جلو به پهلو بعد از اصابت جلو به جلو باعث فوتی خواهد شد که کاملاً منطقی است با توجه به این نمودار و اهمیت‌های متغیرها به این نتیجه میتوان رسید که نحوه برخورد وسیله‌های نقلیه با یکدیگر از اهمیت خیلی زیادی در تعیین شدت تصادف خواهد داشت.

نتایج اعتبارسنجی مدل درخت تصمیم در شکل۳ نشان داده شده است. با توجه به آن مساحت زیر نمودار کلاس صفر (تصادف منجر به جرحی) و مساحت زیر نمودار کلاس ۱ (تصادف منجر به فوتی) برابر با ۰/۸۵ است و مساحت زیرنمودار درخت تصمیم برابر است با ۰/۸۶ و دقت مدل برای پیش بینی برابر ۰/۷۹ است. ماتریس درهم‌ریختگی ارائه‌شده، نشان‌دهنده عملکرد مدل درخت تصمیم در طبقه‌بندی شدت تصادفات به دو دسته اصلی، یعنی تصادف‌های جرحی (کلاس صفر) و فوتی (کلاس ۱) است. این ماتریس در قالب درصد، دقت مدل را در پیش‌بینی صحیح و نادرست هر کلاس نشان می‌دهد و به‌طور مشخص می‌توان گفت که مدل توانسته عملکرد نسبتاً خوبی در تشخیص موارد فوتی ارائه دهد. بر اساس این تصویر، مدل ۶۹ درصد از تصادف‌های واقعاً جرحی را به‌درستی به عنوان کلاس صفر شناسایی کرده و در مقابل، ۳۱ درصد از این تصادف‌ها را به اشتباه به عنوان تصادف فوتی طبقه‌بندی کرده است. این میزان خطای نوع اول، یعنی پیش‌بینی اشتباه جرحی به‌عنوان فوتی، می‌تواند منجر به هشدارهای کاذب شود، اما در سناریوهای مرتبط با ایمنی ترافیک، این نوع اشتباه معمولاً خطر کمتری نسبت به خطای عکس آن دارد.

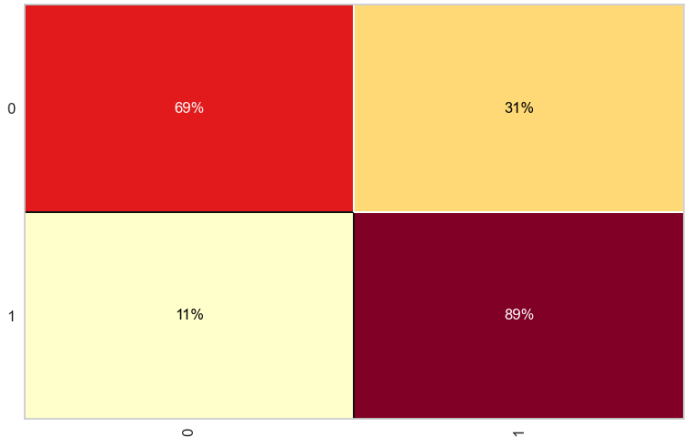
در مقابل، عملکرد مدل در شناسایی تصادف‌های فوتی بسیار مطلوب‌تر است. مدل موفق شده ۸۹ درصد از تصادف‌های واقعاً فوتی را به‌درستی شناسایی کند و تنها در ۱۱ درصد موارد، این حوادث را به اشتباه جرحی در نظر گرفته است.

این یعنی خطای نوع دوم (false negative) در این مدل نسبتاً پایین است که برای کاربردهای ایمنی‌محور از اهمیت فوق‌العاده‌ای برخوردار است؛ چراکه عدم شناسایی یک تصادف مرگبار می‌تواند پیامدهای جدی و خطرناکی به دنبال داشته باشد. به بیان دیگر، اگرچه مدل در برخی موارد تصادف‌های کم‌خطرتر (جرحی) را بیش از حد جدی تلقی کرده، ولی در مقابل از نادیده گرفتن تصادف‌های شدیدتر تا حد زیادی جلوگیری کرده است.

با توجه به این تحلیل، می‌توان نتیجه گرفت که مدل درخت تصمیم از منظر تفکیک بین دو نوع شدت تصادف عملکرد متعادلی داشته، اما به‌طور خاص در تشخیص صحیح تصادف‌های منجر به فوت قوی‌تر عمل کرده است.



شکل ۳. منحنی ROC در مدل درخت تصمیم



شکل ۱. ماتریس درهم ریختگی درخت تصمیم

۵- نتیجه‌گیری

در این پژوهش، با هدف تحلیل شدت تصادفات درون‌شهری استان البرز، از الگوریتم درخت تصمیم بهره گرفته شده و با تکیه بر داده‌های واقعی تصادفات شهری، مدلی توسعه یافته که با دقت ۷۹ درصد توانسته است به‌صورت قابل‌قبول و قابل اتکا، تفاوت میان تصادف‌های جرحی و فوتی را شناسایی و پیش‌بینی کند. این سطح از دقت بیانگر توانمندی بالای مدل در استخراج الگوهای مؤثر و شناسایی متغیرهایی است که در شدت پیامدهای تصادف نقش کلیدی دارند.

تحلیل ساختار درخت تصمیم نشان می‌دهد که متغیر «نحوه برخورد و سایل نقلیه» مهم‌ترین عامل در تعیین شدت تصادفات بوده است. الگوهایی که در آن‌ها نوع برخورد به صورت مستقیم به بخش جلویی خودرو رخ می‌دهد—از جمله برخورد های «جلو به جلو» و «جلو به پهلو» احتمال وقوع تصادف فوتی را به شکل معناداری افزایش می‌دهند. این یافته با منطق فیزیکی تصادف نیز همسو است؛ چراکه در چنین مواردی، ضربه مستقیم‌تر به سرنشینان، به‌ویژه راننده وارد می‌شود و شدت تصادم معمولاً بیشتر خواهد بود. این موضوع می‌تواند هشدار مهمی برای طراحی بدنه خودروها، استقرار ایربگ‌ها، و ارتقاء سیستم‌های ترمز و هشداردهنده به شمار رود.

در رتبه بعدی اهمیت، متغیر «روزهای غیرکاری» قرار دارد. تحلیل مدل نشان می‌دهد که تصادف‌هایی که در روزهای تعطیل (پنج‌شنبه و جمعه) رخ می‌دهند، بیش از روزهای کاری به فوت منجر می‌شوند. این پدیده احتمالاً با تغییر الگوی ترافیکی، افزایش سفرهای بین‌شهری، تمایل به رانندگی با سرعت بالاتر در خیابان‌های خلوت، و کاهش سطح تمرکز یا احتیاط همراه است. یافته‌های مدل در این بخش نشان می‌دهد که مدیریت ایمنی معابر باید تمرکز بیشتری بر روزهای تعطیل داشته باشد و سیاست‌های کنترل سرعت، نظارت پلیسی، و آموزش عمومی در این روزها مورد توجه ویژه قرار گیرد.

یکی دیگر از متغیرهای مؤثر، «ساعت وقوع تصادف» است. مدل نشان داده که وقوع تصادف در ساعات اولیه بامداد یا پیش از طلوع خورشید، با افزایش احتمال فوتی بودن همراه است. این موضوع می‌تواند به دلایلی مانند کاهش دید، خستگی رانندگان، خواب‌آلودگی و خلوتی معابر و در نتیجه افزایش سرعت رانندگی نسبت داده شود. در مقابل، در ساعات پرتردد روز که حضور سایر رانندگان و ترافیک بیشتر است، شدت تصادف‌ها گرایش بیشتری به جرحی بودن دارد. این یافته می‌تواند در برنامه‌ریزی زمان‌بندی خدمات شهری، نورپردازی معابر، و توسعه سیستم‌های هشدار جاده‌ای نقش ایفا کند.

علاوه بر این، متغیرهای محیطی همچون «شرایط جوی»، «روشنایی معبر» و «ماه وقوع تصادف» نیز به‌عنوان فاکتورهای مکمل در تصمیم‌گیری مدل عمل کرده‌اند. به‌ویژه در ترکیب با زمان و نوع برخورد، این عوامل توانسته‌اند پیش‌بینی مدل را دقیق‌تر و حساس‌تر سازند. برای مثال، وقوع تصادف در هوای غیرصاف و در روزهای غیرکاری احتمال فوتی بودن را به میزان قابل‌توجهی افزایش داده است. این نشان می‌دهد که شدت تصادف تنها به ویژگی‌های انسانی یا فنی محدود نیست و محیط فیزیکی نیز نقش تعیین‌کننده‌ای ایفا می‌کند.

در مجموع، مدل ارائه‌شده با دقت بالا و قدرت تبیین روشن، قادر است تعامل پیچیده میان عوامل محیطی، زمانی و رفتاری را آشکار سازد. نتایج این تحلیل می‌تواند به‌عنوان مبنایی برای طراحی و پیاده سازی سیاست‌های ایمنی در حمل‌ونقل شهری به کار گرفته شود. از جمله کاربرد های عملی آن می‌توان به شناسایی بازه‌های زمانی پرخطر، طراحی برنامه‌های آموزشی هدفمند، ارتقاء سیستم های نظارتی در روز های تعطیل و بهینه سازی زیرساخت‌های شهری برای کاهش تصادفات شدید اشاره کرد. مدل توسعه یافته، علاوه بر دقت بالا، به دلیل تفسیرپذیری ساختار درخت تصمیم، امکان بهره‌برداری مستقیم برای مدیران شهری، مهندسان ترافیک، و برنامه‌ریزان ایمنی راه را فراهم می‌آورد.

۶- پی‌نوشت‌ها

1. Vision Zero
2. Big Data
3. Random Forest
4. Support Vector Machines
5. Decision tree
6. Black Spot
7. Extreme Gradient Boosting
8. Overfitting
9. Extreme Gradient Boosting

۷- مراجع

-Abrari Vajari, Mohammad, Kayvan Aghabayk, Mohammad Sadeghian, and Nirajan Shiwakoti (2020). A Multinomial Logit Model of Motorcycle Crash Severity at Australian Intersections. *Journal of Safety Research* 73 (June):17–24.
doi.org/10.1016/j.jsr.2020.02.008

-Aron, Maurice, Romain Billot, Nour-Eddin El Faouzi, and Régine Seidowsky (2015). Traffic Indicators, Accidents and Rain: Some Relationships Calibrated on a French Urban Motorway Network. *Transportation Research Procedia* 10:31–40.
doi.org/10.1016/j.trpro.2015.09.053

-Colagrande, Sandro (2022). A Methodology for the Characterization of Urban Road Safety through Accident Data Analysis." *Transportation Research Procedia* 60:504–11.
doi.org/10.1016/j.trpro.2021.12.065

-Eliane, Hernandez Garcia Michelle, and Lozano Angelica. (2024). Analysis of Road Accidents in Two Mixed Industrial Urban Zones, Using Nested Poisson and Negative Binomial Models. *Transportation Research Procedia* 78:377–83.
doi.org/10.1016/j.trpro.2024.02.048

-EslamiNezhad, S. A., and M. R. Delavar (2019). An Integrated Network-Constrained Spatial Analysis For Car Accidents: A Case Study Of Tehran City, Iran. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-4/W18 (October):335–42.
doi.org/10.5194/isprs-archives-XLII-4-W18-335-2019

-García, María Flor, Armando Ortuño Padilla, Begoña Guirao Abad, and Jairo Casares Blanco. (2021). Urban Road Accidents and Ride-Hailing Services: A Study of Dependence in Madrid. *Transportation Research Procedia* 58:301–8.
doi.org/10.1016/j.trpro.2021.11.041

-Hyodo, Satoshi, and Tomoyuki Todoroki (2018). An Analysis of Risk Factors for Rear-Ender Accident on Urban Expressway Considering Accident Severity. *Transportation Research Procedia* 34:203–10.
doi.org/10.1016/j.trpro.2018.11.033

-Kaplan, Sigal, and Carlo Giacomo Prato (2012). Risk Factors Associated with Bus Accident Severity in the United States: A Generalized Ordered Logit Model. *Journal of Safety Research* 43 (3): 171–80.
doi.org/10.1016/j.jsr.2012.05.003

-Rezapour, Mahdi, Ahmed Farid, Sahima Nazneen, and Khaled Ksaibati (2021). Using Machine Leaning Techniques for Evaluation of Motorcycle Injury Severity. *IATSS Research* 45 (3): 277–85.
doi.org/10.1016/j.iatssr.2020.07.004

-Santos, Kenny, João P. Dias, and Conceição Amado (2022). A Literature Review of Machine Learning Algorithms for Crash Injury Severity Prediction. *Journal of Safety Research* 80 (February), 254–69.
doi.org/10.1016/j.jsr.2021.12.007

-Wang, Deyu, Qinyi Liu, Liang Ma, Yijing Zhang, and Haozhe Cong (2019). Road Traffic Accident Severity Analysis: A Census-Based Study in China. *Journal of Safety Research* 70 (September):135–47.
doi.org/10.1016/j.jsr.2019.06.002

-Yoon, Junho (2025). Prediction of High-Risk Areas Using the Interpretable Machine Learning: Based on Each Determinant for the Severity of Pedestrian Crashes. *Journal of Transport Geography* 126 (June):104216.
doi.org/10.1016/j.jtrangeo.2025.104216

-Yuan, Yalong, Min Yang, Yanyong Guo, Soora Rasouli, Zuoxian Gan, and Yifeng Ren (2021). Risk Factors Associated with Truck-Involved Fatal Crash Severity: Analyzing Their Impact for Different Groups of Truck Drivers. *Journal of Safety Research* 76 (February):154–65.
doi.org/10.1016/j.jsr.2020.12.012

Modeling and Predicting Urban Traffic Accident Severity via Intelligent Algorithms

Alireza Taherian, M.Sc., Grad., Civil and Environmental Engineering, Department of Transportation Planing and Engineering, Tarbiat Modarres University, Tehran, Iran.
Saeid Sherafatipour, Assistant Professor, Civil and Environmental Engineering, Department of Transportation Planing and Engineering, Tarbiat Modarres University, Tehran, Iran.
Amirhossein Gheyami, M.Sc., Grad., Civil and Environmental Engineering, Department of Transportation Planing and Engineering, Tarbiat Modarres University, Tehran, Iran.
Mahmmod Saffarzadeh, Professor, Civil and Environmental Engineering, Department of Transportation Planing and Engineering, Tarbiat Modarres University, Tehran, Iran.

E-mail: sherafatipour@modares.ac.ir

Received: April 2025- Accepted: November 2025

ABSTRACT
Urban traffic accidents represent a significant challenge in Iran, primarily due to inadequate infrastructure, lack of adherence to traffic regulations, and the absence of a well-established driving culture. These factors have led to considerable human and economic losses. In contrast, many developed countries have achieved substantial reductions in urban traffic accidents through the strategic application of technology, public education, and intelligent urban planning. The success of these nations highlights the critical role of effective traffic management in enhancing urban safety. This study aims to identify and predict the severity of urban road accidents in Alborz Province using a decision tree algorithm. For this purpose, a dataset comprising 5,000 recorded accidents from 2018 to 2022 was collected and, after data cleansing, analyzed in the Python programming environment. The results revealed that the three most influential variables in predicting accident severity are the type of vehicle collision, the time of occurrence, and whether the accident took place on a non-working day. The final model achieved a prediction accuracy of 0.79, indicating a relatively high level of precision in distinguishing between injury and fatal accidents. The findings suggest that, in addition to human-related factors, environmental and structural elements such as the nature of vehicle collisions significantly contribute to the likelihood of fatal outcomes. Leveraging real-world data and an interpretable modeling approach, this research offers practical insights for road safety policymaking and the design of effective preventive strategies.

Keywords: Accident Severity, Decision Tree, Urban Accidents, Traffic Safety