فصلنامه مهندسی ترافیک، سال بیست و دوم، شماره ۸۸، تابستان ۱۴۰۱، ص. ۳۵-۲۷

بررسی تاثیر ویژگی‌های رانندگان بر شدت تصادفات عبور از چراغ قرمز – مطالعه موردی اصفهان

علی توکلی کاشانی (مسئول مکاتبات)، استادیار، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، مرکز تحقیقات ایمنی کاربردی حمل و نقل جاده‌ای، دانشگاه علم و صنعت ایران، تهران، ایران

سعیده امیری فر، دانشجوی کارشناسی ارشد مهندسی حمل‌ونقل، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، مرکز تحقیقات ایمنی کاربردی حمل و نقل جاده‌ای، دانشگاه علم و صنعت ایران، تهران، ایران

E-mail: alitavakoli@iust.ac.ir

چکیده

عبور از چراغ قرمز یکی از رایج‌ترین انواع تخلفات در تقاطعات چراغ‌دار است که می‌تواند خسارت‌های جانی و مالی فراوانی را به دنبال داشته باشد. این مطالعه با استفاده از درخت تصمیم CART به شناسایی مهم‌ترین ویژگی‌های رانندگان که بر شدت جراحت آنان در تصادفات عبور از چراغ قرمز شهر اصفهان موثر بوده‌است می‌پردازد. برای این منظور از حدود ده هزار داده‌ی تصادف عبور از چراغ قرمز شهر اصفهان طی پنج سال اخیر بهره گرفته شد. شناسایی عوامل مهم بر مبنای «شاخص اهمیت متغیر» انجام شد و از ماتریس درهم‌ریختگی برای ارزیابی مدل استفاده شد. نتایج نشان داد استفاده از کمربند ایمنی، شغل راننده، تجربه‌ی رانندگی و سن راننده به ترتیب مهم‌ترین عوامل تاثیرگذار در آسیب دیدن رانندگان برای این نوع تصادفات هستند. شناخت این عوامل انسانی می‌تواند در برنامه‌ریزی‌های رسانه‌ای و فرهنگی و بهبود سطح ایمنی کمک شایانی نماید.

واژه‌های کلیدی: تصادفات عبور از چراغ، درخت تصمیم، داده‌کاوی

۱. مقدمه

عبور از چراغ قرمز همه‌ساله موجب کشته و زخمی شدن بسیاری از افراد می‌شود. طبق آمار منتشر شده، رایج‌ترین نوع تصادفات شهری در ایالات متحده مربوط به رانندگانی است که بدون توجه به علائم توقف و چراغ‌های راهنمایی از تقاطع عبور کرده‌اند. هم‌چنین، بین سال‌های ۲۰۰۵ تا ۲۰۱۵، ۸,۳٤۷ نفر در ۷,۵۲۸ تصادف عبور از چراغ کشته شده‌اند. در سال ۲۰۱۵ نیز بطور میانگین روزانه دو نفر در تصادفات عبور از چراغ در امریکا جان باخته‌اند. در ایران، تلفات ناشی از حوادث ترافیکی در شش ماهه‌ی نخست سال ۹۸، ۹۱۰۹ نفر گزارش شده است که این میزان نسبت به سال ۹۷، ۳,۳ درصد رشد داشته‌است. از این رو، مسئله‌ی حاضر تبدیل به یک چالش مهم در کشور شده است. بنابراین شناسایی فاکتورهای موثر در وقوع این نوع تصادفات می‌تواند در تعیین استراتژی‌های مناسب برای بهبود ایمنی راه توسط دولت‌ها کمک شایانی نماید.

اکثر مطالعات به بررسی عوامل انسانی در تصادفات عبور از چراغ پرداخته‌اند که این می‌تواند به دلیل سهم بالای این عوامل در تلفات ترافیکی باشد چرا که عامل انسانی در تعیین الگوی رفتار رانندگی نقش بسزایی دارد. براین اساس، در راستای کاهش خطرات سوانح رانندگی نیازمند شناسایی عوامل موثر در بروز و شدت وقوع تصادف می‌باشیم و این امر خواستار مداخلات پیش‌گیرانه بویژه در خصوص کاستن از عوامل انسانی منجر به تصادف است تا بتوان برنامه‌ریزی مناسبی را در آینده طرح نمود. نتایج یک پژوهش در امریکا نشان داد ۹۳ درصد از تصادفات راه‌ها مربوط به عامل انسانی بوده است. در ایران، پاک‌گهر و همکاران به بررسی رابطه بین عوامل انسانی و شدت تصادفات پرداختند. آنان نتیجه گرفتند عامل انسانی ۹۷,۵ درصد در وقوع تصادفات موثر است.

بسیاری از پژوهش‌ها در خصوص شدت تصادفات از روش‌های داده‌کاوی برای دستیابی به نتایج و تحلیل آن‌ها استفاده نموده‌اند. در مطالعه‌ای که توسط ونگ و همکاران[۱] انجام شد، از لوجیت چندگانه[۲] و الگوریتم جنگل تصادفی[۳] برای شناسایی فاکتورهای دخیل در شدت تصادفات و پیش‌بینی و مقایسه‌ی آن‌ها در ایالت مریلند استفاده شد. وهاب و جیانگ[۴] از سه روش داده کاوی شامل درخت تصمیم[۵]، الگوریتم جنگل و یادگیری بر پایه‌ی نمونه[۶] برای بررسی شدت تصادفات موتور سیکلت در غنا استفاده کردند. در مطالعه‌ی نیز از مدل خطی تعمیم‌یافته‌ی مرکب[۷] و مدل شبکه‌ی عصبی[۸] برای یافتن فاکتورهای موثر در شدت تصادفات تقاطعات در ایتالیا استفاده شد. در ایران نیز توکلی کاشانی و بشارتی از روش درخت کارت به مطالعه‌ی شدت جراحت سرنشینان اتومبیل در راه‌های برون‌شهری پرداختند.

هدف این مطالعه، شناسایی و بررسی ویژگی‌های رانندگان در شدت تصادفات عبور از چراغ قرمز به عنوان یکی از مهم‌ترین علت نامه‌ی تصادفات، با استفاده از الگوریتم طبقه‌بندی و رگرسیون درختی[۹] می‌باشد.

در ادامه، به مروری بر روش درخت تصمیم پرداخته می‌شود، سپس به روش جمع‌آوری داده و ساخت مدل اشاره خواهد شد. در نهایت فاکتورهای معنادار تعیین و بحث خواهند شد.

۲. روش الگوریتم طبقه‌بندی و رگرسیون درختی

درخت‌های طبقه‌بندی و رگرسیونی یا به اصطلاح CART یکی از انواع روش‌های یادگیری ماشین[۱۰] برای ساخت مدل‌های پیش‌بینی کننده از داده‌های موجود هستند که اولین‌بار توسط Breiman و همکاران در سال ۱۹۸٤ مطرح گردید. درخت تصمیم یک مدل خودتوصیف است، یعنی به‌تنهایی و بدون حضور یک فرد متخصص در آن حوزه، نحوه دسته‌بندی را به صورت گرافیکی نشان می‌دهد و به دلیل همین سادگی و قابل‌فهم بودن، روش محبوبی در داده‌کاوی محسوب می‌شود. هم‌چنین دقت آن در مقایسه با بسیاری از روش‌های طبقه‌بندی بالا می‌باشد و قابلیت مدیریت داده‌های با ابعاد زیاد را دارد بنابراین از

الگوریتم درخت تصمیم برای طبقه‌بندی در زمینه‌های بسیاری از جمله پزشکی، تولید و صنعت، تحلیل‌های مالی و غیره استفاده می‌شود. این الگوریتم برای ساخت درخت تصمیم، داده‌ها را به قسمت‌های دوتایی تقسیم کرده و بر اساس آن‌ها درخت دودویی[11] می‌سازد. همان‌طور که از نام آن مشخص است، درخت تصمیم از تعدادی گره[12] و شاخه تشکیل‌شده است به‌گونه‌ای که‌برگ‌ها[13] کلاس‌ها یا دسته‌بندی‌ها را نشان می‌دهند و گره‌های میانی هم برای تصمیم‌گیری با توجه به یک یا چند صفت خاصه به‌کارمی‌روند. این درخت از شاخص افراز جینی[14] به عنوان معیاری برای انتخاب گره‌های درخت استفاده می‌کند. فرآیند رشد درخت تا زمانی ادامه پیدا می‌کند که ماکزیمم اندازه-ی درخت بدست آید و عملیات تقسیم با کمبود داده مواجه نشود. پس از ساخت درخت، عملیات هرس کردن از برگ‌ها به سمت ریشه آغاز می‌گردد. هرس کردن به منظور بهبود دقت طبقه‌بندی داده‌ها توسط درخت انجام می‌شود چرا که درخت طبقه‌بندی بزرگ لزوما درخت مناسبی نیست. باید توجه شود که الگوریتم CART مجموعه‌ای از درخت‌های متوالی را تولید می‌کند و در نهایت یکی از آن‌ها به عنوان درخت بهینه انتخاب می‌گردد. انتخاب درخت بهینه با ارزیابی کارایی آن بر روی داده-های مستقل انجام می‌شود.

در این روش داده‌ها به دو قسمت آموزشی و آزمایشی تقسیم می‌شوند. بر اساس داده‌های آموزشی، درخت مناسب ایجاد می-شود و خطای مربوط به آن با توجه به داده‌های آزمایشی برآورد می‌گردد.

۳. داده‌ها

برای انجام این مطالعه از داده‌های تصادفات سال‌های ۱۳۹۰ تا ۱۳۹۵ شهر اصفهان استفاده شد که از پایگاه داده‌ی پلیس بدست آمده بودند. ابتدا داده‌ها آماده‌سازی شدند، بدین منظور داده‌های نامتعارف و نامشخص حذف شدند. سپس رانندگان درگیر درتصادفات عبور از چراغ فیلتر شدند. در نهایت تعداد ۱۱۲٤۸راننده‌ی درگیر باقی ماندند. متغیرهای مطالعه بر اساس مطالعات پیشین و وجود تعداد نمونه‌ی کافی در سطوح هر متغیر انتخاب شدند. در جدول ۱، متغیرها و توضیحات مربوط به آن‌ها آورده شده است.

همان‌طور که اشاره شد، پیش از ورود داده‌ها به الگوریتم درخت تصمیم، داده‌ها به مجموعه داده‌های آموزشی و آزمایشی تفکیک شدند و بطور تصادفی ۷۰ درصد از نمونه‌ها به داده‌های آموزشی و ۳۰ درصد از آن‌ها نیز به داده‌های آزمایشی اختصاص یافتند.

به منظور دستیابی به هدف مطالعه، با توجه به داده‌هایی که در اختیار بود، ۷ متغیر مربوط به راننده انتخاب شدند که در جدول ۱ نشان داده شده‌اند. شدت آسیب متغیر وابسته و متغیرهای مستقل نیز شامل سن، جنسیت، میزان تجربه‌ی رانندگی، بومی بودن، تحصیلات، استفاده از کمربند ایمنی و شغل راننده می-باشند.

جدول۱. توصیف متغیرها

متغیر	توضیح
شدت آسیب	متغیر هدف: ۱. آسیب دیده ۲. آسیب ندیده
شغل راننده	۱.راننده ۲.کارگر ۳.کارمند ٤.خانه‌دار ۵. نظامی ٦.آزاد ۷.دانشجو ۸.بیکار ۹.سایر ۱۰.نامشخص
میزان تحصیلات راننده	۱.زیر دیپلم ۲.دیپلم ۳.دانشگاهی ٤.نامشخص
جنسیت راننده	۱.مرد ۲.زن

متغیر	توضیح
بومی بودن راننده	۱.بومی ۲.غیربومی
استفاده از کمربند ایمنی	۱.استفاده کرده ۲.استفاده نکرده ۳.نامشخص
تجربه‌ی رانندگی	پیوسته
سن	پیوسته

جدول ۲ تعداد و درصد رانندگان آسیب دیده و آسیب ندیده را به تفکیک تمام متغیرهای انتخاب شده نشان می‌دهد. همان‌طور که ملاحظه می‌شود، بطور کلی، ۱۱۲۴۸ راننده‌ی درگیر در تصادف عبور از چراغ طی سال‌های ۹۰ تا ۹۵ در اصفهان گزارش شده است. هم‌چنین حدود ۱۷ درصد از رانندگان در این تصادفات آسیب دیده‌اند.

جدول ۲. شدت تصادف عبور از چراغ به تفکیک متغیرهای مورد بررسی

متغیر	توضیحات	شدت تصادف		کل
		آسیب دیده (٪)	آسیب ندیده (٪)	
شدت آسیب		۱۸۸۱ (۱۶,۷)	۹۳۶۷ (۸۳,۳)	۱۱۲۴۸
سال	۱۳۹۰	۳۱۹ (۱۶,۶)	۱۶۰۴ (۸۳,۴)	۱۹۲۳
	۱۳۹۱	۳۲۳ (۱۶,۵)	۱۶۳۹ (۸۳,۵)	۱۹۶۲
	۱۳۹۲	۲۹۶ (۱۵,۱)	۱۶۵۹ (۸۴,۹)	۱۹۵۵
	۱۳۹۳	۳۳۶ (۱۵,۹)	۱۷۷۶ (۸۴,۱)	۲۱۱۲
	۱۳۹۴	۳۲۷ (۱۵,۶)	۱۷۶۸ (۸۴,۴)	۲۰۹۵
	۱۳۹۵	۲۸۰ (۲۳,۳)	۹۲۱ (۷۶,۶)	۱۲۰۱
شغل	۱=راننده	۳ (۱۸,۸)	۱۳ (۸۱,۳)	۱۶
	۲=کارگر	۱۷ (۱۶,۵)	۸۶ (۸۳,۵)	۱۰۳
	۳=کارمند	۲۴ (۴,۲)	۵۴۷ (۹۵,۸)	۵۷۱
	۴=خانه‌دار	۲۷ (۵,۳)	۴۸۳ (۹۴,۷)	۵۱۰
	۵=نظامی	۷ (۶,۹)	۹۴ (۹۳,۱)	۱۰۱
	۶=آزاد	۲۹۴ (۵,۱)	۵۴۹۵ (۹۴,۹)	۵۷۸۹
	۷=دانشجو	۸ (۸۰,۰)	۲ (۲۰,۰)	۱۰
	۸=بیکار	۲۱ (۵,۳)	۳۷۶ (۹۴,۷)	۳۹۷
	۹=سایر	۱۱ (۳۳,۳)	۲۲ (۶۶,۷)	۳۳
	۱۰=نامشخص	۱۴۶۹(۳۹,۵)	۲۲۴۹(۶۰,۵)	۳۷۱۸
تحصیلات	۱=زیر دیپلم	۱۵۰ (۹,۸)	۱۳۸۰ (۹۰,۲)	۱۵۳۰
	۲=دیپلم	۶۸۶ (۱۰,۸)	۵۶۵۱ (۸۹,۲)	۶۳۳۷
	۳=دانشگاهی	۳۶ (۴,۹)	۶۹۳ (۹۵,۱)	۷۲۹
	۴=نامشخص	۱۰۰۹(۳۸,۰)	۱۶۴۳(۶۲,۰)	۲۶۵۲

متغیر	توضیحات	شدت تصادف		کل
		آسیب دیده (٪)	آسیب ندیده (٪)	
جنسیت	۱=مرد	۱۷۹٦ (۱۸,۵)	۷۹۲۷ (۸۱,۵)	۹۷۲۳
	۲=زن	۸۵ (۵,٦)	۱٤٤۰ (۹٤,٤)	۱۵۲۵
کمربندایمنی	۱=استفاده کرده	۱۱۹(۱۰,۰)	۱۰۷٦(۹۰,۰)	۱۱۹۵
	۲=استفاده نکرده	۷٦۵(۸۳,۲)	۱۵٤(۱٦,۸)	۹۱۹
	۳=نامشخص	۹۹۷(۱۰,۹)	۸۱۳۷(۸۹,۱)	۹۱۳٤
بومی بودن	۱=بومی	۱٤۷۸ (۱۵,٤)	۸۱۳٤ (۸٤,٦)	۹٦۱۲
	۲=غیربومی	٤۰۳ (۲٤,٦)	۱۲۳۳ (۷۵,٤)	۱٦۳٦

٤. نتایج

تحقیق حاضر از مدل درخت تصمیم برای پیش‌بینی شدت جراحت رانندگان درگیر در تصادفات عبور از چراغ شهر اصفهان استفاده نموده‌است که شامل متغیر وابسته با دو سطح آسیب دیده و آسیب ندیده و ۷ متغیر مستقل مربوط به راننده می‌باشد.

مطابق مدل درخت تصمیم ارائه شده‌ی شکل ۱، در آغاز تمام داده‌ها در گره‌ی ریشه که در ابتدای درخت قرار دارد جمع می‌شوند. سپس درخت بر اساس متغیرکمربند ایمنی به دو دسته تقسیم می‌شود. یعنی بهترین متغیر برای دسته‌بندی و پیش‌بینی شدت تصادف عبور از چراغ، متغیر کمربند ایمنی می‌باشد. گره‌ی ۱ در سمت چپ یک گره‌ی پایانی است و نشان می‌دهد حدود ۸۳ درصد از رانندگانی که از کمربند ایمنی استفاده ننموده‌اند دچار آسیب شده‌اند. بنابراین عدم استفاده رانندگان از کمربند ایمنی احتمال آسیب دیدن آنان را افزایش می‌دهد. در شاخه‌ی سمت راست درخت، چهار گره‌ی پایانی وجود دارد (گره‌های ۳،٦،۷و۸). گره‌ی ۲ مربوط به استفاده‌ی رانندگان از کمربند ایمنی است. این گره بر اساس متغیر شغل راننده به دو گره‌ی دیگر تقسیم می‌شود. مطابق گره‌ی انتهایی ۳، رانندگانی که شغل‌شان کارگر، کارمند، خانه‌دار، نظامی، آزاد، بیکاریا سایر است و از کمربند استفاده ننموده‌اند با احتمال نزدیک به ۹٦ درصد در تصادفات آسیبی نمی‌بینند. هم‌چنین طبق گره‌ی انتهایی ٦، حدود ۸۰ درصد رانندگان با شغل راننده، دانشجو و نامشخص، که بیش از ٦,۵ سال تجربه‌ی رانندگی دارند، در تصادفات آسیبی نخواهند دید این نتیجه می‌تواند به دلیل توجه بیش‌تر رانندگان با تجربه‌تر به قوانین و مقررات راهنمایی و رانندگی و توانایی بیش‌تر آنان درکنترل وسیله نقلیه باشد. در انتها نیز گره‌ی ۵ بر اساس متغیر سن راننده به دو گره‌ی انتهایی ۷ و ۸ تقسیم شده است. مطابق گره‌ی ۷، اگر راننده بیش از ۲٤,۵ سال سن داشته باشد و تجربه‌ی رانندگی‌اش کم‌تر از ٦,۵ سال باشد با احتمال ۸۳,۳ درصد آسیبی نمی‌بینند و در صورتی‌که سن کم‌تر از ۲٤,۵ سال داشته‌باشند با احتمال حدود ۵٤ درصد دچار آسیب خواهند شد (گره‌ی ۸). در مطالعه‌ای نیز تایید شده است که رانندگان با سنین کم در تصادفات آسیب بیش‌تری خواهند دید.

برای ساخت این مدل از نرم‌افزار SPSS Modeler18 استفاده شده است. یکی از خروجی‌های مدل کارت، شاخص اهمیت متغیر(VIM) می‌باشد که برای شناسایی اهمیت متغیرها در دسته‌بندی متغیر هدف انجام می‌شود . اهمیت متغیر x_j به صورت نسبی از رابطه زیر بدست می‌آید:

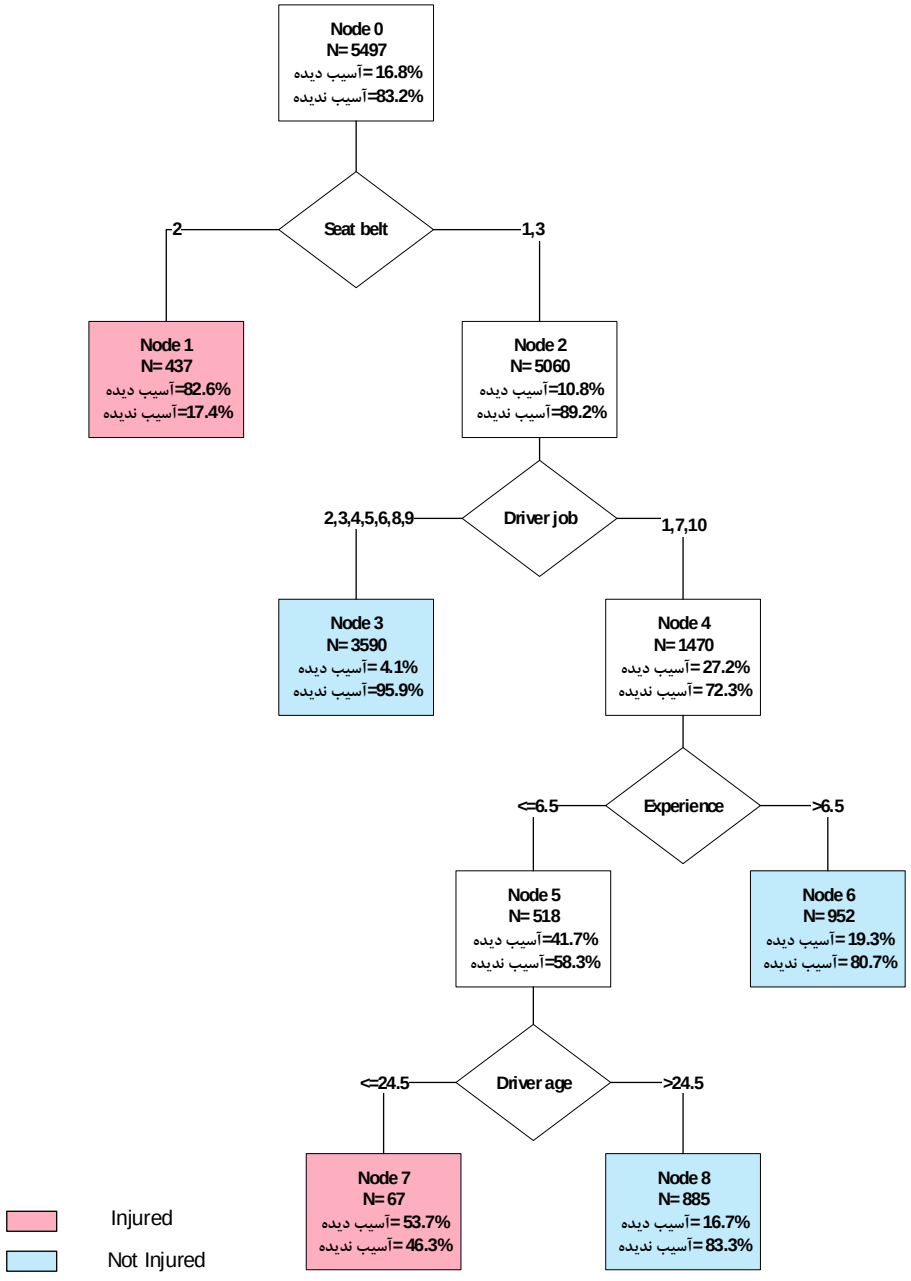
$$VIMx_j = \sum_{m=1}^{M} \frac{n_m}{N} \Delta Gini(S(x_j,m)) \qquad (۱)$$

که در آن $S(x_j,m)$ جداکننده متغیر x_j در محل m و Δ کاهش شاخص جینی در گره m بر اساس متغیر x_j، $\frac{n_m}{N}$ بیان‌گر نسبت مشاهدات در گره m، M تعداد کل گره‌ها و N تعداد کل مشاهدات است. نتایج این شاخص در جدول ۳ نشان داده شده

است. هرچه اهمیت متغیر بیش‌تر باشد، شاخص اهمیت آن نیز بزرگ‌تر است.

جدول ۳. شاخص اهمیت متغیر

متغیر	ضریب VIM
کمربند ایمنی	۰,٦۳
شغل	۰,۲٤
سن	۰,۰٥
تجربه رانندگی	۰,۰۳۱



شکل ۱. درخت تصمیم

٤-١ سنجش عملکرد مدل

جدول ٤ ماتریس درهم‌ریختگی (Confusion Matrix) مدل را نشان می‌دهد. این ماتریس یکی از معیارهای هر الگوریتم دسته‌بندی کننده‌ای می‌باشد و نتایج حاصل از طبقه‌بندی را بر اساس اطلاعات موجود نمایش می‌دهد. هرنمونه در واقعیت متعلق به یکی از کلاس‌های مثبت یا منفی است و از سوی دیگر، از هر الگوریتمی که برای دسته‌بندی داده‌ها استفاده شود، در نهایت هر نمونه عضو یکی از این دو دسته (Class) دسته‌بندی خواهد شد. بنابراین برای هر نمونه داده، یکی از چهار حالتی که در ادامه بیان شده، ممکن است اتفاق بیفتد.

- نمونه عضو دسته مثبت باشد و عضو همین کلاس تشخیص داده شود(مثبت صحیح یاTrue Positive)
- نمونه عضو کلاس مثبت باشد و عضو کلاس منفی تشخیص داده شود(منفی کاذب یاFalse Negative)
- نمونه عضو کلاس منفی باشد و عضو همین کلاس تشخیص داده شود(منفی صحیح یاTrue Negative)
- و در نهایت، نمونه عضو کلاس منفی باشد و عضو کلاس مثبت تشخیص داده شود(مثبت کاذب یا False Positive)

پس از اجرای الگوریتم دسته‌بندی، با توجه به توضیحات و تعاریف ذکر شده، می‌توان عملکرد یک طبقه‌بند را به کمک جدولی به شکل زیر بررسی کرد.

جدول ٤. نمونه ای از ماتریس درهم ریختگی

		پیش‌بینی شده	
		مثبت	منفی
شناخته شده	مثبت	TP	FN
	منفی	FP	TN

در این جدول، آسیب دیدن رانندگان به دسته‌ی *مثبت* و عدم آسیب به دسته‌ی *منفی* اختصاص داده شده است. بنابراین مقادیر مربوطه بصورت جدول ٥ جایگذاری می‌شوند:

جدول ٥. ماتریس درهم ریختگی داده‌ها

		پیش‌بینی شده	
		مثبت	منفی
شناخته شده	مثبت	٥٨٧	٧١٢
	منفی	١٥٦	٦٣٩٥

حال بر اساس این مقادیر می‌توان معیارهای مختلف ارزیابی دسته بند و اندازه‌گیری دقت را تعریف کرد. پارامتر دقت (Accuracy)، متداول‌ترین، اساسی‌ترین و ساده‌ترین معیار اندازه‌گیری کیفیت یک دسته‌بند است و عبارت است از میزان تشخیص صحیح دسته‌بند در مجموع دو دسته. این پارامتر در واقع نشان‌گر میزان الگوهایی است که درست تشخیص داده شده‌اند و بر اساس ماتریس ارائه شده در بالا، به شکل زیر فرموله و تعریف می‌شود:

$$\text{Accuracy} = (TP+TN)/(TP+FN+FP+TN) \quad (٢)$$

پارامتر دقت معمولا به صورت درصد بیان می‌شود. اما پارامترهای دیگری نیز علاوه بر معیار دقت وجود دارند که می‌توان به سادگی از این ماتریس استخراج کرد. یکی از متداول‌ترین آن‌ها، معیار حساسیت (Sensitivity) است که آن را نرخ پاسخ‌های مثبت درست[١٥](TPR) نیز می‌گویند. حساسیت به معنی نسبتی از موارد مثبت است که آزمایش آن‌ها را به درستی به عنوان نمونه مثبت تشخیص داده است. این پارامتر به صورت زیر محاسبه می‌شود:

$$\text{Sensitivity (TPR)} = TP / (TP+FN) \quad (٣)$$

در نقطه مقابل این پارامتر، ممکن است در مواقعی دقت تشخیص کلاس منفی حائز اهمیت باشد. از متداول‌ترین پارامترها که معمولا در کنار حساسیت بررسی می‌شود، پارامتر خاصیت (Specificity)، است که به آن نرخ پاسخ‌های منفی درست[١٦] (TNR) نیز می‌گویند. خاصیت به معنی نسبتی از موارد منفی است که آزمایش آن‌ها را به درستی به عنوان نمونه منفی تشخیص داده است. این پارامتر به صورت زیر محاسبه می‌شود:

$$\text{Specificity (TNR)} = TN / (TN+FP) \quad (٤)$$

پارامترهای حساسیت و خاصیت، بنابر ماهیتی که دارند همواره در رقابت با یکدیگر هستند. یعنی افزایش یکی با کاهش دیگری همراه است و برعکس. همین وضعیت منجر به تولید ابزاری دیگر برای ارزیابی کیفیت دسته‌بندها شده است. منحنی مشخصه عملکرد سیستم[17] (ROC) عبارت است از منحنی که ارتباط بین دو پارامتر حساسیت و خاصیت را بیان می‌کند. محور عمودی این نمودار نشان‌دهنده نرخ مثبت صحیح (Sensitivity)، و محور افقی نشان‌دهنده مقدار نرخ مثبت غلط (One-Specificity) است. نتایج مختلف دسته‌بندی نشانگر نقاط مختلف بر روی این نمودار هستند و در نهایت یک منحنی را تشکیل می‌دهند. مساحت زیر این نمودار(AUC) [18] ، به عنوان یک معیار برای ارزیابی عملکرد دسته‌بند مورد استفاده قرار می‌گیرد. در حالت ایده‌آل، مساحت زیر منحنی برابر با بیشترین مقدار خود، یعنی یک است. بنابراین، هر چه مساحت زیر نمودار به عدد یک نزدیکتر باشد، به معنای بهتر بودن عملکرد دسته‌بند است. هریک از پارامترهای معرفی شده در جدول ٦ محاسبه شده است.

جدول ٦. ارزیابی مدل

Accuracy	TPR	TNR	AUC
٪٨٩	٪٤٥	٪٩٨	۰٫٨٦

٥. جمع‌بندی و نتیجه‌گیری

در این مطالعه به شناسایی و بررسی مهم‌ترین عوامل تاثیر گذار در شدت تصادفات عبور از چراغ شهر اصفهان پرداخته شد. با توجه به نتایج بدست آمده، استفاده از کمربند ایمنی مهم‌ترین عاملی است که در ریسک آسیب دیدن رانندگان موثر می‌باشد و بی‌تردید تاثیر بسزایی در کاهش تلفات و آسیب‌های ناشی از حرکات ناگهانی خودرو دارد و نتایج بدست آمده از مدل نیز آن را تایید می‌کند. عوامل بعدی حائز اهمیت به ترتیب شامل شغل راننده، تجربه‌ی رانندگی فرد و در نهایت سن راننده می‌باشند. شغل می‌تواند نشان دهنده‌ی موقعیت اجتماعی افراد باشد. افرادی که دارای موقعیت اجتماعی بالاتری هستند، احساس همنوایی بیشتری نسبت به قوانین اجرایی جامعه دارند و درواقع رانندگی را هم بخشی از زندگی خود دانســته، به‌طوری‌که خود را ملزم به رعایت قوانین در هنگام رانندگی می‌دانند و با چنین نگرشی کم‌تر دچار ناهنجاری رفتاری از جمله نادیده گرفتن قوانین و مقررات راهنمایی و رانندگی می‌شوند. افزایش احتمال آسیب رانندگان در سنین کم می‌تواند به دلیل زیاد بودن رفتارهای پرخطر، سرعت بالای رانندگی و تجربه‌ی اندک آنان باشد که منجر می‌شود کنترل وسیله نقلیه در تقاطعات به سختی صورت گیرد. طبق مدل ارائه شده، رانندگانی که تجربه‌ی بیش‌تری دارند کم‌تر دچار آسیب‌دیدگی خواهند شد. این نتیجه می‌تواند به دلیل آن باشد که رانندگان کم‌تجربه‌تر معمولا با سرعت بالاتری رانندگی می‌کنند و اغلب از گوشی همراه در حین رانندگی استفاده می‌کنند. درکل، رانندگان باتجربه در موقعیت یکسان با رانندگان کم‌تجربه، محتاط‌تر عمل می‌کنند. بنابراین لازم است در برنامه‌ریزی‌ها ، قوانین ترافیکی و طرح‌های پیش‌گیرانه با تشکیل کارگروه‌های ویژه به این فاکتورها توجه شود.

٦. پی‌نوشت‌ها

1. Wang et al.
2. Multinomial logit
3. Random Forest (RF)
4. Wahab and Jiang
5. Decision Tree
6. Instance-Based learning
7. Generalized Linear Mixed model
8. Neural Network model
9. Classification And Regression Tree (CART)
10. Machine Learning
11. Binary
12. Node
13. Leaf
14. Gini Index
15. True Positive Rate
16. True Negative Rate
17. Receiver Operating Characteristic
18. Area Under Curve

۷. منابع

– آمار متوفیات و مصدومین حوادث رانندگی طی شش ماهه ی اول ۹۸. ۱۳۹۸, سازمان پزشکی قانونی کشور.

– مهیمنی, ا، نیکو,ن. شناسایی مهمترین عوامل موثر در شدت تصادفات موتورسیکلت در جاده های اصلی با استفاده از مدل درخت تصمیم, کنفرانس ترافیک. ۱۳۹۲.

– حیدری, ح، خسروی,م. بررسی رابطه پایگاه اجتماعی اقتصادی و میزان احترام به قوانین راهنمایی و رانندگی. فصلنامه علمی ترویجی راهور, ۱۳۹۵. ۱۳(۳۵).

- FHWA. 2016; Available from: https://safety.fhwa.dot.gov/intersection/conventional/signalized/rlr/.

- Petridou, E. and M. Moustaki, Human factors in the causation of road traffic crashes. European journal of epidemiology, 2000. 16(9): p. 819-826.

- Pakgohar, A., et al., 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, 2011. 3: p. 764-769.

- Wang, X. and S.H. Kim, Prediction and Factor Identification for Crash Severity: Comparison of Discrete Choice and Tree-Based Models. Transportation Research Record, 2019: p. 0361198119844456.

- Wahab, L. and H. Jiang, A comparative study on machine learning based algorithms for prediction of motorcycle crash severity. PloS one, 2019. 14(4): p. e0214966.

- Mussone, L., M. Bassani, and P. Masci, Analysis of factors affecting the severity of crashes in urban road intersections. Accident Analysis & Prevention, 2017. 103: p. 112-122.

- Tavakoli Kashani, A. and M.M. Besharati, An analysis of vehicle occupants' injury severity in crashes occurred on rural freeways and multilane highways in Iran. International Journal of Transportation Engineering, 2016. 4(2): p. 137-146.

- Breiman, l., et al., Classification and regression trees. 1998: CHAPMAN & HALL/CRC.

- Wu, Q., et al., Analysis of driver injury severity in single-vehicle crashes on ruraland urban roadways. Accident Analysis & Prevention, 2016.