

پیش‌بینی و بررسی عوامل تصادفات جاده‌ای با استفاده از الگوریتم ماشین بردار پشتیبان

محسن خواجه‌سلیمی^۱، محمدمهدی خبیری^{۲*} و محمّدصابر فلاح‌نژاد^۳

^۱ کارشناس ارشد راه و ترابری، پردیس فنی و مهندسی، دانشکده مهندسی عمران، دانشگاه یزد

^۲ دانشیار پردیس فنی و مهندسی، دانشکده مهندسی عمران، دانشگاه یزد

^۳ دانشیار پردیس فنی و مهندسی، دانشکده مهندسی صنایع، دانشگاه یزد

(دریافت: ۹۶/۴/۱۸، پذیرش: ۹۶/۱۱/۲، نشر آنلاین: ۹۶/۱۱/۳)

چکیده

شدت جراحات یکی از مهم‌ترین معیارهای سنجش هزینه تصادفات است. شناسایی و اثرسنجی پارامترهای مؤثر بر شدت جراحات تصادفات می‌تواند به عنوان ابزاری جهت افزایش ایمنی در اختیار سیاست‌گذاران قرار گیرد. محققین زیادی در بخش ایمنی، سعی داشته‌اند با تحلیل تصادفات با رویکردهای و مدل‌های مختلف، به بررسی این مهم بپردازند. هدف از این تحقیق بررسی عملکرد الگوریتم ماشین بردار پشتیبان در پیش‌بینی شدت تصادفات جاده‌های بین شهری ایران و تعیین عوامل تأثیرگذار بر شدت جراحات تصادفات است. مدل ماشین بردار پشتیبان تکنیک مدل‌سازی نسبتاً جدیدی در حل مسائل طبقه‌بندی و رگرسیون است که عملکردی دقیق و قابل قبول نشان می‌دهد. در این مطالعه با استفاده از داده‌های تصادفات جاده‌ای پلیس راهور برای هفت استان اصفهان، خوزستان، خراسان جنوبی، قم، قزوین، کرمان و مازندران، با به کارگیری الگوریتم ماشین بردار پشتیبان به پیش‌بینی شدت تصادفات در دو سطح جرحی یا فوتی پرداخته شد. نتایج نشان داد حضور وسیله نقلیه سنگین و خودروی سواری مهم‌ترین عامل در تخمین شدت تصادفات موتورسیکلت‌ها است و احتمال آسیب‌دیدگی را افزایش می‌دهد. در تصادفات مربوط به خودروی سواری عوامل خستگی و خواب‌آلودگی، حضور عابر پیاده، انحراف به چپ خودروی سواری بر شدت تصادفات خودروی سواری تأثیر مستقیم دارد. همچنین در تصادفات مربوط به کامیون‌ها، نتایج بیان می‌کنند انحراف به چپ کامیون، سن بالای راننده و افزایش سرعت خودرو، از مهم‌ترین عوامل افزایش سطح جراحات در این نوع تصادفات است. زمانی که تصادف کامیون به شکل واژگونی اتفاق بیافتد می‌توان انتظار سطح جراحات کم‌تری داشت.

کلیدواژه‌ها: شدت تصادفات، ماشین بردار پشتیبان، ایمنی جاده‌ای، تصادفات فوتی.

۱- مقدمه

به بررسی و مدل‌سازی شدت تصادفات در یک ناحیه محدود (اکثراً یک مسیر) از جاده‌های برون شهری یا درون شهری با استفاده از مدل‌های گسسته رگرسیون لوجیت و یا الگوریتم‌های عصبی تمرکز داشته‌اند. مدل ماشین بردار پشتیبان (SVM) تکنیک مدل‌سازی نسبتاً جدیدی است که برای حل مسائل طبقه‌بندی و رگرسیون ارائه شده است. در سال‌های اخیر، مدل SVM در مطالعات حمل و نقل مانند پیش‌بینی شدت جریان ترافیک، پیش‌بینی تصادفات، مدل‌سازی انتخاب مد سفر و پیش‌بینی تعداد تصادفات مورد استفاده قرار گرفته است. می‌توان یافت که مدل SVM ظرفیت زیادی برای حل مسائل طبقه‌بندی دارد. بنابراین می‌توان از آن برای مدل‌سازی شدت جراحات تصادفات استفاده کرد (Zhibin و همکاران، ۲۰۱۲). لیکن هدف اصلی این تحقیق

وقوع تصادف و نتایج آن می‌تواند جبران‌ناپذیر باشد. سالیانه آمار و اطلاعاتی در رابطه با نرخ تصادفات جرحی، فوتی و غیرفوتی توسط مراکز پلیس راهور و پزشک قانونی اعلام می‌گردد. آمار و ارقام حاکی از آن است که علی‌رغم اتخاذ خط و مشی‌های مختلف در راستای کاهش تصادفات، هنوز هم تصادفات در بخش‌های مختلف شبکه راه‌ها، خسارات سنگینی در پی دارد. بنابراین در صورت بررسی شدت جراحات تصادفات و عوامل مؤثر بر آن، می‌توان از شدت خسارات وارده کاست.

با کمک چنین رویکردی می‌توان با شناخت بالقوه متغیرهای تأثیرگذار و نحوه عملکردشان، گام مهم و مؤثری در کاهش شدت تصادفات جاده‌ای برداشت. اکثر مطالعات انجام شده قبلی بیشتر

* نویسنده مسئول؛ شماره تماس: ۰۳۵۳۱۲۳۲۴۶۱

رضایی مقدم و همکاران (۲۰۱۱) در بررسی شدت تصادفات، از مدل شبکه عصبی استفاده نمودند. آنها انواع مختلف شبکه‌های عصبی پرسپترون چندلایه و جلو رونده تعمیم یافته را برحسب تعداد لایه‌های پنهان و تعداد نرون‌های هر لایه مورد امتحان قرار داده و بهترین مدل را بر مبنای دقت بالاتر ارائه دادند. دقت پیش‌بینی بهترین مدل در این مطالعه در حدود ۸۵ درصد گزارش شده است که با استفاده از ۲۵ متغیر توصیفی دارای اهمیت صورت گرفته است.

Chong و همکاران (۲۰۰۵) با استفاده از مدل شبکه عصبی به بررسی شدت تصادفات بر حسب شدت آسیب‌دیدگی راننده در پنج سطح پرداختند. متغیرهای مستقل این مطالعه شامل خصوصیات راننده، ویژگی‌های خودرو و اطلاعات مربوط به تصادف بوده و علاوه بر شبکه عصبی مرسوم، مدل‌های ترکیب شده با شبکه عصبی نیز آزمایش شده‌اند. نتایج نشان داد که مدل نهایی مشاهدات مربوط به هر پنج دسته از شدت تصادفات را با دقت بالای ۵۸ درصد پیش‌بینی کرده است.

به طور کلی کاربرد مدل‌های شبکه عصبی با توجه به قابلیت بالا در برقراری رابطه بین تعداد بالایی از متغیرها و به خصوص روابط پیچیده نامشخص، می‌تواند بسیار مفید باشد. از طرفی به نظر می‌رسد علاوه بر زمان بر بودن، مشکل اصلی این مدل‌ها ارائه یک چهارچوب در بسته‌ای از محاسبات است که امکان محاسبات درون شبکه‌ای و دخالت در روند پردازش مدل را از محقق می‌گیرد. این مشکل تا به آنجا ادامه دارد که حتی این مدل‌ها قادر به تحلیل حساسیت آماری نبوده و معمولاً اهمیت نسبی پارامترها را تخمین نمی‌زنند.

Martin و همکاران (۲۰۱۴) برای شناسایی رابطه بین قطعات خطرناک راه و تعداد تصادفات، داده‌های جمع‌آوری شده را به کمک داده‌کاوی تحلیل و بررسی کردند که منجر به شناسایی عوامل مؤثر بر وقوع تصادفات شد.

Dong و همکاران (۲۰۱۵) استفاده از مدل‌های ماشین بردار پشتیبان را در پیش‌بینی شدت تصادفات پیشنهاد دادند. به این منظور برای بررسی عملکرد SVM^۱، مدل‌ها با مدل‌های دوگانه بیزی مقایسه شدند. نتایج نشان داد که مدل‌های SVM نتایج دقیق‌تر و مناسب‌تری ارائه می‌دهند، بهترین مدل پیش‌بینی کننده حاصل مدل SVM با کرنل RBF^۲ بود.

Abdel-Aty و همکاران (۲۰۱۳) طی مطالعه‌ای به ایجاد مدل‌های آنالیز شدت جراحات تصادفات برای مقاطع آزادراهی پرداخت. شدت جراحات تصادفات به صورت خروجی دوتایی (تصادفات شدید و تصادفات غیرشدید) در نظر گرفته شد و با استفاده از مدل جنگل تصادفی، متغیرهای توصیفی مهم مرتبط با شدت تصادفات

معطوف به بررسی و مدل‌سازی شدت تصادفات به وقوع پیوسته در جامعه آماری بسیار بزرگ‌تر و محدوده وسیع‌تر در سطح چند استان نمونه کشور بوده است که بر مبنای روشی نسبتاً جدید در داده کاوی انجام خواهد شد. فرآیند بررسی عوامل مؤثر در شدت انواع مختلف تصادف در چند بخش، به شرح زیر صورت خواهد گرفت: ابتدا از مدل ماشین بردار در فرآیند ساخت مدل‌های شدت تصادف برای سه نوع تصادف موتورسیکلت، سواری، کامیون‌ها به صورت جداگانه استفاده می‌شود. نتایج مدل‌های حاصل، در مشخص شدن نقش متغیرهای توصیفی تعریف شده در شدت تصادفات مورد بررسی قرار می‌گیرند. سپس به بررسی و تشریح عملکرد مدل ماشین بردار پشتیبان در پیش‌بینی شدت تصادفات پرداخته می‌شود.

۲- پیشینه پژوهش

بخش زیادی از مطالعات انجام شده در زمینه ایمنی ترافیک، به بررسی تصادفات پرداخته است. در مطالعات پیشین، تصادفات بیشتر در دو بخش تعداد تصادفات و شدت جراحات ناشی از تصادف مورد بررسی قرار گرفته است. با این وجود، از آنجا که شدت جراحات معیاری از هزینه تصادفات بر جامعه است، بیشتر مطالعات انجام شده، به این جنبه از تصادفات توجه داشته‌اند.

در این بخش به مرور مطالعاتی پرداخته می‌شود که شدت تصادف را با استفاده از مدل‌های مختلف مورد بررسی قرار داده‌اند. معین‌الدینی و حبیبیان (۱۳۹۴) در پژوهشی جراحات ناشی از تصادفات وسایل نقلیه سنگین را مورد بررسی قرار دادند. و با استفاده از داده‌های تصادفات جاده‌ای پلیس راهور به مدل‌سازی لوجیت دوگانه برای پیش‌بینی شدت‌های جراحات تصادفات، پرداختند. پس از بررسی نتایج به دست آمده از مدل‌سازی و اثر متغیرها بر روی شدت تصادفات با ضرائب معنادار، بیان کردند که عواملی همچون خستگی و خواب‌آلودگی بر احتمال ایجاد شدت جراحات فوتی می‌افزاید.

راضی اردکانی و همکاران (۱۳۹۰) براساس اطلاعات تصادفات بزرگراه‌های درون شهری تهران با هدف بررسی تأثیر جنسیت در پارامترهای مؤثر بر شدت تصادفات، تحقیقی انجام دادند. این متغیر وابسته در دو سطح جرحی و فوتی بررسی شد. نتایج نشان داد که سن کم‌تر از ۲۵ سال و بیشتر از ۵۰ سال، بی‌سوادی رانندگان، در رانندگان مرد دو عامل تخلف عمدی و عجله و همچنین تصادف در روزهای تعطیلی، تصادف در شب، تصادفات در آب و هوای بارانی، حجم ترافیک و میانگین سرعت ۲۰ تا ۴۰ کیلومتر برای هر دو جنس، شدت تصادف را افزایش داده‌اند.

مدل پرابیت ترتیبی عملکرد بهتری در پیش‌بینی شدت جراحات تصادفات داشته است.

آقایان و همکاران (۲۰۱۵) برای ارزیابی عملکرد مدل‌های ماشین بردار پشتیبان با توابع کرنل مختلف در پیش‌بینی شدت تصادفات مطالعاتی انجام دادند. سه سطح مصدومیت برای این مطالعه تعریف شد (بدون جراحات، جرحی، فوتی) و دوازده متغیر از داده‌های موجود انتخاب شد. از آنجا که پارامترهای طرح هندسی جاده در دسترس نبود، برای تمام تصادفات در مجموعه داده به صورت مشترک رعایت گردید. از مدل‌سازی SVM برای پیش‌بینی شدت جراحات تصادفات استفاده شد که دقت پیش‌بینی آن با آن با نتایج مدل‌های الگوریتم ژنتیک و الگوریتم ژنتیک چند لایه و الگوریتم پرسپترون چندلایه مقایسه شد. و نتایج نشان داد که برای پیش‌بینی شدت جراحات تصادفات، عملکرد الگوریتم پرسپترون چندلایه حدود ۵ درصد بهتر از عملکرد SVM است. با این وجود، ماشین بردار پشتیبان به علت استفاده مستقیم از الگوریتم در فاز یادگیری، شامل هزینه‌های محاسباتی پایین‌تری نسبت به پرسپترون چند لایه است. درصد صحت پیشگویی برای مدل پرسپترون چندلایه ۸۶٫۲٪ بود، که بیشتر از مدل ماشین بردار پشتیبان با کرنل چندجمله‌ای (۸۱٫۴٪)، مدل الگوریتم ژنتیک ترکیبی (۷۸٫۶٪) و مدل الگوریتم ژنتیک (۷۸٫۱٪) بود. دسته‌بندی دو سطحی ماشین بردار پشتیبان (بدون خسارت و فوتی/جرحی) - با درصد ۸۵٫۳٪- نتایج دقیق‌تری نسبت به دسته‌بندی چندکلاسی ۸۱٫۴٪ ارائه داد.

در تحقیقی دیگر Chen و همکارانش (۲۰۱۶) با هدف ارزیابی الگوی شدت جراحات رانندگان با استفاده از مدل‌های ماشین بردار پشتیبان، اطلاعات دو ساله تصادفات واژگونی در نیومکزیکو را مورد مطالعه قرار دادند و متغیرهای توصیفی از شرایط محیطی، خصوصیات وسایل نقلیه و راننده تعریف شد که با استفاده از درخت تصمیم متغیرهای تأثیرگذار تعیین شدند. پرداخت و تست مدل‌های SVM با کرنل گوسی (RBF) و چند جمله‌ای حاکی از برتری دقت کرنل RBF است. نتایج نشان داد شرایط محیطی، استفاده از کمربند ایمنی، مصرف مشروبات الکلی، زمان تصادف تأثیر قابل توجهی بر شدت جراحات راننده دارد. در جدول (۱) جمع‌بندی کلی از مطالعات مرتبط با این پژوهش ارائه شده است. مطالعات بررسی شده، روش‌های مختلفی را برای مدل‌سازی مورد استفاده قرار داده‌اند. علاوه بر آن تحقیقات دیگری با هدف مقایسه و تعیین روش مناسب جهت مدل‌سازی شدت تصادفات انجام شده بود که در راستای شناسایی و بررسی دقت روش‌های متداول در زمینه پیش‌بینی شدت تصادفات مورد مطالعه قرار گرفتند. این روش‌ها شامل مدل‌های شبکه عصبی و ماشین بردار

(فصل برفی، درجه سراسیعی، انحراف معیار سرعت و دما) به عنوان ورودی تعیین شدند. برای تعیین رابطه واقعی بین شدت تصادفات و متغیرهای توصیفی و افزایش انطباق مدل با واقعیت سه مدل مورد بررسی قرار گرفت:

۱- مدل لوجیت با ضرایب ثابت

۲- مدل ماشین بردار پشتیبان با تابع کرنل گوسی RBF

۳- مدل لوجیت با ضرایب تصادفی

با مقایسه این سه مدل، مشخص شد مدل SVM نسبت به مدل‌های لوجیت رابطه غیرخطی بین متغیرهای توصیفی و متغیرهای وابسته را به صورت بهتر و دقیق‌تری نشان می‌دهد. یافته‌ها نشان می‌دهد متغیرهای شرایط جوی و ترافیک، تأثیر مهمی بر روی شدت جراحات تصادفات دارد که می‌تواند برای پیش‌بینی شدت تصادفات استفاده شوند.

Li و همکاران (۲۰۰۸) در جهت ارزیابی توابع عملکرد ایمنی برای تصادفات موتورسیکلت، پتانسیل مدل SVM را بررسی کردند. نتایج نشان داد مدل‌های SVM در مقایسه با مدل‌های دو جمله‌ای منفی^۳ عملکرد بهتر و مناسب‌تری دارند.

Cheu و Yuan (۲۰۰۳) SVM را در تعیین شدت تصادفات به کار بردند و نتایج مدل‌های SVM را با مدل‌های شبکه عصبی مقایسه کردند که دقت مدل‌های SVM در تعیین شدت تصادفات نشان داده شد. همچنین Chen و همکاران (۲۰۰۹) در تحقیقی به بررسی قابلیت عملکردی مدل‌های ماشین بردار پشتیبان برای پیش‌بینی تصادفات پرداختند. با توجه به ماهیت مقایسه‌ای تحقیق، مدل‌سازی براساس یکی از معروف‌ترین و پرکاربردترین داده‌های تصادفات آزادراهی که در مطالعات قبلی مورد توجه محققین قرار گرفته بود انجام شد که شامل تصادفات مقاطع هفت مایلی آزادراهی در سال ۱۹۹۳ در منطقه سانفرانسیسکو کالیفرنیا بود. الگوریتم در نرم‌افزار MATLAB نوشته شد که با آن آموزش SVM و دسته‌بندی انجام شد. با هدف بهبود دقت، به جای یک مدل SVM، ترکیب چند مدل SVM مختلف را پیشنهاد دادند.

علاوه بر تعیین و پیش‌بینی تصادفات، از مدل‌های SVM در مطالعات فراوانی تصادفات نیز استفاده شده است. Zhou و Ren (۲۰۱۱) از SVM به عنوان روشی بهینه در پیش‌بینی تعداد تصادفات استفاده کردند. Zhibin و همکاران (۲۰۱۲) برای تحلیل شدت تصادفات در پنج سطح شدت جراحات (بدون جراحات، جراحات احتمالی، جراحات سطحی، جراحات شدید و فوتی) - مدل‌های SVM را به کار بردند و برای تعیین مدل SVM از نرم‌افزار LIBSVM استفاده کردند. در این تحقیق با مقایسه دقت مدل SVM و مدل پروبیت ترتیبی، نتایج نشان داد که مدل SVM نسبت به

$$\text{Min}(\omega, \omega) + C \sum_i \varepsilon_i^\delta, \forall \delta \geq 0 \quad (۳)$$

$$\text{st: } y_i[(\omega \cdot x_i) + b] \geq 1 - \varepsilon_i, \forall \varepsilon_i \geq 0 \quad (۴)$$

در میان این ابرصفحه‌ها یک ابرصفحه با بیش‌ترین حاشیه وجود دارد که به عنوان ابرصفحه جداکننده بهینه شناخته می‌شود. این ابرصفحه با توجه به بردارهای حاشیه‌ای تعیین می‌شود که به این بردارها بردار پشتیبان گفته می‌شود. الگوریتم ماشین بردار پشتیبان رابطه (۳) با در نظر گرفتن محدودیت مورد اشاره در رابطه (۴) را بهینه می‌کند. در حقیقت، تابع هدف (رابطه (۳)) کمینه کردن نرم ضرایب ابرصفحه به علاوه خطای وزن دار است و محدودیت (رابطه (۴)) درستی دسته‌بندی ابر صفحه با اجازه خطای ε_i است.

با توجه به این که شرایط کوهن- تاکر برقرار است با استفاده از روش لاگرانژ این مسئله به فرم زیر در می‌آید:

$$W(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j (h(x_i), h(x_j)) \quad (۵)$$

که شامل محدودیت‌های زیر است:

$$\sum_i \alpha_i y_i = 0 \quad (۶)$$

$$0 \leq \alpha_i \leq C \quad (۷)$$

و معادله ابرصفحه جدا کننده به صورت زیر است:

$$f(x) = \sum_{i=1}^N \alpha_i y_i (h(x_i), h(x_j)) \quad (۸)$$

در هر دو معادله بالا $h(x)$ ها ورودی مستقیم هستند. می‌توان با دانستن تابع کرنل فرم $h(x)$ را تغییر داد:

$$K(x_i, x_j) = (h(x_i), h(x_j)) \quad (۹)$$

بنابراین با تعریف تابع کرنل می‌توان بر کارایی دسته‌بندی افزود (Chen و همکاران، ۲۰۰۹).

در این مطالعه از کرنل RBF در بسته libsvm در نرم‌افزار Matlab 2015b استفاده شده است. این کرنل به صورت رابطه (۱۰) است.

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (۱۰)$$

۴- معرفی داده‌ها

اطلاعات مورد استفاده در این مطالعه شامل تصادفات جاده‌ای ثبت شده در حوزه راهنمایی رانندگی هفت استان (اصفهان، قم، قزوین، خراسان جنوبی، کرمان، مازندران و خوزستان) می‌باشد. که

پشتیبان می‌باشد. نتایج این بررسی حاکی از دقت بالا و پیچیدگی مدل‌های شبکه عصبی، ماشین بردار پشتیبان (SVM) در پیش-بینی تصادفات می‌باشد. همچنین نتایج نشان داد مدل ماشین بردار پشتیبان SVM با کرنل گوسی RBF در مقایسه با دیگر مدل‌ها همواره از دقت و برتری برخوردار است.

۳- مبانی نظری و روش مدل سازی

مدل ماشین بردار پشتیبان (SVM) (رضانی، ۱۳۹۲) تکنیک مدل‌سازی نسبتاً جدیدی است که برای حل مسائل طبقه‌بندی و رگرسیون ارائه شده است. اصول نظری مربوط به این الگوریتم‌ها را می‌بایست در تئوری یادگیری آماری جستجو کرد. یکی از ویژگی‌های قابل توجه الگوریتم‌های مذکور آن است که در اغلب موارد هیچ حساسیتی را نسبت به ابعاد داده‌ها از خود نشان نمی‌دهند. در ادامه، مفاهیم و معادلاتی که تشکیل دهنده شالوده این الگوریتم‌ها هستند ارائه شده است. که در سال‌های اخیر، در جنبه‌های مختلف مطالعات حمل و نقل به کار گرفته می‌شود. می‌توان یافت که مدل SVM ظرفیت زیادی برای حل مسائل طبقه-بندی دارد. بنابر این می‌توان از آن برای مدل‌سازی شدت جراحات تصادفات استفاده کرد (Zhibin و همکاران، ۲۰۱۲).

الگوریتم ماشین بردار پشتیبان بر اساس نظریه یادگیری آماری و در چهارچوب کمینه کردن خطای دسته‌بندی طراحی شده است. این الگوریتم بر اساس پیدا کردن ابرصفحه‌ای جدا کننده، سعی می‌کند فاصله اعدادی که اشتباه دسته‌بندی شده‌اند تا مرز تصمیم را کمینه کند.

با توجه به تعریف دو سطح شدت جراحات (فوتی، جرحی)، برای مسئله دسته‌بندی داده‌های یادگیری به صورت مجموعه رابطه (۱) تعریف می‌شوند:

$$(x_1, y_1), \dots, (x_i, y_i), y_i \in -1, 1 \quad (۱)$$

که در آن، y_i متغیر هدف برای نمونه مشاهده شده i ام است. فرض می‌شود که برای تصادفات فوتی، $y_i = 1$ و برای تصادفات غیرفوتی $y_i = -1$ است. همچنین فرض می‌شود x_i بردار متغیرهای توصیفی برای نمونه مشاهده شده i ام است.

ماشین بردار پشتیبان به دنبال یافتن ابرصفحه‌ای در فضای n بعدی است (n تعداد متغیرهای توضیحی در یک ردیف داده است) به نحوی که داده‌های آموزشی را به دو دسته تصادفات فوتی و جرحی تقسیم کند. ماشین بردار پشتیبان سعی می‌کند که تابع ابرصفحه $f(x, \alpha) \in f(x, \alpha_0)$ را بیابد به طوری که بهترین تخمین از تابع $y = f(x)$ باشد. معادله این ابرصفحه به صورت زیر نوشته می‌شود:

$$f(x, \alpha) = (\omega_\alpha \cdot x) + b \quad (۲)$$

۴-۱- اطلاعات ثبت شده در هر گزارش تصادف

● سن راننده: این متغیر بیانگر سن راننده مقصر در تصادف می- باشد. سن عابر پیاده و راننده در بسیاری از مطالعات بررسی شده پیشین، به صورت یک متغیر پیوسته دیده شده است. همواره این موضوع مورد بررسی است که ممکن است گروه‌های مختلف سنی، تأثیر متفاوتی بر شدت جراحات تصادفات داشته باشند. که این می‌تواند بر مبنای معیارهایی چون توانایی جسمی، رفتارهای اجتماعی و بلوغ ذهنی- جسمی راننده، عابر و سرنشین قابل توجیه باشد. Demetriades و همکاران (۲۰۰۴) دسته‌بندی سن راننده را پیشنهاد داده‌اند، با این پیشنهاد که بر مبنای معیارهایی چون توانایی جسمی، رفتارهای اجتماعی و بلوغ ذهنی- جسمی راننده صورت پذیرفته است سن راننده در پنج دسته کم‌تر از ۱۶ سال، ۱۶ تا ۲۴ سال، ۲۵ تا ۵۴ سال، ۵۵ تا ۷۴ سال و بزرگ‌تر از ۷۴ سال تقسیم‌بندی می‌شود.

● خستگی و خواب‌آلودگی: همان‌طور که در فصل دو بیان شد، خستگی و خواب‌آلودگی بر عملکرد راننده تأثیر منفی دارد. این متغیر به عنوان تعریف کننده یکی از خصوصیات راننده بیان شده است. این معیار تحت مجموعه دلایل تصادفات ثبت می‌شود.

● تخطی از سرعت مجاز: سرعت وسایل نقلیه تأثیر غیرقابل انکاری در شدت تصادف دارد، در هیچ کدام از تصادفات ثبت شده و اطلاعات موجود، مقدار سرعت تخمینی و یا سرعت لحظه‌ای قید نشده است در نتیجه تنها پارامتری که تأثیر سرعت بر شدت تصادف را بیان می‌کند تخطی از سرعت مجاز می‌باشد. همان‌طور که از اسم آن برمی‌آید، این پارامتر بیان کننده اتخاذ سرعتی بیشتر از سرعت مجاز است. البته با توجه به ماهیت صفر و یک معرفی شدن داده‌ها در مدل‌ها، مشکلی رخ نمی‌دهد.

● نوع خودرو: با توجه به این که نوع خودرو یکی از مهم‌ترین عوامل تأثیرگذار در شدت جراحات تصادفات است، خودروها بسته به نوعشان تفکیک شده‌اند. هرچند متغیر عمر خودرو می‌تواند، به عنوان معیاری از فن‌آوری به کار رفته در آن مورد استفاده قرار گیرد، این ویژگی در داده‌ها موجود نبود. جدول (۲) نمایشی توصیفی از حضور وسایل نقلیه مختلف در آمار تصادفات مورد بررسی است.

جدول ۲- درصد مشارکت کلاس خودروهای مختلف در تصادفات

۵۰/۵۶	خودروی سواری
۱۷/۰۶	موتورسیکلت
۱۳/۳۸	کامیون
۱۱/۳۲	وانت
۴/۹۷	عابر پیاده
۱/۹۵	اتوبوس
۰/۴۳	شاسی بلند
۰/۰۵	ون

بازه زمانی چهار ساله (۱۳۹۳-۱۳۹۰) را در بر می‌گیرد. که مجموعاً شامل ۴۰۴۲۹ مشاهده تصادف می‌شود. هرکدام از تصادفات ثبت شده، حاوی اطلاعاتی نظیر زمان وقوع تصادف، تعداد مجروح و کشته، نوع وسیله مقصر، علت تصادف، سن راننده مقصر، نوع محور و شرایط آب و هوایی می‌باشد که در جدول (۲) نمایش داده شده- اند. در این مجموعه داده نوع خسارت به صورت تعداد جرحی یا فوتی بودن سرنشینان ثبت شده است. در بانک اطلاعاتی مورد مطالعه، آسیب‌دیدگی سرنشینان در قالب دو دسته تعداد سرنشینان مجروح و متوفی ثبت شده است. با توجه به این که روش دسته‌بندی دوگانه و چندگانه شدت تصادفات بر اساس بیش- ترین شدت آسیب‌دیدگی سرنشینان وضعیت کلی حاکم بر تصادف را بیان می‌کند لذا امکان بررسی و مدل‌سازی تصادفات در سطح بالاتر و کلی‌تر نسبت به بررسی تک‌تک سرنشینان را فراهم می- کند.

از این رو در این تحقیق، بیش‌ترین شدت آسیب‌دیدگی سرنشینان به عنوان معیار مشخص کننده شدت تصادف انتخاب شد و به تقسیم تصادفات، در دو سطح تصادفات جرحی (تصادفاتی که منجر به جراحات حداقل یک سرنشین می‌گردد همچنین هیچ- کدام از سرنشینان فوت نشوند) و تصادفات فوتی (تصادفاتی که منجر به فوت حداقل یک سرنشین می‌گردد)، می‌پردازد.

جدول (۱) آماری کلی از تمامی تصادفات ثبت شده در بانک اطلاعاتی را طی دوره چهار ساله نشان می‌دهد بدین صورت که تعداد تصادف فوتی در برابر تصادف جرحی را در هر سال بیان می‌کند. همان‌طور که مشاهده می‌شود از تعداد کل تصادفات در سال ۱۳۹۰ تعداد ۱۸۲۴ تصادف فوتی و شناخته شده است نشان می‌دهد. به طور کلی تعداد کل تصادفات ثبت شده در دوره چهار سال برای هفت استان برابر با ۴۰۴۲۹ تصادف است که ازین تعداد ۶۹۶۲ تصادف منجر به فوت شدند به عبارتی حدود ۱۷ درصد از تصادفات ثبت شده فوتی و ۸۳ درصد جرحی می‌باشند. این تصادفات منجر به ۸۹۴۰ کشته و ۵۱۰۵۲ مجروح شده‌اند.

جدول ۱- آمار چهارساله نوع تصادفات ثبت شده برای کل

سال	تعداد تصادفات			تعداد مجروحی متوفی	
	کل	جرحی	فوتی	مجروح	متوفی
۱۳۹۰	۸۷۰۰	۶۸۷۶	۱۸۲۴	۱۰۶۹۱	۲۴۰۶
۱۳۹۱	۱۰۸۹۳	۸۸۴۶	۲۰۴۷	۱۳۶۶۴	۲۶۸۶
۱۳۹۲	۱۰۱۰۹	۸۸۰۶	۱۳۰۳	۱۲۹۲۶	۱۷۶۴
۱۳۹۳	۱۰۷۲۷	۹۲۰۵	۱۵۲۲	۱۳۷۹۱	۲۰۸۴

۵- مدل‌سازی

در این مطالعه از الگوریتم ماشین‌بردار پشتیبان‌بررسی عوامل مؤثر و پیش‌بینی شدت تصادفات و مدل‌سازی داده‌های موتور، خودروهای سواری و کامیون پرداخته می‌شود. با توجه به مطالب بیان شده در قسمت دوم این تحقیق استفاده از کرنل گوسی RBF در مدل‌سازی ماشین‌بردار پشتیبان همواره بهترین نتایج را ارائه داده است (Dong و همکاران، ۲۰۱۵؛ Aghayan و همکاران، ۲۰۱۵) لذا کرنل در مدل‌سازی به کار برده شد.

برای ساخت مدل SVM پیشنهاد شده است که اطلاعات مورد استفاده در قالب سه دسته جهت آموزش^۴، تست^۵ و اعتبارسنجی^۶ تقسیم شوند. مدل براساس اطلاعات بخش اول پرداخت می‌شود (آموزش می‌بیند). اطلاعات دسته دوم در جهت تعیین و تنظیم ضرایب ثابت C و σ به کار می‌رود هدف از انجام این عملیات تنظیم ضرایب ثابت به صورتی است که مدل پرداخت شده کم‌ترین خطای ممکن را داشته باشد و نهایتاً با استفاده از اطلاعات بخش سوم مدل پرداخت شده مورد اعتبارسنجی قرار می‌گیرد. نسبت این دسته‌بندی به این صورت است که ۷۰ درصد داده‌ها جهت آموزش، ۱۵ درصد داده‌ها جهت تست و ۱۵ درصد دیگر داده‌ها در اعتبارسنجی استفاده می‌شوند.

۵-۱- مدل ماشین‌بردار پشتیبان شدت تصادفات

موتورسیکلت

در این قسمت داده‌های مربوط به تصادفات که موتورسیکلت‌ها در آن حضور داشتند. جهت ساخت مدل به الگوریتم SVM معرفی شدند. مدل با ۷۰ درصد داده‌ها آموزش داده شد. در مرحله تست، به روش شبکه‌ای، ضرایب ثابت C و σ به گونه‌ای تنظیم شدند که مدل کم‌ترین خطای ممکن را ارائه دهد. که مقادیر بهینه و مطلوب C و σ به ترتیب برابر با ۰/۰۱ و ۱۴ به دست آمد. نهایتاً برادر ابرصفحه دسته‌بندی کننده تشکیل شد، که وزن پارامترها در جدول (۴) نمایش داده شده است.

با توجه به جدول (۳) ضریب مثبت متغیر حضور وسیله نقلیه سنگین hgv نشان می‌دهد که حضور وسایل نقلیه سنگین احتمال آسیب‌دیدگی جراحات تصادف در تصادفات موتورسیکلت را افزایش می‌دهد به عبارتی اگر در تصادفی موتورسیکلت با کامیون برخورد کند احتمال منجر شدن تصادف به فوت، قوت می‌گیرد. همچنین ضریب مثبت متغیر حضور وسیله نقلیه سواری Sedan نیز بیان می‌کند که حضور وسایل نقلیه سواری، شدت جراحات را در تصادفات موتوری افزایش می‌دهد.

همان‌طور که مشاهده می‌شود در بیشتر از ۵۰ درصد از تصادفات ثبت شده، خودروی سواری حضور داشته است. همچنین تصادفات موتورسیکلت و کامیون‌ها از لحاظ تعداد در رتبه‌های بعدی قرار دارند.

● خصوصیات مسیر: به محدود بودن اطلاعات ثبت شده در فرم ثبت تصادفات، تنها عامل وابسته به جاده، نوع مسیر است، در این داده‌ها دسته‌بندی نوع مسیر توسط پلیس راهور به سه دسته اصلی، فرعی و روستایی انجام شده است.

همچنین پارامترهای ثبت شده دیگر مثل واژگونی بعد از برخورد، وجود شی ثابت، وجود عابر پیاده و وجود ماشین سنگین از پارامترهایی هستند که می‌توان به عنوان متغیر مستقل معرفی کرد.

● واژگونی بعد از برخورد: این متغیر بیانگر این است که آیا وسایل نقلیه حاضر در تصادف دچار واژگونی شده‌اند.

● وجود شی ثابت: این متغیر نشان دهنده وضعیتی است که وسیله نقلیه شریک در تصادف به شی ثابتی برخورد کند.

● حضور عابر پیاده: این متغیر بیان کننده حالتی است که در تصادف عابر پیاده حضور داشته باشد.

● حضور کامیون: این متغیر بیان کننده حالتی است که در تصادف کامیون حضور داشته باشد.

جدول ۳- معرفی نام و نوع متغیرها

متغیر	نوع
شدت جراحات	وابسته
سن	مستقل
خواب آلودگی	مستقل
نوع خودرو	مستقل
آب و هوا	مستقل
روشنایی	مستقل
نوع مسیر	مستقل
انحراف به چپ	مستقل
تغییر مسیر ناگهانی	مستقل
حرکت در خلاف جهت	مستقل
عبور از محل ممنوعه	مستقل
عدم توانایی در کنترل وسیله	مستقل
عدم توجه به جلو	مستقل
عدم رعایت حق تقدم	مستقل
عدم رعایت فاصله طولی	مستقل
نقض ماده ۴ آیین‌نامه راه‌ها	مستقل
تخطی از سرعت مجازنقص فنی	مستقل
واژگونی	مستقل
وجود شی ثابت	مستقل
حضور عابر پیاده	مستقل
حضور وسیله نقلیه سنگین	مستقل

ضرایب متغیر سن راننده (more74, to55, less16) بیان می- کنند در دو رده سنی زیر ۱۶ سال و بالای ۷۴ سال احتمال جراحت نسبت به رانندگان با سن کمتر از ۵۵ سال، بیشتر است، که می- توان ضعف جسمانی را دلیلی بر این نتیجه بیان کرد.

۵-۲- مدل ماشین بردار پشتیبان شدت تصادفات خودروی سواری

داده‌های مربوط به تصادفات که خودروی سواری در آن‌ها حضور داشتند جهت ساخت مدل SVM استفاده شدند. در اینجا نیز مدل با ۷۰ درصد داده‌ها آموزش داده شد. در مرحله تست، با استفاده از ۱۵ درصد مشاهدات و به روش شبکه‌ای، ضرایب ثابت C برابر با ۰/۰۵ و σ برابر با ۱۰ تنظیم شدند. از ۱۵ درصد باقی‌مانده اطلاعات جهت اعتبارسنجی استفاده شد. جدول (۵)، اوزان بردار ابرصفحه دسته‌بندی کننده را نشان می‌دهد.

جدول ۵- وزن متغیرهای مستقل در مدل SVM تصادفات خودروی سواری

متغیر	ضریب	بازه اطمینان ۹۵٪	
Pedestrian	۱۷/۳۵۱	۱۶/۹۰۷	۱۹/۶۲۰
hgv	۳۳/۷۱۷	۳۳/۲۶۸	۳۵/۲۴۰
MTL	۳۴/۲۲۵	۳۱/۶۹۰	۳۳/۸۷۰
SM	-۸/۰۲۴	-۷/۷۱۳	-۷/۰۳۱
TAS	۲۳/۶۰۶	۲۲/۱۶۳	۲۴/۱۳۰
DCD	-۸/۰۱۳	-۸/۸۲۳	-۸/۲۰۳
CRR	۸/۸۱۴	۸/۳۷۰	۹/۴۴۷
SR	۸/۵۵۸	۳/۰۷۳	۵/۴۶۳
N	-۰/۶۴۳	-۰/۰۴۲	۱/۶۸۱
MW	۷/۱۷۷	۵/۳۷۳	۶/۶۹۵
BW	-۱/۷۴۰	-۱/۴۹۷	-۰/۶۰۱
Less16	-۹/۷۵۰	-۹/۰۹۱	-۶/۵۶۹
to24	-۴/۸۱۳	-۷/۰۲۹	-۵/۴۰۷
to55	-۶/۸۲۸	-۶/۳۲	-۴/۵۶۷
more74	۸/۳۵۸	۷/۹۳۳	۹/۷۸۰
bias	-۰/۷۷۳	-۰/۸۰۷	-۰/۷۳۴

با توجه به جدول (۵) ضریب Pedestrian بیان می‌کند که حضور عابر پیاده از عواملی است که در افزایش سطح شدت جراحت تصادف تأثیرگذار است اگر در تصادفات بین جاده‌ای خودروهای سواری، عابر پیاده حضور داشته باشد این تصادف با احتمال بیشتری منجر به فوت می‌شود. از طرفی ضریب hgv با علامت مثبت نشان دهنده تأثیر مستقیم حضور وسایل نقلیه سنگین در سطح شدت جراحت تصادفات خودروهای سواری با وسایل نقلیه سنگین است. این بدان معناست اگر در جاده‌های بین

جدول ۴- متغیرها و ضریب وزنی پارامترهای به کار رفته

متغیر	ضریب	بازه اطمینان ۹۵٪	
hgv	۲۴/۰۲۲	۲۳/۵۳۱	۲۴/۳۸
Sedan	۳/۴۲۹	۳/۹۷۶	۵/۰۳۰
MTL	۱۱/۶۵۶	۱۰/۸۸۱	۱۱/۷۷۱
SM	-۳/۲۷۲	-۳/۹۹۱	-۳/۲۸۶
S	۶/۱۷	۵/۸۵۱	۶/۵۸۰
SR	۵/۴۸۰	۶/۲۴۰	۷/۴۳۰
D	-۳/۲۷۰	-۳/۳۳۹	-۲/۳۱
SS	-۴/۰۹۱	-۴/۳۰۵	-۳/۶۵۰
RW	-۷/۷۰۰	-۷/۷۱۰	-۷/۰۲۲
Less16	-۲/۶۷۳	-۳/۴۲۲	-۲/۸۳۱
to55	-۹/۱۱۵	-۸/۷۵۰	-۷/۹۵۰
more74	-۲/۸۰	۱/۹۰۸	۲/۶۸۰
bias	-۰/۱۳	-۰/۱۳۲	-۰/۱۱۶

این دو ضریب بیانگر این مسئله است که افزایش وزن وسیله نقلیه در تصادفات مربوط به موتورسیکلت، افزایش شدت جراحت تصادف را در پی دارد.

نتیجه ای که از ضریب متغیر انحراف به چپ MTL بر می‌آید این است که اگر تصادفی بین موتورسیکلت و سایر وسایل نقلیه مانند سواری، شاسی‌بلند و غیره اتفاق افتد و دلیل این تصادف انحراف به چپ آن وسیله نقلیه (غیر موتورسیکلت) باشد احتمال منجر به فوتی شدن این تصادف افزایش می‌یابد. همچنین ضریب منفی متغیر مربوط به حرکت ناگهانی SM در امتداد حرکت نشان می‌دهد اگر تصادفی بین موتورسیکلت و سایر وسایل نقلیه رخ دهد که در آن دلیل تصادف حرکت ناگهانی وسیله نقلیه غیر موتورسیکلت باشد احتمال منجر به فوت شدن این تصادف کم می‌شود. همان‌طور که انتظار می‌رود ضریب متغیر تخطی از سرعت مجاز S حاکی از افزایش احتمال جراحت است. به این معنا که اگر در تصادف موتورسیکلت، دلیل تصادف تخطی از سرعت مجاز دیگر وسایل نقلیه باشد احتمال فوتی بودن تصادف افزایش می‌یابد.

ضریب SR نشان می‌دهد که در زمان طلوع خورشید، احتمال این که تصادفات موتورسیکلت به فوت ختم شود بسیار بالاست. اما با توجه به ضرایب D و SS این احتمال در طول روز و زمان غروب خورشید کاهش می‌یابد و انتظار به جرحی بودن تصادف می‌رود.

با توجه به ضریب منفی RW می‌توان اذعان داشت چنانچه تصادف موتورسیکلت در جاده‌های روستایی اتفاق افتد انتظار به شدت جراحت کم‌تری می‌رود که به نظر می‌رسد دلیل وضعیت نامناسب روسازی و مشخصات هندسی راه برای اخذ سرعت‌های بالا است.

مقدار ۰/۰۶ به عنوان بهترین مقدار برای C و مقدار ۱۰ به عنوان بهترین مقدار برای σ به روش شبکه‌ای، به گونه‌ای تنظیم شدند که مدل کم‌ترین خطای ممکن را ارائه دهد. نهایتاً بردار ابرصفحه دسته‌بندی کننده تشکیل شد، که وزن پارامترها در جدول (۶) نمایش داده شده است.

جدول ۶- وزن متغیرهای مستقل در مدل SVM تصادفات کامیون

متغیر	ضریب	بازه اطمینان ۹۵٪	
Rollover	-۷/۸۵۳	-۸/۵۷۷	-۷/۸۰۱
Pedestrian	۱۲/۷۰۸	۱۱/۹۱۴	۱۳/۱۱۱
Motorcycle	۵/۷۲۸	۵/۱۶۴	۶/۲۵۳
Sedan	-۱۳/۲۱۰	-۱۱/۸۱۰	-۱۰/۷۸۰
MTL	۲۸/۸۸۰	۲۸/۰۴۴	۲۹/۱۰۷
SM	-۱۰/۹۳۰	-۱۱/۱۰۲	-۱۰/۱۷۰
TAS	۴/۰۷۳	۴/۹۳۰	۶/۱۱۳
DCD	-۷/۸۹۸	-۸/۳۶۰	-۷/۷۱۰
S	-۴/۲۷۸	-۴/۰۱۲	-۳/۰۹۷
to55	-۱۰/۳۰	-۱۱/۳۰	-۱۰/۲۳
more74	۵/۲۶۰	۶/۳۳۹	۷/۳۲۹
bias	-۰/۰۰۳	-۰/۰۰۳	-۰/۰۰۲

جدول (۶) عوامل مؤثر بر شدت تصادفات وسایل نقلیه سنگین را نشان می‌دهد. ضریب منفی متغیر واژگونی Rollover بیان می‌کند که اگر در تصادف وسیله نقلیه سنگین واژگونی رخ دهد شدت جراحت تصادف کاهش می‌یابد و احتمال فوتی بودن تصادف کم‌تر می‌شود. ضریب متغیر حضور عابر پیاده Pedestrian نشان می‌دهد که در تصادفات وسایل نقلیه سنگین در صورت حضور عابر پیاده، احتمال فوتی بودن تصادف قوت می‌یابد. با توجه به ضریب موتورسیکلت Motorcycle می‌توان ادعان داشت که اگر در تصادف کامیون‌ها، موتورسیکلت درگیر باشد احتمال فوتی بودن تصادف افزایش می‌یابد. همچنین از ضریب منفی متغیر مربوط به سواری Sedan می‌توان نتیجه گرفت که اگر وسیله مقصر سواری باشد تصادف با احتمال کم‌تری منجر به فوت می‌شود.

از ضریب مثبت متغیر مربوط به انحراف به چپ MTL می‌توان این‌چنین نتیجه گرفت که اگر تصادف کامیون با سایر وسایل نقلیه صورت گیرد، به طوری که کامیون مقصر باشد احتمال منجر به فوت شدن این تصادف افزایش می‌یابد. ضریب منفی متغیر مربوط به حرکت ناگهانی SM در امتداد حرکت نشان‌دهنده این امر است که اگر تصادفی بین کامیون و سایر وسایل نقلیه رخ دهد که در آن دلیل تصادف حرکت ناگهانی وسیله نقلیه کامیون باشد احتمال منجر به فوت شدن این تصادف کم می‌شود. با توجه به ضریب مثبت متغیر خواب‌آلودگی TAS می‌توان نتیجه گرفت که اگر تصادفی بین کامیون و سایر وسایل نقلیه صورت گیرد در صورتی

شهری تصادفی بین خودروهای سواری و وسایل نقلیه سنگین رخ دهد، افزایش شدت جراحت فوتی را به دنبال دارد.

از ضریب متغیر انحراف به چپ MTL می‌توان نتیجه گرفت که اگر تصادفی بین وسایل نقلیه سواری و سایر وسایل نقلیه مانند کامیون، شاسی‌بلند و غیره اتفاق افتد و دلیل این تصادف انحراف به چپ باشد احتمال منجر به فوتی شدن این تصادف قوت می‌گیرد. همچنین ضریب منفی متغیر مربوط به تغییر مسیر ناگهانی SM در امتداد حرکت نشان‌دهنده این امر است که اگر تصادفی بین وسایل نقلیه سواری و سایر وسایل نقلیه رخ دهد که در آن دلیل تصادف حرکت ناگهانی وسیله نقلیه سواری باشد احتمال منجر به فوت شدن این تصادف کم می‌شود.

با توجه به ضریب مثبت متغیر خواب‌آلودگی TAS می‌توان نتیجه گرفت که اگر تصادفی بین وسیله نقلیه سواری و سایر وسایل نقلیه صورت گیرد در صورتی که دلیل تصادف خواب‌آلودگی راننده باشد این تصادف با احتمال بیشتری منجر به فوت می‌شود.

همچنین از ضریب منفی متغیر عدم رعایت فاصله طولی DCD می‌توان نتیجه گرفت اگر عدم رعایت فاصله طولی توسط وسایل نقلیه سواری دلیل تصادف آنها با دیگر وسایل نقلیه باشد، احتمال منجر به فوتی شدن تصادف کم می‌شود. در این مدل با توجه به ضریب مثبت متغیر نقض ماده ۴ آیین‌نامه ایمنی راه‌ها CRR می‌توان دریافت که اگر دلیل تصادف عدم توجه به جلوی راننده وسیله نقلیه سواری باشد، احتمال منجر به فوت شدن تصادف افزایش می‌یابد.

ضرایب مثبت SR و N بیان می‌کنند که از نظر زمانی، اگر تصادفی در شب هنگام یا زمان طلوع خورشید رخ دهد این تصادف با احتمال بالاتری منجر به فوت می‌گردد. ضریب مثبت MW نشان می‌دهد که تصادفات وسایل نقلیه سواری در جاده‌های اصلی، احتمال فوتی بودن سرنشینان را افزایش می‌دهد. ضریب منفی BW بیان می‌کند که تصادفات وسیله نقلیه سواری در جاده‌های فرعی، نسبت به جاده‌های اصلی منجر به جراحات کم‌تر می‌شوند. سن همواره یکی از عوامل تأثیرگذار بر سطح شدت جراحت تصادفات بوده است، ضرایب Less16، to24، to55، more74 بیان می‌کنند که افزایش سن، موجب افزایش احتمال آسیب‌دیدگی سرنشینان می‌شود.

۴-۵- مدل ماشین بردار پشتیبان شدت تصادفات کامیون

در این قسمت با استفاده از داده‌های مربوط به تصادفات وسایل نقلیه سنگین مدل ماشین بردار پشتیبان کامیون‌ها تهیه شد. مدل با ۷۰ درصد داده‌ها آموزش داده شد. برای تعیین ضرایب ثابت C و σ مدل با ۱۵ درصد از داده‌ها مورد تست قرار گرفت،

شدند که در افزایش احتمال فوتی بودن تصادفات نقش مستقیم دارند.

۷- مراجع

اردکانی ح، حاتمی ع، نجف پ، "بررسی تأثیر جنسیت رانندگان در عوامل مؤثر بر شدت تصادفات درون شهری"، یازدهمین کنفرانس بین‌المللی مهندسی حمل و نقل و ترافیک، تهران، ۱۳۹۰.

معین‌الدینی ا، حبیبیان م، "بررسی عوامل مؤثر بر شدت جراحت حادثه دیدگان در تصادفات کامیون‌ها"، پانزدهمین کنفرانس بین‌المللی مهندسی حمل و نقل و ترافیک، ۱۳۹۴.

Abdel-Aty M, Yu R, "Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data", Safety Science, 2013.

Aghayan I, Hadji Hosseinlou M, Kunt M, "Application of Support Vector Machine for Crash Injury Severity Prediction: A Model Comparison Approach", Journal of Civil Engineering and Urbanism, 2015.

Chen C, Zhang G, Qian Z, Rafiqul A, Tian Z, "Investigating driver injury severity patterns in rollover crashes using support vector machine models", Accident Analysis and Prevention, 2016.

Chen S, Wang W, Van Z, "Construct support vector machine ensemble to detect traffic incident", Expert Systems with Applications, 2009.

Chong M, Abraham A, Paprzycky M, "Traffic Accident Data Mining Using Machine Learning Paradigms", Informatica, 2005.

Dong N, Huang H, Zheng L, "Support vector machine in crash prediction at the level of traffic analysis zones: Assessing the spatial proximity effects", Accident Analysis and Prevention, 2015.

Li X, Lord D, Zhang Y, "Predicting Motor Vehicle Crashes Using Support Vector Machine Models", Accident Analysis and Prevention, 2008.

Martín L, Baena L, Garach L, López G, Oña J, "Using data mining techniques to road safety improvement in Spanish roads", Procedia- Social and Behavioral Sciences, 2014.

Moghaddam F, Afandizadeh Sh, Ziyari M, "Prediction of Accident Severity Using Artificial Neural Networks", International Journal of Civil Engineering. 2011, 9 (1), 41-49.

Ren G, Zhou Z, "Traffic safety forecasting method by particle swarm optimization and support vector machine", Expert Systems with Applications, 2011.

Yuan F, Cheu R, "Incident detection using support vector machines", Transportation Research Part C. 2003.

Zhibin L, Pan L, Wang Wei, Chengcheng Xu, "Using support vector machine models for crash injury severity analysis", Accident Analysis and Prevention, 2012, 478-486.

که دلیل تصادف خواب‌آلودگی راننده کامیون باشد این تصادف با احتمال بیشتری منجر به فوت می‌شود. همچنین از ضریب منفی متغیر رعایت فاصله طولی DCD می‌توان نتیجه گرفت اگر عدم رعایت فاصله طولی توسط کامیون دلیل تصادف آن با سایر وسایل نقلیه باشد، احتمال منجر به فوتی شدن تصادف کم می‌شود.

ضریب مثبت متغیر مربوط به تخطی از سرعت مجاز S نشان-دهنده تأثیر مستقیم افزایش سرعت بر افزایش سطح جراحت تصادف می‌باشد. یعنی اگر دلیل تصادف کامیون، تخطی از سرعت مجاز باشد احتمال فوتی بودن تصادف قوت می‌گیرد. ضرایب متغیرهای مربوط به سن راننده نشان می‌دهد که در تصادفات رانندگان با سن بیشتر از ۷۴ سال احتمال فوتی بودن تصادف افزایش می‌یابد.

همان‌گونه که پس از بررسی منابع مطالعاتی در بخش دوم انتظار می‌رفت، حضور وسایل نقلیه سنگین در معابر، همواره از میزان ایمنی ترافیکی می‌کاهد و منجر به مصدومیت‌های شدیدتری در تصادفات می‌گردد (معین‌الدینی و حبیبیان، ۱۳۹۴). که این به دلیل اختلاف وزنی وسایل درگیر در تصادفات است چرا که افزایش فاصله وزن دو خودروی درگیر در تصادفات، احتمال مرگ راننده خودروی سبک‌تر افزایش می‌یابد. از طرفی انحراف به چپ از جمله مهم‌ترین عوامل منجر به افزایش شدت تصادفات می‌باشند (Zhibin و همکاران، ۲۰۱۲)، که در نتایج مدل‌سازی همواره یکی از تأثیرگذارترین پارامترها بر افزایش شدت تصادفات شناخته شد. نقش سن راننده در شدت تصادفات تمامی مدل‌ها مؤثر بود به طوری که با بالا رفتن سن راننده، شدت تصادفات افزایش می‌یابد (Chong و همکاران، ۲۰۰۵).

۶- نتیجه‌گیری

نتایج این مطالعه نشان می‌دهد، در حالت‌های مختلف تصادفات وسایل نقلیه مختلف (خودروی سواری، موتور سیکلت و کامیون) عوامل متفاوتی دخیل هستند که الزاماً همسو نیستند. این امر می‌تواند نشانه درستی تفکیک تصادفات به لحاظ ماهیت در این مطالعه باشد.

در تصادفات مربوط به موتورسیکلت‌ها، عواملی به ترتیب اولویت حضور وسایل نقلیه سنگین و خودروی سواری، تخطی از سرعت مطمئنه و افزایش سن راننده احتمال فوتی بودن تصادف را افزایش می‌دهد. در تصادفات خودروی سواری، نقض قانون ۴ آیین‌نامه ایمنی راه‌ها، خستگی و خواب‌آلودگی، حضور عابر پیاده و وسایل نقلیه سنگین افزایش سطح شدت جراحت تصادفات را افزایش می‌دهد. همچنین در تصادفات وسایل نقلیه سنگین به ترتیب اولویت، سن بالای راننده، انحراف به چپ کامیون‌ها، تخطی از سرعت مجاز و حضور موتورسیکلت به عنوان عواملی شناسایی

EXTENDED ABSTRACT

Prediction and Investigation of Road Traffic Accident Severity Factors using Support Vector Machine Algorithm

Mohsen Khajahsalimi^a, Mohammad Mehdi Khabiri^{a,*}, Mohammad Saber Fallah Nezhad^b

^a Faculty of Civil Engineering, Yazd University, Yazd, 89195741, Iran

^b Department of Industrial Engineering, Yazd University, Yazd, 89195741, Iran

Received: 09 July 2017; Accepted: 22 January 2018

Keywords:

Severity of crashes, Support vector model, Road safety, Fatal crashes, Kernel algorithm.

1. Introduction

Accidents as a threat factor to the transport system have widespread political, social, and economic dimensions, which is increasing in developing countries. Iran, as a developing country, has not escaped this danger, but in recent years it has taken preventive measures and crash statistics has been decreasing. Crash injury severity is one of the most important criterions of measuring the costs of accidents. Different methods have been used for predict modeling. Support Vector Machine is a relatively new modeling technique which was proposed to solve the classification and regression problems. That shows accurate and acceptable performance. In the present study, it has been tried to model the severity of accidents with a combination clustering and classification approach, with the help of neural network algorithms, simple parsing, SVM, KNN, and C4.5 algorithms and by comparing the algorithms used assess the ability of each of the algorithms in the prediction of the severity of accidents. The purpose of this study is to evaluate the performance of support vector machine algorithm to predict the severity of road accidents and identify the factors that affect the severity of accidents.

2. Methodology

2.1. Data collection

Several approaches have been used to model and predict the severity of accidents. The support vector model is a relatively new modeling technique for solving regression and classification problems, which provides a precise and acceptable operation. The purpose of this study was to evaluate the performance of backup vector machine algorithm in predicting the severity of road accidents and determining the factors influencing the severity of accidents. In this study, road traffic crash data for seven provinces of Isfahan, Khuzestan, South Khorasan, Qom, Qazvin, Kerman and Mazandaran has been used, and support vector machine algorithm has been used to predict the injuries or casualties.

2.2. SVM modeling

Ren and Zhou used the SVM as an optimal way to predict the number of accidents (Ren, and Zhuping, 2011). Lee et al., applied the SVM models to analyze the severity of accidents at five levels of injury severity (no injury, possible injury, superficial injury, severe injury and fatal injury), and to determine the SVM model, they used LIBSVM software (Zhibin et al., 2012), as well. Aghaeian et al. (2015), conducted a study to evaluate the performance of Support Vector Machines models with various kernel functions in case of predicting the severity of accidents (Aghayan et al., 2015). Three levels of injury were defined for this study (no injury,

* Corresponding Author

E-mail addresses: mohsen.khajehsalimi@gmail.com (Mohsen Khajahsalimi), mkhabiri@yazd.ac.ir (Mohammad Mehdi Khabiri), Fallahnezhad@yazd.ac.ir (Mohammad Saber Fallah Nezhad).

wounding injury and fatal injury) and twelve variables were selected from existing data. In another study, Chen et al., aiming to assess the model of severity of drivers injuries using the of Support Vector Machines models, studied the two-year data of reversal accidents in New Mexico, and descriptive variables of environmental conditions, characteristics of vehicle and driver were defined then, using the decision tree, the effective variables were determined (Chen et al., 2016). The data related to accident in which a riding car existed was used to build the SVM model. At this point, the model was trained by 70% of data as well. At the testing stage, using 15 percent of the observations through a network approach, the constant coefficient of C was determined 0.05 and σ was set equal to 10. The remaining 15 percent of information was used for validation.

3. Results and discussion

3.1. SVM to predict the severity of car crashes

According to Fig. 1, the Pedestrian Coefficient suggests that the presence of pedestrians is one of the factors affective in increase in the severity level of the accident injury. If in road traffic accidents of riding cars, pedestrians exist, these accidents are more prone to death. On the other hand, HCAR Coefficient with a plus sign indicates the direct impact of the presence of heavy vehicles on the level of severity of injury accidents of riding cars with heavy vehicles. This means that if throughout interurban roads, an accident happens between cars and heavy vehicles; it will lead to increase in the severity of fatal injuries.

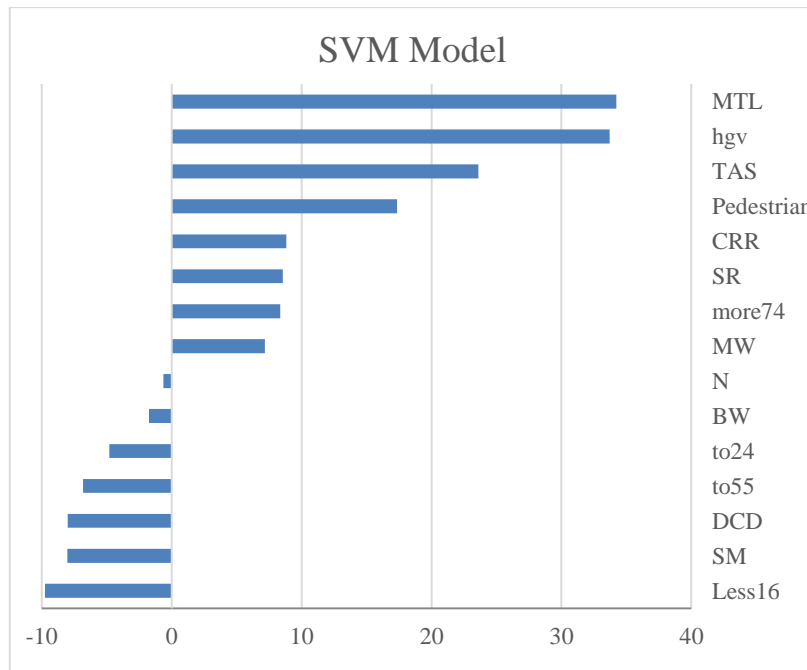


Fig. 1. Weighting coefficients of independent variables in the model of severity of car accidents

3.2. Identify important variables

The data related to crashes in which the car was present were used to build the SVM model, the effect of important variables on the severity of accidents, each of the three types of accidents related to motorcycles, cars and trucks were investigated. After comparing the results obtained from the support vector machine model, the results indicated that backup vector machine algorithm has the Performance superiority, higher accuracy, and low operating error. Age has always been one of the factors affecting the severity of accidents. Coefficients Less16, to24, to55, and more74 indicate that increasing age increases the likelihood of injuries to occupants, are presented in Table 1.

Table 1. Weighting coefficients of independent variables in the truck crash intensity model

Parameter	coefficients	A 95% confidence interval	
Pedestrian	17.351	16.907	19.62
Hgv	33.717	33.268	35.24
MTL	34.225	31.69	33.87
SM	-8.024	-7.713	-7.031
bias	-0.733	-0.807	-0.734

4. Conclusions

The results of the research indicate that algorithms such as neural networks and support vector machine algorithms (often non-linear degrees) usually provide the highest rate of classification, while they do not provide the proper interpretation capability, instead, tree-based algorithms are rule-based if and then have the interpretation capability for humans. Rules are important because they can be made available to traffic and safety professionals, and by taking a strategic decision in the field of reducing the severity of the accidents and the resulting losses, has taken an important and effective step. The results showed that the data volume does not affect the performance of the backup vector machine and can consistently predict more than 70% correctly for Gang types of accidents, injuries and casualties. In other words, due to the high number of crash injuries in comparison with the crash casualties, there is a strong tendency to anticipate injury crashes. Also, factors such as the difference in the weight of vehicles involved in crashes, deviation of the vehicles towards left, elderly age of the driver, pedestrian presence and violating the legal speed limit lead to increased accident severity and fatal accidents.

5. References

- Aghayan I, Hadji Hosseinlou M, Metin Kunt M, "Application of support vector machine for crash injury severity prediction: A model comparison approach", *Journal Civil Engineering Urban*, 2015, 5, 193-199.
- Gang R, Zhou Z, "Traffic safety forecasting method by particle swarm optimization and support vector machine", *Expert Systems with Applications*, 2011, 38 (8), 10420-10424.
- Ren G, Zhou Z, "Traffic safety forecasting method by particle swarm optimization and support vector machine", *Expert Systems with Applications*, 2011.
- Zhibin L, Pan L, Wang Wei, Chengcheng Xu, "Using support vector machine models for crash injury severity analysis", *Accident Analysis and Prevention*, 2012, 478-486.