مدل‌سازی و پیش‌بینی تخلفات عدم رعایت فاصله زمانی ایمن در محورهای برون‌شهری با استفاده از داده‌های ساعتی ترددشماری (مطالعه موردی: استان خراسان رضوی در سال ۱۴۰۴)

مقاله علمی-پژوهشی

حسین محمدی، دانش آموخته کارشناسی ارشد، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه فردوسی، مشهد، ایران
*سید علی ضیائی (نویسنده مسئول)، دانشیار، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، ایران
سحر علیشاهی، دانش آموخته کارشناس ارشد، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه فردوسی، مشهد، ایران
*پست الکترونیکی نویسنده مسئول: sa-ziaee@um.ac.ir

دریافت: ۱۴۰۵/۰۱/۲۸- پذیرش: ۱۴۰۵/۰۵/۰۵
صفحه ۲۴۶-۲۳۵

چکیده

هدف** این پژوهش، ارائه چارچوبی داده‌محور برای مدل‌سازی و پیش‌بینی تخلفات عدم رعایت فاصله زمانی ایمن در محورهای برون‌شهری با استفاده از داده‌های ساعتی ترددشماری سامانه ۱۴۱ است. متغیر «تخلف فاصله غیرمجاز» بر اساس تعریف سامانه، عدم رعایت حداقل فاصله زمانی دو ثانیه‌ای با وسیله نقلیه جلویی را نشان می‌دهد و در این مطالعه به‌عنوان شاخص جانشین ایمنی برای پایش پیش‌دستانه ریسک به‌کار رفت. قلمرو مطالعه شامل ۱۳۶ محور یک‌جهته استان خراسان رضوی در سال ۱۴۰۴ است. داده‌های خام دوازده‌ماهه مجدداً خوانده و بر اساس چارچوب کنترل کیفیت شامل بررسی مدت کارکرد، حجم تردد غیرمثبت، سرعت نامعتبر، ناسازگاری شمارش کلاس‌ها، اختلاف تعداد برآوردشده و تکرار کلید محور-زمان پالایش شدند. از میان ۱،۱۶۶،۱۰۲ رکورد ساعتی، ۱،۱۴۳،۱۶۹ رکورد به‌عنوان داده آماده تحلیل حفظ شد. متغیر هدف، نرخ تخلف فاصله غیرمجاز به ازای هر ۱۰۰۰ وسیله نقلیه تعریف گردید. نتایج نشان داد نرخ شبکه‌ای این تخلف برابر با ۲۰۳٫۹۰ تخلف به ازای هر ۱۰۰۰ وسیله نقلیه است. برای پیش‌بینی نرخ تخلف، داده‌ها در سطح محور-ماه-ساعت-روز هفته تجمیع و ماه‌های ۱ تا ۹ برای آموزش و ماه‌های ۱۰ تا ۱۲ برای آزمون استفاده شد. مدل جنگل تصادفی با R^2 برابر ۰٫۹۴۴، MAE برابر ۱۳٫۸۲ و $RMSE$ برابر ۲۱٫۲۴ عملکرد مناسبی در آزمون زمانی نشان داد. یافته‌ها حاکی از نقش برجسته حجم تردد ساعتی، سرعت متوسط، نرخ تخلف سبقت و تفاوت‌های محوری در توضیح نرخ تخلف فاصله است. نتایج پژوهش می‌تواند مبنایی برای طراحی سامانه هشدار زودهنگام، رتبه‌بندی محورهای پرریسک و برنامه‌ریزی مداخله‌های مدیریتی در ایمنی راه باشد.

واژه‌های کلیدی: ترددشماری، تخلف فاصله غیرمجاز، فاصله زمانی دو ثانیه‌ای، شاخص جانشین ایمنی، یادگیری ماشین

۱- مقدمه

ایمنی راه یکی از حوزه‌های اصلی سیاست‌گذاری حمل‌ونقل است؛ زیرا خسارت‌های انسانی، اقتصادی و اجتماعی تصادفات جاده‌ای همچنان در سطح جهانی بسیار بالاست. گزارش جهانی ایمنی راه سازمان جهانی بهداشت نشان می‌دهد که سالانه حدود ۱٫۱۹ میلیون نفر در تصادفات جاده‌ای جان خود را از دست می‌دهند و آسیب‌های ترافیکی از مهم‌ترین علل مرگ در گروه سنی کودکان و جوانان است (سازمان جهانی بهداشت، ۲۰۲۳). گزارش‌های سالانه ITF/OECD نیز بر پایش عملکرد ایمنی، استفاده از داده‌های معتبر و سیاست‌گذاری مبتنی بر شواهد تأکید می‌کنند (مجمع بین‌المللی حمل‌ونقل، ۲۰۲۴، ۲۰۲۵).

در چنین چارچوبی، داده‌های تردشماری برخط می‌توانند به ابزار فعال پیش‌بینی، کنترل و مدیریت ریسک تبدیل شوند، نه صرفاً منبعی برای گزارش حجم تردد. در ایران، سامانه‌های تردشماری جاده‌ای حجم تردد، سرعت متوسط، ترکیب ناوگان و برخی تخلفات رفتاری را ثبت می‌کنند. یکی از متغیرهای ارزشمند این سامانه، «تخلف فاصله غیرمجاز» است که بر اساس عدم رعایت فاصله زمانی ایمن دو ثانیه‌ای تعریف می‌شود. این متغیر از نظر مهندسی ترافیک اهمیت دارد؛ زیرا به‌طور غیرمستقیم شدت فشردگی جریان، ریسک برخورد از عقب و رفتار پرریسک رانندگان را نشان می‌دهد. بر خلاف داده تصادف که رخدادمحور، کم‌فراوان و گاه با تأخیر در دسترس است، تخلف فاصله در سطح ساعتی و محوری ثبت می‌شود و برای تحلیل پیش‌دستانه ایمنی ظرفیت بالایی دارد.

پژوهش حاضر به سه پرسش پاسخ می‌دهد:

– الگوی زمانی و محوری تخلفات عدم رعایت فاصله زمانی ایمن در شبکه راه‌های استان خراسان رضوی چگونه است؟

– کدام متغیرهای جریان ترافیک و ترکیب ناوگان با نرخ این تخلف ارتباط بیشتری دارند؟

– آیا می‌توان با استفاده از داده‌های ساعتی تردشماری، نرخ تخلف فاصله را در دوره زمانی آینده با دقت قابل قبول پیش‌بینی کرد؟

نوآوری پژوهش در چهار محور خلاصه می‌شود: نخست، بازخوانی و پالایش داده‌های خام دوازده‌ماهه سامانه ۱۴۱ در مقیاس محور یک‌جهته؛ دوم، تبدیل تعداد خام تخلف فاصله به نرخ قابل مقایسه به ازای هر ۱۰۰۰ وسیله نقلیه؛ سوم، استفاده از تخلف فاصله به‌عنوان شاخص جانشین ایمنی و نه فقط یک متغیر آماری؛ و چهارم، آزمون مدل پیش‌بینی با تقسیم زمانی آموزش و آزمون برای جلوگیری از ارزیابی خوش‌بینانه.

۲- پیشینه تحقیق

۲-۱- شاخص‌های جانشین ایمنی

تحلیل ایمنی سنتی عمدتاً بر آمار تصادفات متکی است، اما تصادفات رخدادهایی نسبتاً نادر و تحت تأثیر گزارش‌دهی، شرایط محیطی و تصادف آماری هستند. به همین دلیل، ادبیات مهندسی ترافیک از شاخص‌های جانشین ایمنی مانند زمان تا برخورد، زمان پس از تجاوز، نرخ تعارض، فاصله مکانی و فاصله زمانی برای ارزیابی پیش‌دستانه ریسک استفاده می‌کند (هایوارد، ۱۹۷۲؛ گتمن و هد، ۲۰۰۳؛ گتمن و همکاران، ۲۰۰۸؛ میندرهود و بووی، ۲۰۰۱؛ هیدن، ۱۹۸۷؛ نیکولائو و همکاران، ۲۰۲۳). مدل SSAM اداره فدرال بزرگراه‌های آمریکا نمونه‌ای شناخته‌شده برای استخراج تعارضات و شاخص‌های شدت از داده‌های مسیر حرکت وسایل نقلیه است (گتمن و هد، ۲۰۰۳؛ گتمن و همکاران، ۲۰۰۸). اگرچه داده این پژوهش از نوع تردشماری و نه مسیر حرکت لحظه‌ای است، متغیر فاصله غیرمجاز به دلیل مبتنی بودن بر فاصله زمانی دو ثانیه‌ای، در خانواده شاخص‌های رفتاری جانشین ایمنی قرار می‌گیرد.

۲-۲-داده‌های تردشماری و پیش‌بینی ترافیک

پیش‌بینی جریان و رفتار ترافیک از ارکان سامانه‌های حمل‌ونقل هوشمند است. مرورهای جدید نشان می‌دهند که روش‌های آماری، یادگیری ماشین و یادگیری عمیق در پیش‌بینی جریان ترافیک، سرعت و شاخص‌های ایمنی کاربرد گسترده دارند (افندی‌زاده و همکاران، ۲۰۲٤؛ لیو و همکاران، ۲۰۲۵؛ روی و حسن، ۲۰۲٤؛ وانگ و همکاران، ۲۰۲٤؛ لی و همکاران، ۲۰۲۵). در حوزه ایمنی، مدل‌های یادگیری ماشین برای پیش‌بینی ریسک تصادف، شدت تصادف و شناسایی عوامل مؤثر استفاده شده‌اند و پژوهش‌های اخیر بر تفسیرپذیری مدل و نقش متغیرهایی مانند زمان، مکان، نوع وسیله نقلیه، ویژگی راه و رفتار راننده تأکید دارند (عبدالعاطی و پاندی، ۲۰۰۵؛ لرد و منرینگ، ۲۰۱۰؛ یو و عبدالعاطی، ۲۰۱٤؛ مصطفی و همکاران، ۲۰۲۵؛ ریفات و همکاران، ۲۰۲٤).

بیشتر مطالعات پیش‌بینی ایمنی یا به داده تصادف وابسته‌اند یا به داده‌های ریزدانه مسیر حرکت نیاز دارند. پژوهش حاضر در فاصله میان این دو رویکرد قرار می‌گیرد: از یک سو، به‌جای تصادف از یک شاخص رفتاری و پرتکرار استفاده می‌کند؛ و از سوی دیگر، به‌جای داده‌های مسیر حرکت بسیار پرهزینه، از داده‌های موجود و رسمی تردشماری بهره می‌برد. این ویژگی، قابلیت پیاده‌سازی عملی چارچوب پیشنهادی را برای راهداری و مدیریت ایمنی افزایش می‌دهد.

۳- داده‌ها و قلمرو مطالعه

قلمرو مکانی پژوهش شامل محورهای برون‌شهری استان خراسان رضوی است. داده‌ها مربوط به سال ۱٤۰٤ و شامل

اطلاعات ساعتی و روزانه سامانه ترددشماری ۱۴۱ هستند. هر کد شش‌رقمی محور به‌عنوان یک جهت حرکت مستقل در نظر گرفته شد؛ بنابراین جهت رفت و برگشت یک مسیر، دو واحد تحلیل جداگانه محسوب شدند. در مجموع، فایل ساعتی و فایل روزانه خوانده شد و ۱۳۶ محور یکتا در پایگاه داده شناسایی شد.

۳-۱- تعریف کلاس‌های وسیله نقلیه

در داده‌های ترددشماری، کلاس ۱ شامل سواری و وانت؛ کلاس ۲ شامل کامیونت، کامیون‌های کوچک و مینی‌بوس؛ کلاس ۳ شامل کامیون معمولی کمتر از ۱۰ متر و کامیون سه‌محوره؛ کلاس ٤ شامل اتوبوس؛ و کلاس ۵ شامل تریلی و وسایل باربری با بیش از سه محور است. بر اساس این تعریف، متغیرهای سهم وسایل سبک، سهم وسایل تجاری کلاس ۲ تا ۵، سهم وسایل سنگین کلاس ۳ تا ۵، سهم باربری سنگین کلاس ۳ و ۵ و سهم اتوبوس ساخته شدند.

۳-۲- تعریف متغیر هدف

متغیر هدف پژوهش نرخ تخلف فاصله غیرمجاز به ازای هر ۱۰۰۰ وسیله نقلیه است. این نرخ از رابطه زیر محاسبه شد.

Unsafe Following-Distance Rate = Unsafe Following-Distance Violations / Total Vehicles × ۱۰۰۰

این تبدیل برای مقایسه منصفانه محورها ضروری است؛ زیرا تعداد خام تخلف در محورهای پرتردد به‌طور طبیعی بیشتر است و بدون استانداردسازی می‌تواند رتبه‌بندی ریسک را دچار سوگیری کند.

چارچوب کنترل کیفیت و آماده‌سازی داده

با توجه به ماهیت خام داده‌های ترددشماری، کنترل کیفیت بخشی اساسی از روش پژوهش بود. رکوردها بر اساس قواعد قطعی و قابل بازتولید برچسب‌گذاری شدند. معیارهای حذف شامل تاریخ نامعتبر، مدت کارکرد ناقص، حجم تردد غیرمثبت، سرعت متوسط غیرواقعی، مقدار منفی در شمارش‌ها، بزرگ‌تر بودن شمارش تخلف از تردد کل و ناسازگاری شدید مجموع کلاس‌ها با تعداد کل بود. رکوردهای دارای هشدار اما فاقد خطای قطعی با برچسب REVIEW حفظ شدند تا تحلیل حساسیت در مراحل بعدامکان‌پذیر باشد.

جدول ۱. خلاصه پایگاه داده پس از بازخوانی و کنترل کیفیت

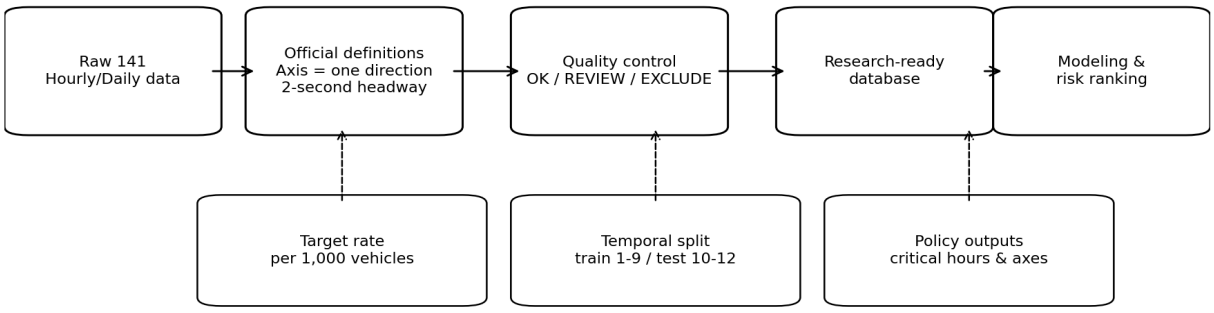
شاخص	مقدار
رکورد ساعتی خام	۱،۱۶۶،۱۰۲
رکورد ساعتی آماده تحلیل	۱،۱٤۳،۱٦۹
رکورد روزانه خام	٤۹،۲۸۰
رکورد روزانه آماده تحلیل	٤۲،۳۰۸
تعداد محور یکتا	۱۳٦
کلید تکراری محور-زمان	۰
نرخ شبکه‌ای تخلف فاصله/۱۰۰۰	۲۰۳٫۹۰

جدول ۲. توزیع رکوردها بر اساس شدت کنترل کیفیت

مقیاس داده	وضعیت کنترل کیفیت	تعداد رکورد
Daily	EXCLUDE	٦٩٧٢
Daily	OK	٤١٧٠١
Daily	REVIEW	٦٠٧
Hourly	EXCLUDE	٢٢٩٣٣
Hourly	OK	١١٣٤٦٧٠
Hourly	REVIEW	٨٤٩٩

جدول ۳. مهم‌ترین کدهای کنترل کیفیت شناسایی‌شده در داده خام

مقیاس داده	کد مسئله کیفیت	تعداد رکورد
Daily	OK	٤١٧٠١
Daily	BAD_DURATION_EXPECTED_١٤٤٠	٦٩٧٢
Daily	ESTIMATED_COUNT_GAP_GT١٠PCT	٣١٨٥
Hourly	OK	١١٣٤٦٧٠
Hourly	ESTIMATED_COUNT_GAP_GT١٠PCT	٢٤١٨٠
Hourly	BAD_DURATION_EXPECTED_٦٠	١٩٠٤٢
Hourly	NON_POSITIVE_TOTAL	٣٩٩٨
Hourly	AVG_SPEED_IMPLAUSIBLE	٤



شکل ۱. چارچوب علمی پژوهش از داده خام تا خروجی مدیریتی

جدول ٤. طراحی آزمایش پیش‌بینی و ساختار مدل

مقدار	مؤلفه
محور–ماه–ساعت–روز هفته	سطح تحلیل مدل
ماه‌های ۱ تا ۹	دوره آموزش
ماه‌های ۱۰ تا ۱۲	دوره آزمون
۲۷۳،۳٥٤	تعداد گروه‌های مدل‌سازی
۸۰،۰۰۰	نمونه آموزش
۳۰،۰۰۰	نمونه آزمون
نرخ تخلف فاصله/۱۰۰۰	متغیر هدف
Random Forest	مدل اصلی

روش‌شناسی تحلیل و مدل‌سازی

پس از کنترل کیفیت، رکوردهای ساعتی آماده تحلیل در سطح محور–ماه–ساعت–روز هفته تجمیع شدند. این سطح تجمیع از یک سو ساختار زمانی و محوری داده را حفظ می‌کند و از سوی دیگر نوسانات بسیار ریز و خطاهای تصادفی را کاهش می‌دهد. در مجموع، مشاهده تحلیلی برای مدل‌سازی ساخته شد. برای آزمون تعمیم‌پذیری زمانی، ماه‌های ۱ تا ۹ به‌عنوان دوره آموزش و ماه‌های ۱۰ تا ۱۲ به‌عنوان دوره آزمون استفاده شدند. مدل پایه میانگین و مدل جنگل تصادفی با نمونه‌گیری ثابت از گروه‌های آموزش و آزمون مقایسه شدند.

جدول ۵. عملکرد مدل‌ها در آزمون زمانی ماه‌های ۱۰ تا ۱۲

مدل	داده آموزش	داده آزمون	MAE	RMSE	R^2
Mean baseline	۸۰۰۰۰	۳۰۰۰۰	۷۳,۷۳٦	۹۲,۲٥	۰,۰٥۱–
Random Forest	۸۰۰۰۰	۳۰۰۰۰	۱۳,۸۲۳	۲۱,۲۳۸	۰,۹٤٤

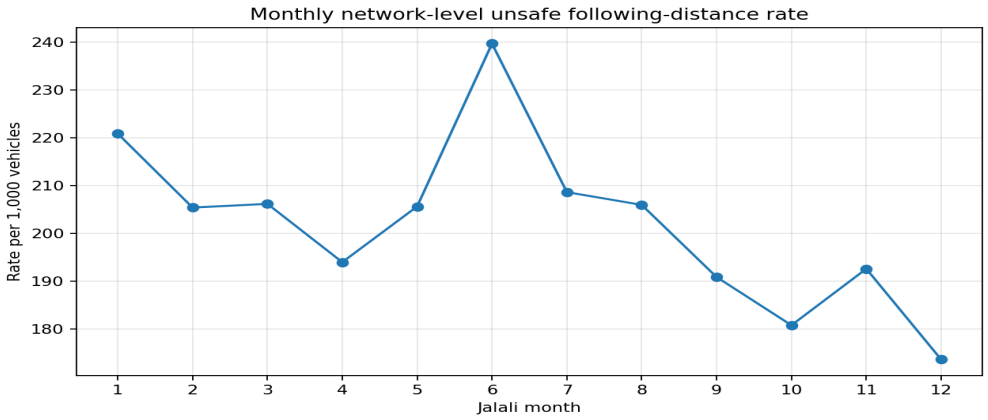
٤–یافته‌ها

الگوی ماهانه و شبکه‌ای

در داده‌های ساعتی آماده تحلیل، مجموع تردد ثبت‌شده برابر با ٤۲۷،٥٤۰،٥٦۰ وسیله نقلیه و مجموع تخلفات فاصله غیرمجاز برابر با ۸۷،۱۷۷،۳۹۷ مورد بود. نرخ شبکه‌ای سالانه تخلف فاصله برابر با ۲۰۳,۹۰ تخلف به ازای هر ۱۰۰۰ وسیله نقلیه محاسبه شد. جدول ٦ و شکل ۲ الگوی ماهانه این نرخ را نشان می‌دهند.

جدول ٦. خلاصه ماهانه داده‌های ساعتی آماده تحلیل

ماه	نام ماه	رکورد آماده	محور	تردد کل	تخلف فاصله	نرخ فاصله/۱۰۰۰
۱	فروردین	۹۷٬٦٦۸	۱۳٦	۳۸٬۰۵۳٬٦۵۹	۸٬٤۰٦٬۳۵۵	۲۲۰٫۹۱
۲	اردیبهشت	۹۸٬٤۱۵	۱۳٦	۳٦٬۰۷٤٬٦۲۰	۷٬٤۱۰٬٦۲۷	۲۰۵٫٤۳
۳	خرداد	۹۸٬٦٤۳	۱۳٦	۳۷٬٦۲۸٬۰۳۸	۷٬۷۵۸٬۷۵۲	۲۰٦٫۲۰
٤	تیر	۹۸٬۲٦٦	۱۳٦	۳٦٬٦۰۱٬۵۰۹	۷٬۱۰۰٬۸۰۸	۱۹٤٫۰۰
۵	مرداد	۹٦٬۷۲٤	۱۳٦	۳۸٬۲۵۸٬۹٤۱	۷٬۸٦۸٬۱۱۵	۲۰۵٫٦۵
٦	شهریور	۹۵٬۹۳۲	۱۳۵	٤۵٬۱۳٦٬۵۸۷	۱۰٬۸۱۹٬۹۹۷	۲۳۹٫۷۲
۷	مهر	۹۲٬۷۱۵	۱۳٦	۳٦٬٦٤۵٬٦۰۱	۷٬٦٤٦٬۲۷۸	۲۰۸٫٦٦
۸	آبان	۹۳٬۷۰٤	۱۳٦	۳۵٬۹۰۸٬۰٤۰	۷٬۳۹٦٬٤٦۵	۲۰۵٫۹۸
۹	آذر	۹۱٬۵۷۳	۱۳٦	۳۱٬٤۰۹٬۷۳۳	۵٬۹۹۵٬٦۱۰	۱۹۰٫۸۸
۱۰	دی	۹٤٬۱٤۰	۱۳٦	۲۹٬۹۰۸٬۲۲۵	۵٬٤۰۷٬٦۸۰	۱۸۰٫۸۱
۱۱	بهمن	۹۳٬۲۳۳	۱۳٦	۳۲٬٤۵٤٬٤۷۵	٦٬۲۵۱٬۲٤٦	۱۹۲٫٦۲
۱۲	اسفند	۹۲٬۱۵٦	۱۳٦	۲۹٬٤٦۱٬۱۳۲	۵٬۱۱۵٬٤٦٤	۱۷۳٫٦۳

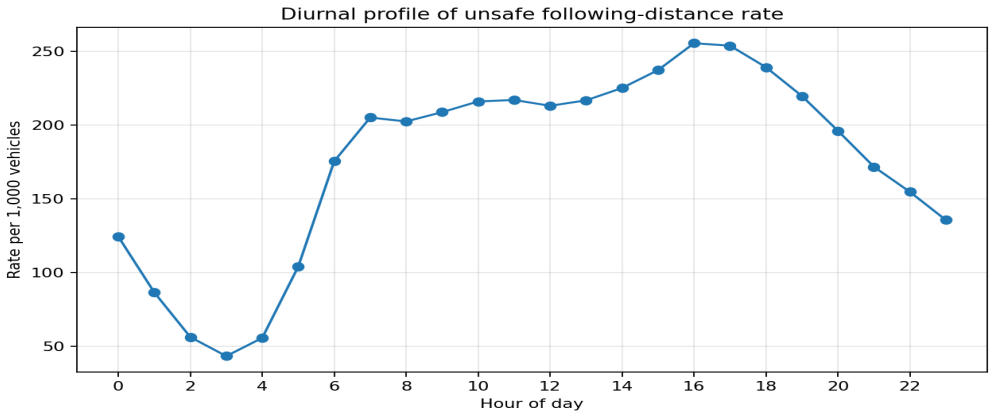


شکل ۲. روند ماهانه نرخ تخلف فاصله غیرمجاز در سطح شبکه

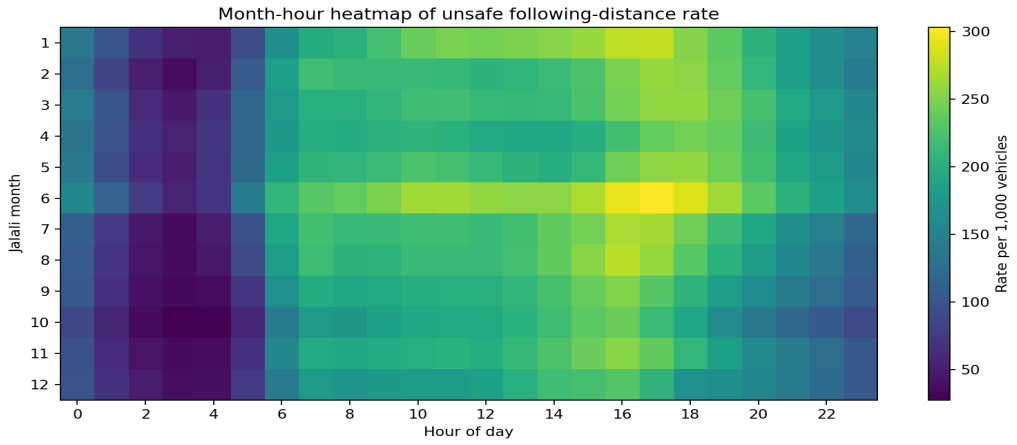
الگوی ساعتی و ماه–ساعت

الگوی ساعتی نشان داد که تخلف فاصله متأثر از ساختار شبانه‌روزی تقاضا و جریان ترافیک است. پروفیل ساعتی و نقشه حرارتی ماه–ساعت می‌توانند برای طراحی هشدارهای زمانی، کنترل هوشمند و برنامه‌ریزی گشت‌های ایمنی استفاده شوند.

.



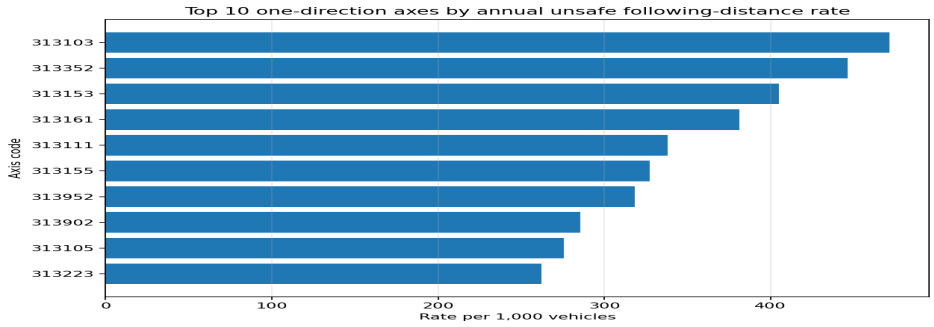
شکل ۳. پروفیل ساعتی نرخ تخلف فاصله غیرمجاز



شکل ٤. نقشه حرارتی ماه–ساعت نرخ تخلف فاصله غیرمجاز

محورهای اولویت‌دار

از آنجا که هر محور یک جهت حرکت مستقل است، رتبه‌بندی محورهای پرریسک باید در سطح کد محور انجام شود. جدول ۷ و شکل ۵ ده محور با بیشترین نرخ تخلف فاصله غیرمجاز را نشان می‌دهند. این محورها اولویت بالاتری برای بررسی میدانی، تحلیل هندسه مسیر، کنترل رفتاری و مداخله مدیریتی دارند.



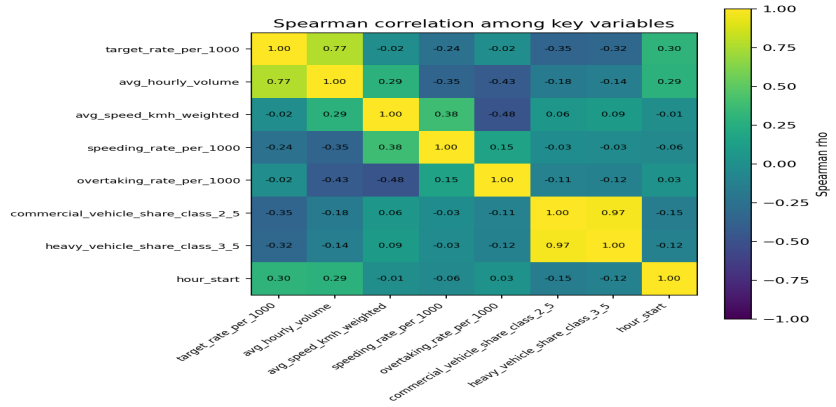
شکل ۵. محورهای دارای بیشترین نرخ تخلف فاصله غیرمجاز

جدول ۷. ده محور یک‌جهته با بیشترین نرخ سالانه تخلف فاصله غیرمجاز

کد محور	نام محور	تردد کل	تخلف فاصله	نرخ فاصله/۱۰۰۰
۳۱۳۱۰۳	مشهد – چناران (مشهد – تقاطع آزادراه)	۲۰٬۱۸۰٬۸۲۱	۹٬۵۲۲٬۲۶۱	۴۷۱٫۸۵
۳۱۳۳۵۲	فریمان – مشهد (عوارضی فریمان)	۳٬۵۲۸٬۳۸۱	۱٬۵۷۵٬۳۹۶	۴۴۶٫۴۹
۳۱۳۱۵۳	چناران – مشهد (تقاطع آزادراه – مشهد)	۱۷٬۰۹۴٬۴۲۲	۶٬۹۲۲٬۲۶۶	۴۰۴٫۹۴
۳۱۳۱۶۱	شاندیز – مشهد	۱۰٬۷۹۵٬۷۴۳	۴٬۱۱۶٬۸۱۰	۳۸۱٫۳۴
۳۱۳۱۱۱	مشهد – شاندیز	۱۱٬۸۶۴٬۷۸۲	۴٬۰۱۳٬۴۷۶	۳۳۸٫۲۷
۳۱۳۱۵۵	آزادراه باغچه – مشهد (مشهد)	۱۰٬۸۴۴٬۰۶۵	۳٬۵۵۱٬۶۱۶	۳۲۷٫۵۲
۳۱۳۹۵۲	چناران – مشهد (چناران – گلبهار)	۱۰٬۴۰۸٬۷۶۵	۳٬۳۱۳٬۲۱۲	۳۱۸٫۳۱
۳۱۳۹۰۲	مشهد – چناران (گلبهار – چناران)	۹٬۸۷۴٬۵۸۰	۲٬۸۲۰٬۵۰۲	۲۸۵٫۶۳
۳۱۳۱۰۵	آزادراه مشهد – باغچه (مشهد)	۱۲٬۶۰۳٬۹۰۶	۳٬۴۷۳٬۵۷۰	۲۷۵٫۶۰
۳۱۳۲۲۳	نیشابور – مشهد (نیشابور) (متناظر توزین)	۱۰٬۷۲۷٬۸۵۷	۲٬۸۱۲٬۷۵۶	۲۶۲٫۱۹

روابط آماری و اهمیت متغیرها

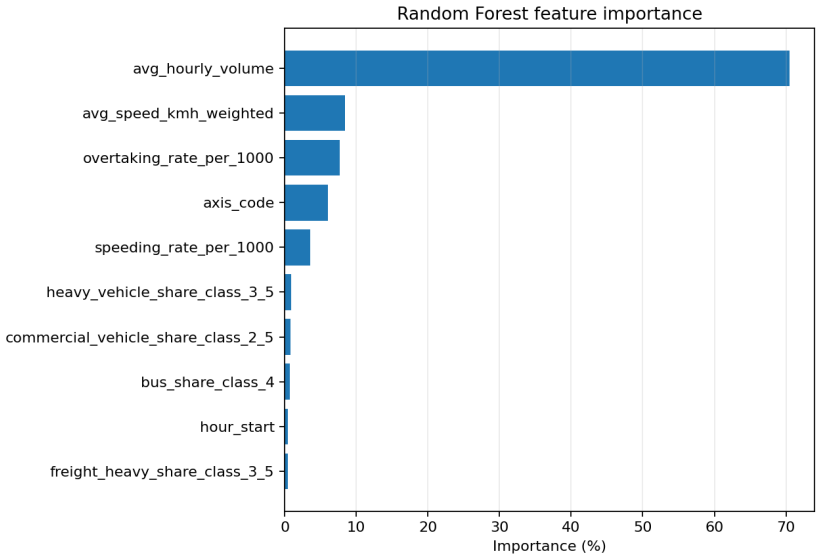
نتایج همبستگی و اهمیت متغیرها نشان داد که حجم تردد ساعتی مهم‌ترین عامل توضیح‌دهنده نرخ تخلف فاصله است. پس از آن سرعت متوسط، نرخ تخلف سبقت، تفاوت‌های محوری و نرخ تخلف سرعت در مدل اهمیت دارند. این یافته از نظر مهندسی قابل تفسیر است؛ زیرا افزایش حجم جریان احتمال کاهش فاصله زمانی و افزایش برهم‌کنش وسایل نقلیه را بیشتر می‌کند.



شکل ٦. ماتریس همبستگی اسپیرمن متغیرهای کلیدی

جدول ۸. اهمیت نسبی متغیرها در مدل جنگل تصادفی

متغیر	اهمیت نسبی (٪)
avg_hourly_volume	۷۰٫۴۵۷
avg_speed_kmh_weighted	۸٫۴۴۶
overtaking_rate_per_۱۰۰۰	۷٫۷۲۵
axis_code	۶٫۰۴۶
speeding_rate_per_۱۰۰۰	۳٫۵۸۶
heavy_vehicle_share_class_۳_۵	۰٫۸۶۵
commercial_vehicle_share_class_۲_۵	۰٫۸۰۱
bus_share_class_۴	۰٫۶۹۶
hour_start	۰٫۴۸۲
freight_heavy_share_class_۳_۵	۰٫۴۵۶
month_no	۰٫۳۱۴
weekday_iso_monday۰	۰٫۰۶۸



شکل ۷. اهمیت متغیرها در پیش‌بینی نرخ تخلف فاصله

بحث

نتایج پژوهش نشان می‌دهد که تخلف فاصله غیرمجاز یک شاخص رفتاری حساس به شرایط جریان ترافیک است و می‌تواند مکمل مناسبی برای شاخص‌های سنتی مانند سرعت متوسط و حجم تردد باشد. اهمیت بسیار زیاد حجم تردد در مدل، با منطق نظری فاصله زمانی سازگار است؛ زیرا با افزایش تعداد وسایل نقلیه در یک بازه زمانی ثابت، متوسط فاصله زمانی میان وسایل نقلیه کاهش یافته و احتمال ثبت فاصله کمتر از آستانه دو ثانیه افزایش می‌یابد. در عین حال، اهمیت کد محور

نشان می‌دهد که ویژگی‌های ثابت مکانی، مانند کارکرد محور، موقعیت نسبت به شهرهای بزرگ، هندسه مسیر، دسترسی‌ها و محدودیت سرعت، در رفتار رانندگان مؤثر است؛ هرچند همه این متغیرها در داده فعلی به‌طور مستقیم موجود نیستند.از نظر کاربردی، نتایج پیشنهاد می‌کند که مدیریت ایمنی راه نباید صرفاً بر تخلف سرعت متمرکز باشد. محورهایی با نرخ بالای تخلف فاصله ممکن است مستعد برخورد از عقب، ناپایداری جریان و افزایش ریسک تعارض باشند. بنابراین، شاخص پیشنهادی می‌تواند به‌صورت ماهانه و ساعتی در داشبورد راهداری پایش شود و برای اولویت‌بندی مداخلات نرم و سخت مورد استفاده قرار گیرد.

۵- نتیجه‌گیری

این مقاله نشان داد که داده‌های ترددشماری سامانه ۱۴۱، در صورت بازخوانی و کنترل کیفیت دقیق، می‌توانند مبنای یک پژوهش داده‌محور و قابل دفاع در حوزه ایمنی راه باشند. چارچوب پیشنهادی، داده خام را به پایگاه داده آماده مقاله تبدیل کرد، متغیر هدف استاندارد ساخت و الگوی زمانی–محوری تخلف فاصله را مدل‌سازی نمود. مدل جنگل تصادفی در آزمون زمانی عملکرد مطلوبی نشان داد و می‌تواند پایه اولیه برای سامانه هشدار زودهنگام باشد.

پیشنهادهای اجرایی پژوهش عبارت‌اند از: ۱) افزودن نرخ تخلف فاصله به داشبوردهای ایمنی راه ۲) رتبه‌بندی ماهانه محورهای پرریسک ۳) بررسی میدانی محورهای اولویت‌دار ٤) ترکیب داده ترددشماری با داده تصادف، وضعیت آب‌وهوا، هندسه راه و حد سرعت مجاز ٥) توسعه مدل‌های تفسیرپذیر مانند SHAP برای استخراج آستانه‌های مداخله و ٦) طراحی سامانه هشدار برای ساعت‌های بحرانی بر اساس محور، ساعت و حجم تردد.

٦-سپاسگزاری

از سازمان راهداری و حمل‌ونقل جاده‌ای بابت انتشار داده‌های ترددشماری و فراهم‌کردن امکان استفاده پژوهشی از داده‌های سامانه ۱٤۱ قدردانی می‌شود. مسئولیت پالایش، تحلیل و تفسیر داده‌ها بر عهده نویسندگان است.

۷- مراجع

-AASHTO (2010). Highway Safety Manual. American Association of State Highway and Transportation Officials.

-Abdel-Aty, M., & Pande, A. (2005). Identifying crash propensity using specific traffic speed conditions. *Journal of Safety ,Research(36)*1, 97-108.

-Afandizadeh, S., et al., (2024). Deep learning algorithms for traffic forecasting: A comprehensive review. *Journal of Advanced Transportation*.

-Breiman, L. (2001). Random forests. *Machine Learning*(45), 5-32.

-Chen, T., & Guestrin, C,XGBoost (2016). A scalable tree boosting system. *Proceedings of KDD*.

-Federal Highway Administration (2026). Surrogate Safety Assessment Model(SSAM) overview. *Turner-Fairbank Highway Research Center*.

-Gettman, D., & Head, L, Surrogate (2003). safety measures from traffic simulation models *FHWA-RD03-05*.

-Gettman, D., Pu, L., Sayed, T., & Shelby, S, Surrogate (2008). Safety Assessment Model and Validation. *FHWA-HRT08-051*.

-Minderhoud, M. M., & Bovy, P. H. L, (2001). Extended time-to-collision measures for road traffic *safety assessment Accident Analysis & Prevention33,* 89-97.

-Mostafa, A. M., et al., (2025). AI-based prediction of traffic crash severity for improving road safety. *Scientific Reports*.

-Nikolaou, D., et al, (2023). Exploiting surrogate safety measures and road design characteristics for motorway safety analysis. *Infrastructures*(8)3, 40-41.

-Rifat, M. A. K., Kabir, A., & Huq, A. S. (2024). An explainable machine learning approach to traffic accident fatality prediction. arXiv 2409, 11929.

-Roy, S., & Hasan, S. (2024). Advance real-time detection of traffic incidents in highways using vehicle trajectory data. arXiv 2408, 16773.

-Wang, J., et al., (2024). Can historical accident data improve sustainable urban traffic accident prediction *Sustainability* (16)22, 9642-9643.

-World Health Organization, (2023).Global status report on road safety.

-Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science 63*, 50-56.

-Hayward, J. C. (1972). Near-miss determination through use of a scale of danger. *Highway Research Record 384.* 24-34.

-Hydén, C. (1987). The development of a method for traffic safety evaluation: *The Swedish traffic conflicts technique Lund University.*

-International Transport Forum, (2024).Road Safety Annual Report.

-International Transport Forum, (2025). Road Safety Annual Report.

-Ke, G., Meng, Q., Finley, T., et al., LightGBM (2017). A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*.

-Li, J., et al., (2025). Traffic safety measures ,and assessment. *Applied Sciences* (15), 19, 10532-10533.

-Liu, R., et al., (2025). A review of traffic flow prediction methods in intelligent transportation ,systems. *Applied Sciences*(15)7, 3866-3867.

-Lord, D., & Mannering, F, (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A 5*(44), 261-305.

-Lundberg, S. M., & Lee, S. I, A. (2017). Unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*.

Modeling and Forecasting Unsafe Following-Distance Violations on Intercity Roads Using Hourly Traffic Count Data: Evidence from Razavi Khorasan Province

Hossein Mohammadi, M.Sc., Grad., Department of Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Iran.

Seyed Ali Ziaee, *Associate Professor, Department of Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Iran.*

Sahar Alishahi, M.Sc., Grad., Department of Civil Engineering, Faculty of Engineering, Hakim Sabzevari University of Sabsevar, Iran.

E-mail: sa-ziaee@um.ac.ir

Received: May 2026- Accepted: November 2026

ABSTRACT

The aim of this study is to provide a data-driven framework for modeling and predicting safe time headway violations on intercity axes using hourly traffic count data from the 141 system. The variable "illegal gap violation," as defined by the system, indicates non-compliance with the minimum two-second time headway from the preceding vehicle and was employed in this study as a safety surrogate measure for proactive risk monitoring. The study domain includes 136 one-directional axes in Razavi Khorasan Province during the year (2025) Persian calendar). Raw twelve-month data were reprocessed and refined according to a quality control framework encompassing operational duration checks, non-positive traffic volume, invalid speed, class count inconsistencies, estimated count discrepancies, and axis–time key duplication. Out of 1, 166, 102 hourly records, (1, 143, 169) records were retained as analysis-ready data. The target variable was defined as the illegal gap violation rate per 1000 vehicles. Results indicated that the network-wide violation rate was 203, 9 violations per 1000 vehicles. For violation rate prediction, data were aggregated at the axis–month–hour–day-of-week level, with months 1 through 9 used for training and months 10 through 12 for testing. The Random Forest model demonstrated strong performance in temporal testing, achieving an R^2 of 0/944, an MAE of 13/82, and an RMSE of *21/24.* Findings highlight the prominent roles of hourly traffic volume, average speed, overtaking violation rate, and inter-axis differences in explaining the gap violation rate. The results of this research can serve as a foundation for designing early warning systems, ranking high-risk axes, and planning managerial interventions in road safety.

Keywords: Traffic Counting, Illegal Gap Violation, Two-Second Time Headway, Safety Surrogate Measure, Machine Learning