پیش‌بینی شدت تصادفات جاده‌ای با استفاده از شبکه عصبی مصنوعی و مقایسه آن با روش آنالیز چند متغیره

مقاله پژوهشی

محمدرضا کی منش*، استادیار، گروه عمران، دانشکده فنی و مهندسی، دانشگاه پیام نور، تهران، ایران

نصیر برادران رحمانیان، دانشجوی دکتری، دانشکده فنی و مهندسی، دانشگاه پیام نور، تهران، ایران

*پست الکترونیکی نویسنده مسئول: mrkeymanesh@pnu.ac.ir

دریافت: ۹۹/۰۵/۲۷– پذیرش: ۹۹/۱۰/۰۵

صفحه ۱۰-۱

چکیده

پیش بینی شدت و یا تعداد تصادفات در صورتی که با متدلوژی های دارای مبانی علمی انجام گیرد بدون شک می تواند به عنوان ابزاری مهم در مدیریت ایمنی راهها و مهندسی حمل و نقل محسوب شود. در این مقاله روشی برای شناسایی و پیش بینی وضعیت ایمنی از حیث شدت تصادفات در راههای بین شهری ارائه شده است. این روش مبتنی بر یک تحقیق تجربی است که در یک کریدور مهم و شریانی واقع در استان خراسان رضوی انجام شده است. تجزیه و تحلیل حوادث رخ داده مربوط به سال‌های ۱۳۹۲ تا ۱۳۹۵ و داده‌هایی است که از گزارش پلیس راه بدست آمده است. سایر اطلاعات از جمله داده‌های هندسی از طریق برداشت میدانی و نیز بهره‌گیری از سیستم اطلاعات جغرافیایی و اطلاعات ترافیکی از اداره مربوطه در استان یاد شده جمع آوری شده است. داده ها در سیستم GIS، ثبت و ذخیره شد و با استفاده از یک سری روش‌های آماری پردازش گردید و نتایج از دو مدل مورد بررسی به دست آمد. مدل۱ از طریق آنالیز چند متغیره (MVA)[1] و مدل۲ با استفاده از تکنیک شبکه عصبی مصنوعی(ANN)[2] ایجاد گردید. با مقایسه این دو مدل ، مشخص شد که مدل ۲ از مدل ۱ دارای نتایج بهتری است به این دلیل که جمع کل مقدار باقیمانده که تفاوت بین میزان واقعی مشاهده شده و میزان پیش بینی شده توسط مدل است در مدل ۲ کمتر است. اگرچه که به نظر می‌رسد مدل ۱ در تخمین مقاطع پرتصادف خطرناک تر(دارای تعداد زیاد فقره تصادف) کارایی و عملکرد بهتری دارد.

واژه‌های کلیدی: پیش بینی، شدت تصادفات، شبکه عصبی مصنوعی، آنالیز چند متغیره

۱- مقدمه

در سال‌های اخیر، هوش مصنوعی و محاسباتی کاربردهای فزاینده ای در تحلیل‌ها و پیش بینی‌ها و بطور کلی مدیریت زیرساخت‌های حمل و نقلی پیدا کرده است. بسیاری از محققان، روشهای شبکه عصبی مصنوعی (ANN) را برای تحلیل عوامل مرتبط با پردازش داده‌ها به کار گرفته‌اند. در بررسی پیشینه و ادبیات موضوع مشخص شد که در بسیاری از تحقیقات که به مسائل مربوط به ایمنی راه‌ها از جمله عوامل رفتاری، سرعت و جریان ترافیک پرداخته است (Vujanić et al.2013) از چنین تکنیک‌هایی استفاده شده است. در تحقیقی دیگر(Mario De Luca.2015) قدرت محاسباتی دو تکنیک شبکه عصبی و آنالیز چند متغیره در مدیریت ایمنی راه مورد بررسی و مقایسه با یکدیگر قرار گرفت و نشان داد که مدل بدست آمده از شبکه عصبی، تخمین بهتری را برای نقاط پرتصادف نشان می‌دهد. در این تحقیق درجه قوس افقی، شیب طولی، وضعیت روسازی، عرض راه و نیز شهری یا غیر شهری بودن موقعیت تصادف به عنوان متغیرهای مستقل و حاصل ضرب تعداد تصادف در شدت آن به عنوان متغیر وابسته درنظر گرفته شده بود. مطالعه دیگر

(Chiou.2006) از ANN برای توسعه سیستم تخصصی ارزیابی تصادفات استفاده کرده است. نتایج نشان می‌دهد که مدل ANN می‌تواند به دقت بالایی از ٨٥,٧٢٪ در آموزش و ٧٧,٩١٪ دراعتبارسنجی برسد در مقابل، مدل بیزین شوارتز دارای دقت پایین با میزان‌های ٠/٨٢- در آموزش و ٠,١٣ در اعتبار سنجی است که نشان می‌دهد مدل ANN برای ارزیابی تصادفات مناسب‌تر است. علاوه بر این، در تکنیک مذکور به منظور سنجش اهمیت هر متغیر توصیفی، یک شاخص تأثیرگذار کلی بر اساس وزنهای آموزش دیده ANN محاسبه می‌شود. بررسی تحقیقی دیگر (Chang.2005) نیز نشان می‌دهد که مدل رگرسیون دوتایی پواسون یا دو جمله‌ای منفی می‌تواند برای تجزیه و تحلیل تکرار تصادف برای تعداد سال‌های بسیاری مورد استفاده قرار گیرد. هرچند این مدل‌ها رابطه اساسی از پیش تعریف شده‌ای بین متغیرهای وابسته و مستقل دارند. اگر این فرض نقض شود، این مدل می تواند به برآورد اشتباه از احتمال تصادف منجر شود. در مقابل، ANN که نیازی به رابطه اساسی از پیش تعریف شده بین متغیرهای وابسته و مستقل ندارد، نشان داده است که ابزاری قدرتمند برای مقابله با مشکلات پیش‌بینی و طبقه‌بندی است. در مطالعه دیگر بررسی شده در این مقاله (Delen.2006) از یک سری شبکه‌های عصبی مصنوعی برای مدل سازی روابط بالقوه غیرخطی بین سطح شدت آسیب و عوامل مرتبط با تصادف استفاده شده است. آنالیز حساسیت بر روی مدل‌های شبکه عصبی آموزش دیده انجام شده است تا اهمیت عوامل مرتبط با تصادفات و موثر در میزان شدت صدمات مختلف مشخص شود. در این فرآیند، مشکل پیش بینی پنج کلاسه به یک دسته از مدل های پیش بینی دوتایی تقسیم شد تا اطلاعات مفیدی را برای شناسایی روابط "واقعی" علت و معلولی بین عوامل مرتبط با تصادف و سطوح مختلف شدت آسیب بدست آورد. نتایج تأیید شده توسط مطالعات قبلی، نشان دهنده تغییر اهمیت عوامل تصادف با تغییر میزان شدت آسیب است. در تحقیقی مشابه دیگر (Abdel Ati.2004) در مورد استفاده از دو الگوی مشهور ANN با عنوان شبکه‌های عصبی پرسپترون چند لایه و تئوری منطق فازی برای آنالیز شدت آسیب راننده بررسی های لازم صورت گرفته است. هدف از این مطالعه بررسی میزان زنده ماندن و مزایای احتمالی استفاده از ANN برای پیش‌بینی شدت آسیب راننده در صورت تصادف بود. عملکرد ANN با یک مدل پروبیت کالیبره شده مرتب شده مقایسه شد. نتایج مدل‌سازی نشان داد که دقت طبقه بندی تست برای پرسپترون چند لایه برابر ٪٧٣/٥ ، برای تئوری منطق فازی ٪٧٠/٦ و برای مدل پروبیت کالیبره شده ٪٦١/٧ می باشد. این نتیجه توانایی پیش‌بینی دقیق‌تر شدت صدمات را برای ANN (بخصوص پرسپترون چند لایه) در مقایسه با سایر روش‌های سنتی نشان می‌دهد. بنابراین، در این مطالعه از یک مدل رگرسیون دوجمله‌ای منفی و یک مدل ANN برای تجزیه و تحلیل داده‌های تصادف سالهای ١٣٩٥-١٣٩٢ در بزرگراه مشهد-نیشابور- سبزوار واقع در استان خراسان رضوی استفاده گردید. با مقایسه عملکرد پیش بینی بین مدل رگرسیون دوجمله‌ای منفی و مدل ANN، این مطالعه نشان می دهد که ANN یک روش جایگزین پایدار برای تجزیه و تحلیل تکرار تصادف جاده ای است.

٢- مروری بر تکنیک‌های مورد استفاده در تجزیه و تحلیل داده‌ها

در این مطالعه از سه روش مختلف برای تجزیه و تحلیل اطلاعات استفاده شده است:

آنالیز خوشه ای با استفاده از الگوریتم "k-means" برای تصادفات گروهی با خصوصیات یکسان و تکنیک های ANN و MVA پس از جمع آوری داده‌ها در قالب دو مدل (مدل ١ MVA و مدل٢ ANN). در ادامه اصول اساسی که دو مدل را توصیف می کند توضیح داده خواهد شد.

٢-١- تکنیک k-means

هدف از تجزیه و تحلیل خوشه‌ای شامل شناسایی یک پارتیش U مشخص در تعداد k گروه ($2 \le k \le n$) از مجموعه U است که توسط n داده تشکیل شده است. فرضیه ای که این روش مبتنی بر آن است به این صورت می‌باشد: داده‌های فضای X که متعلق به یک گروه هستند، با یک وابستگی ریاضی مشخص می‌شوند و این وابستگی بیشتر از داده‌های گروه‌های مختلف است. هر داده موجود در نمونه را می‌توان به عنوان نقطه‌ای توسط m مشخصه طرح‌ریزی کرد و هر مشخصه یک ویژگی از همین داده را تشکیل می‌دهد. یکی

از مقیاس‌های ساده‌تر وابستگی با فاصله اندازه گیری شده بین دو نقطه نشان داده شده است که به فضای داده تعلق دارد. اندازه گیری مناسب برای فاصله تعریف و مشخص می شود و این مورد بین هر واحد مشاهده و کلیه واحدها به عنوان یک کل اندازه گیری می‌شود. البته فاصله بین نقاط متعلق به یک گروه مشابه از فاصله بین نقاط موجود در گروه‌های مختلف کوچکتر است. مجموعه n داده به تعداد k گروه تقسیم می‌شود $X = \{x_1, x_2, x_3, \dots x_n\}$ و هر داده x_i بوسیله m مشخصه تعریف می‌شود ($x_i = \{x_1, x_2, x_3, \dots x_{im}\}$). به همین دلیل x_i (که x_i نشان دهنده تصادف i است) می‌تواند توسط یک نقطه در فضای Rm نمایش داده شود. این روش مبتنی بر استفاده از یک تابع هدف J است که تمایل به ایجاد گروه‌های "کروی" برای تقریب‌های پی در پی دارد. تابع هدف از دو نتیجه به طور همزمان پیروی می‌کند: اولاً فاصله اقلیدسی را بین نقاط هر گروه و مرکز گروه یکسان به حداقل می‌رساند (که به طور کلی با هیچ یک از نقاط مجموعه مطابقت ندارد)، دوم اینکه فاصله اقلیدسی را بین مراکز همه گروه‌ها به حداکثر می‌رساند. U نشان دهنده پارتیشن عمومی است و *U بهینه ای است که به فضای Mc مربوط به پارتیشن احتمالی X تعلق دارد. مقدار $J = J(U)$ ، فرض شده توسط تابع هدف برای هر پارتیشن U یک اندازه نسبی از میزان نزدیک بودن به حد مطلوب تشکیل می‌دهد. تابع هدف برای این است که حداقل جمع مربع مسافت‌های اقلیدسی بین همه نقاط و مرکز هر گروه اندازه‌گیری شود. پیدا کردن پارتیشن U دشوار است زیرا کاردینال بودن قسمت Mc از پارتیشن‌های احتمالی X به سرعت به بی‌نهایت گرایش پیدا می‌کند. جستجوی بهینه سازی در مشکلات در ابعاد مهم، بدون محاسبات دشوار امکان پذیر نیست، بنابراین مشکل با استفاده از یک الگوریتم بهینه سازی تکراری حل می‌شود. با فرض نخستین تلاش با پارتیشن U(r=0)، تعداد k گروه و یک مقدار تحمل تکرار ε (دقت مورد نیاز برای حل کردن) موقعیت مرکز گروه را می‌توان تعیین کرد. دوباره نسبت هر نقطه به گروه‌های مختلف محاسبه می‌شود و محاسبه جدیدی برای ماتریس U(r=1) بدست می‌آید. سپس دو تعیین کننده متوالی ماتریس U مقایسه و روند تکرار می‌شود تا اختلاف بین پارتیشن ها که در دو چرخه متوالی به دست آمده از حد از پیش تعریف شده فراتر رود. این تکنیک فرض می‌کند که از قبل تعداد خوشه‌ها معلوم شده است (Dell'Acqua.2013)، اما به عنوان تعداد بهینه خوشه‌هایی که برای طبقه بندی نهایی با آنها شناخته می‌شوند، به علت قابل توجه بودن عدم وجود اطلاعات اولیه در مورد ساختار خوشه‌هایی که در واحدهای مشاهده قرار می‌گیرند مشخص نیست و لذا به طور تصادفی ادامه می‌یابد. نویسنده بخش‌های مختلف بانک اطلاعاتی را فرض کرده و سپس مقداری برای شاخص S انتخاب کرده است که به عنوان بهترین شاخص گروه‌بندی تعریف شده است (Cokorilo.2014).

۲-۲- رویکرد و ساختار مدل ANN

ANN از ساختار و اصول عملکردی مغز انسان الهام گرفته شده است و از سلولهای عصبی مصنوعی بهم پیوسته ساخته شده است که برخی از خصوصیات نورونهای بیولوژیکی را تقلید می‌کنند. عملکرد یک نورون بیولوژیکی اضافه کردن ورودی آن و تولید یک خروجی است. این خروجی تنها در صورتی که سیگنال منتقل شده زیاد باشد (یعنی بزرگ‌تر از یک مقدار از پیش تعیین شده باشد) از طریق اتصالات سینوپتیک به سلولهای عصبی بعدی منتقل می‌شود در غیر این صورت سیگنال به نورون بعدی منتقل نمی‌شود. در شبکه ای که ایجاد می‌شود، یک نورون مصنوعی در واقع سامانه‌ای با تعداد زیادی ورودی و تنها یک خروجی می باشد و هر نورون دارای دو حالت آموزش و عملکرد می‌باشد. در حالت آموزش نورون یاد می‌گیرد که در مقابل الگوهای ورودی خاص برانگیخته شود و در حالت عملکرد وقتی یک الگوی ورودی شناسایی شده وارد شود، خروجی متناظر با آن ارائه می‌شود. بنابراین در شبکه، یک نورون با استفاده از رابطه۱، جمع وزنی را محاسبه می‌کند. با توجه به ورودی x_i و وزن w_i و مقایسه آن با مقدار آستانه: اگر جمع بیشتر از مقدار آستانه باشد، نورون روشن می شود و سیگنال منتقل می‌شود. در غیر این صورت نورون روشن نمی‌شود و جریان متوقف می‌شود.

$$I = \sum_{i=1}^{n} w_i . x_i \qquad (۱)$$

که در آن:

I: جمع وزنی (بدون بعد)

w_i : وزن (بدون بعد)

x_i: ورودی (بدون بعد)

پس به هر x وزنی به نام w نسبت داده می‌شود. کار شبکه‌ی عصبی این است که این وزن ها و بایاس را مطابق رابطه۲ طوری پیدا کرده و تغییر بدهد که خروجی حاصله از شبکه ی عصبی نسبت به مقدار واقعی کمترین خطا را داشته باشد (بهینه کردن وزن ورودی ها و بایاس).

$$Network = b + \sum_{i=1}^{n} w_i \,.\, x_i \qquad (۲)$$

در این تکنیک بخشی از داده ها مثلا به میزان ۷۰ درصد در اختیار شبکه برای یادگیری و بخشی (مثلاً ۱۵ درصد) در اختیار اعتبارسنجی قرار می‌گیرد تا شبکه‌ی عصبی با توجه به آموزش‌هایی که می‌بیند خودش را اعتبارسنجی کند که آموزش چقدر به داده‌های واقعی نزدیک است و اگر آموزش ایرادی دارد اصلاح کند و در نهایت بخشی (مثلاً ۱۵٪ داده‌ها) برای انجام فرآیند تست تخصیص داده می‌شود تا مشخص شود نتیجه حاصل از شبکه‌ی عصبی چقدر به نتیجه واقعی شبیه است (Zilioniene.2014).

۲-۳- رویکرد و ساختار مدل MVA

اگر در تحقیقی هدف بر این باشد که تأثیر یک یا چند متغیر مستقل بر روی بیش از یک متغیر وابسته مقایسه شود از تحلیل چند متغیره استفاده می‌شود (مقایسه دو گروه تحت تأثیر چند متغیر). این تحلیل برای پاسخ دادن به دو سؤال مناسب است:آیا تغییر در متغیرهای مستقل اثرات مشخصی روی متغیرهای وابسته می‌گذارد؟ رابطه بین متغیرهای مستقل و وابسته چیست؟ این مدل مجموعه‌ای از مدل‌های آماری است که به بررسی میانگین در گروه‌ها و توابع وابسته به آنها (مثل واریانس در یک گروه یا بین چند گروه) می‌پردازد. در این روش واریانس بدست آمده از یک متغیر تصادفی به اجزاء کوچکتری که منابع واریانس هستند تقسیم می‌شوند. در این روش، ANOVA[۳] آزمون آماری را فراهم می‌کند که برابری میانگین‌های گروه های متفاوت را می‌آزماید.

۳- جمع‌آوری داده‌ها در محیط GIS

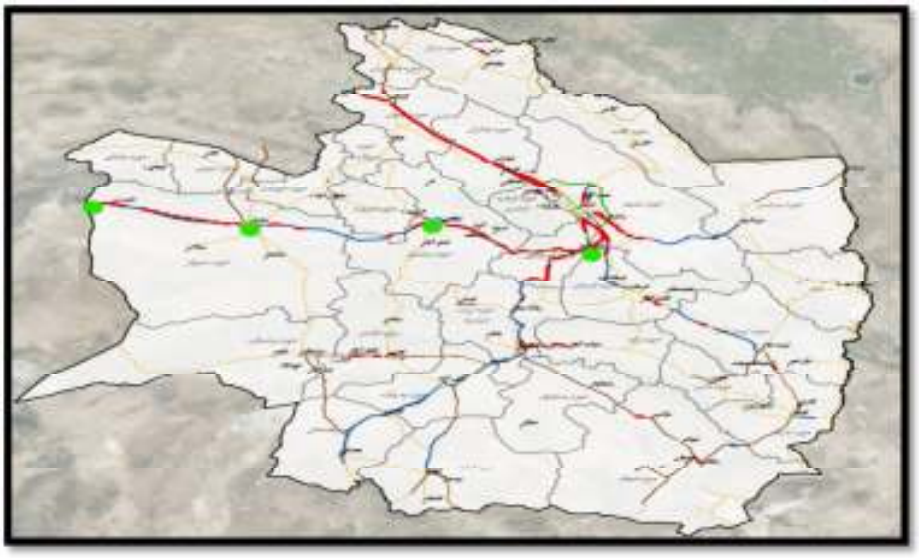
بخش‌های مورد بررسی متعلق به کریدور غربی-شرقی محور باغچه-نیشابور-سبزوار در استان خراسان رضوی است (شکل۱). مسیر تحلیل شده در ۳ قطعه در جدول ۱ نشان داده شده است. برای هر بخش، داده‌های مربوط به تعداد قوس افقی، تعداد تقاطع، نوع کاربری، متوسط سرعت، متوسط ترافیک روزانه در سال و تعداد فقره تصادفات جرحی و فوتی در دوره بین سالهای ۱۳۹۲ تا ۱۳۹۵ جمع آوری شد. سپس داده‌ها در سیستم اطلاعات جغرافیایی (GIS) به همراه مختصات جغرافیایی تصادفات ثبت و ذخیره شدند. راه مورد بررسی از باغچه تا انتهای محور بصورت لایه جی آی اسی در نرم افزار ARCGIS در نظر گرفته شد. لایه مذکور به صورت مقاطع سه کیلومتری قطعه‌بندی گردید. فایلی با فرمت dxf، حاوی محور راه و قطعات به ARCGIS وارد شد ، براساس مختصات UTM به مرجع جغرافیایی اختصاص داده شد و به فرمت مورد استفاده در ARCGIS تبدیل شد (فرمت shp-file). درنهایت، مشخصات هر المنت در محیط GIS بارگذاری شد و کلیه اطلاعات مربوط به ویژگی‌های پارامترهای معرفی شده در لایه‌های سطح مقطع قرار گرفت.

جدول۱. مشخصات قطعات آنالیز شده

شماره قطعه	ابتدای قطعه (km)	انتهای قطعه (km)	نام قطعه
۱	۰	۷٦	باغچه-نیشابور
۲	۷٦	۱۷۸	نیشابور- سبزوار
۳	۱۷۸	۲٦۹	سبزوار-کاهک

٤- برنامه تجزیه و تحلیل خوشه ای

تجزیه و تحلیل خوشه ای برای ماتریس داده‌های بارگذاری شده در نرم افزار GIS و مطابق متغیرهای معرفی شده در جدول۲، اعمال می‌شود. این تکنیک اجازه می دهد تا تصادفات در گروه‌هایی بصورت خوشه ای با یک سطح بالایی از وابستگی تجمیع شوند (به عنوان مثال تصادفات از نظر تعداد قوس افقی و تعداد تقاطع در وضعیت مشابه قرار می‌گیرند). بهترین تجمیع در ۸ کلاستر یا گروه بدست آمد. هر یک از ۸ گروه نشان داده شده در جدول ۳ که با محاسبه میانگین مقدار گروه بدست آمده است، می‌تواند به عنوان یک مقطع پرتصادف در نظر گرفته شود. به طور خاص، طول مقطع پرتصادف (مشخص شده با نام اختصاری L_i) به عنوان یک محدوده‌ای از تاثیر (Dell'Acqua.2013) هر تصادف که در فاصله مشابه رخ داده است محاسبه شده است. مجموع محدوده های تاثیر به عنوان طول (L_i) مقاطع پرتصادف فرض می‌شود.



شکل ۱. سگمنت‌های آنالیز شده

۵- مدل ۱ (MVA)

مدل ۱ که با استفاده از مبانی MVA (Highway Safety Manual.2009) و (Mario De Luca.2015) بدست آمده است و ساختار و متغیرهای مورد استفاده آن به صورت رابطه۳ می باشد در این مقاله بکار گرفته شد:

$$N_i.Sev = (365.AADT.L).b1 + e^{(N.Curve.b2+N.Int.b3+Land.b4+A.Sp.b5)} \quad (۳)$$

که در آن:

N.Curve، N.Int، Land، A.Sp و L_i : متغیرهای مستقل پیش بینی کننده

N_i.Sev : متغیر وابسته

نتایج حاصل از رگرسیون چندگانه غیر خطی در جداول ٤ و ۵ نشان داده شده است.

جدول ۲. متغیرهای مورداستفاده در آنالیز خوشه‌ای

متغیرها	علامت اختصاری	نوع متغیر	واحد	کد متغییر
متوسط تعداد قوس افقی	N.Curve	عددی	تعداد	–
متوسط تعداد تقاطع	N.Int	عددی	تعداد	–
نوع کاربری	Land	غیرعددی	–	۱: مسکونی، ۲: صنعتی، ۳: فاقد کاربری
متوسط سرعت	A.Sp	عددی	Km/hr	–

٦-مدل۲ (ANN)

مدل ۲ با استفاده از تکنیک ANN نشان داده شده در بخش ۱ با بکارگیری همان متغیرهای مورد استفاده در بخش ٤ بدست آمد. این مدل با استفاده از ۷۰ درصد از مجموعه داده‌ها برای آموزش ، ۱۵درصد برای تست و ۱۵ درصد برای اعتبارسنجی بدست آمده است. شکل ۲ ساختار مدل ANN و جدول ٦ پارامترهای برآورد ANN را نشان می‌دهد.

جدول٣. نتایج بدست آمده از آنالیز خوشه‌بندی با تکنیک "k-mean"

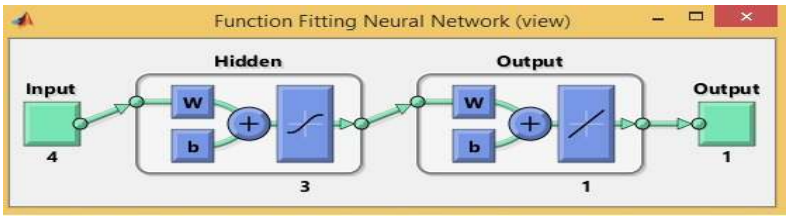
کلاستر	متوسط فاصله	متوسط تعداد قوس افقی	متوسط تعداد تقاطع	متوسط نوع کاربری	متوسط سرعت	AADT	متوسط شدت (Sev)	تعداد تصادف (L_i)
١	١/٤٣	٢/٣٧	٤/١٢	١/٨٧	١/٦٧	٢/٥	٣	٣٧٧
٢	١/٤٧	٣	٧	٢	٢/٥	١/١٥	٣	٢٢٤
٣	١/٦٨	٠/٠٨	١/٣٣	٢/٥١	١/٤١	٣/٥٢	٢/٥٤	٨٦٨
٤	١/٥٥	١/١٧	١	٢/٣٧	١/٢٣	٢/٧	٢/٦٨	٥٧٢
٥	١/٨٧	٠/٣٦	٥/١٨	٢/٣٦	١/٢٨	٣/٢٤	٣/٨٨	٤٢٧
٦	٢/٠٨	٣	٨	١	٢/٥	١	٣	١٠٥
٧	٢/٠٣	١/٥	١٠/٧٥	١/٢٥	١	١/٥	٣	٢٢١
٨	١/٨١	١/٨٧	٢/٣٧	٢/٨١	١/٦٧	٣	٢/٨٦	٣٣٧

جدول٤. پارامترهای تخمین زده شده با استفاده از مدل غیرخطی

پارامتر	تخمین	خطای استاندارد	فاصله اطمینان ٩٥ درصد	
			حد پایین	حد بالا
$b_1(L_i$ *AADT)	١/٥٢٦	٠/٧٨١	−١/٦٣٣	٥/٠٨٥
b_2(N.Curve)	−٠/٣٣٢	٠/١٤٦	−٠/٩٦١	٠/٢٩٧
b_3(N.Int)	−٠/٤٢٩	٠/٢٦٦	−١/٥٧٣	٠/٧١٥
b_4(Land)	−٠/٥٩٧	٠/٧٤٨	−٣/٨١٦	٢/٦٢٢
b_5(A.Sp)	٠/٧٩٤	١/٠٦٠	−٣/٧٦٥	٥/٣٥٣

جدول٥. تست ANOVA برای مدل غیرخطی

نتایج تست	میانگین مربعات	df	جمع مربعات
رگرسیون	٢/٠١٧	٥	١٠/٠٨٦
باقیمانده	٠/١٤٣	٢	٠/٢٨٧
جمع کل	_	٧	١٠/٣٧٣
متغیر وابسته:	(Ni•Sev); R^2=0.972		



شکل ٢. ساختار مدل ANN

جدول٦. پارامترهای ورودی و خروجی (تخمین زده شده) در مدل ANN

پیش‌بینی کننده		کلاستر١	کلاستر٢	کلاستر٣	کلاستر٤	کلاستر٥	کلاستر٦	کلاستر٧	کلاستر٨
لایه ورودی	N.Curve	٢/٣٧	٣	٠/٠٨	١/١٧	٠/٣٦	٣	١/٥	١/٨٧
	N.Int	٤/١٢	٧	١/٣٣	١	٥/١٨	٨	١٠/٧٥	٢/٣٧
	Land	١/٨٧	٢	٢/٥١	٢/٣٧	٢/٣٦	١	١/٢٥	٢/٨١
	A.Sp	١/٦٧	٢/٥	١/٤١	١/٢٣	١/٢٨	٢/٥	١	١/٦٧
	AADT	٢/٥	١/١٥	٣/٥٢	٢/٧	٣/٢٤	١	١/٥	٣
	N_i.Sev	١١٣١	٦٧٢	٢٢٠٤/٧٢	١٥٣٢/٩٦	١٩٢٨/٣٦	٣٦٧/٥	٦٦٣	٩٦٣/٨٢
لایه خروجی	N_i.Sev	١٠٩٢/٥	٦٦٤/٤٧	٢١٦٤/٦	١٥١٠/٣	١٨٨١/٧	٣٦٢/٥	٦٤٥/٣٤	٩٥٤/١٩

۷-مقایسه نتایج مدل ANN در مقابل MVA

جدول ۷ مقایسه نتایج بین دو مدل مورد بررسی را نشان می‌دهد و حاکی از این است که مدل ۲ از مدل ۱ دارای نتایج بهتری است به این دلیل که جمع کل مقدار باقیمانده که تفاوت بین میزان واقعی مشاهده شده و میزان پیش بینی شده توسط مدل است در مدل ۲ کمتر است (دو ستون آخر جدول ۷). از طرفی مشاهدات نشان می‌دهد که مدل ۱ تخمین بهتری را برای مقاطع دارای تعداد تصادف بیشتر نشان می‌دهد (جدول ۷، ردیف‌های ۱ ، ۳، ٤ و ۵). به عبارتی دیگر در این جدول می‌توان نتیجه‌گیری کرد که تکنیک ANN در خوشه‌های ۱، ۳، ٤ و ۵ مقدار باقیمانده بالاتری نسبت به روش MVA دارد. این یک محدودیت برای تکنیک ANN است که برای آنالیز شبکه اعتبار دارد، اما برای خوشه‌های دارای بیشترین خطر یا بیشترین تصادف اطمینان حاصل نمی‌کند، لذا به نظر می‌رسد که مدل ۱ در تخمین مقاطع پرتصادف خطرناک‌تر (دارای تعداد زیاد فقره تصادف) کارایی و عملکرد بهتری دارد.

۸- نتیجه‌گیری

این مطالعه نتایج حاصل از پیش بینی تصادفات جاده‌ای را با مقایسه تکنیک شبکه عصبی مصنوعی (ANN) و تکنیک آنالیز چند متغیره (MVA) نشان می‌دهد. این مطالعه حوادث رخ داده در یک بزرگراه دو خطه واقع در استان خراسان رضوی شامل تصادفات جرحی و فوتی در فاصله زمانی بین سال‌های ۱۳۹۲ و ۱۳۹۵ را در نظر می‌گیرد. داده‌های جمع‌آوری شده با تجزیه و تحلیل خوشه‌ای (با استفاده از پارتیشن دوتایی الگوریتم "k-means") در گروه‌هایی به صورت خوشه‌ای با یک سطح بالایی از وابستگی تجمیع شدند و بهترین تجمیع در ۸ کلاستر یا گروه بدست آمد. سپس با تکنیک‌های MVA و ANN دو مدل به دست آمده است: مدل ۱ (مدل MVA) و مدل ۲ (مدل ANN). با مقایسه نتایج دو مدل، مشخص شد که مدل ۲ از مدل ۱ دارای نتایج نزدیک‌تر به واقعیت و با دقت بهتری است. زیرا از کل میزان باقیمانده پایین‌تری نیز برخوردار است، اگرچه به نظر می‌رسد که مدل ۱ برای توصیف مقاطع دارای تعداد تصادف بیشتر، عملکرد بهتری را دارا می‌باشد.

جدول۷. مقایسه نتایج مدل‌ها و میزان واقعی مشاهده شده به تفکیک هر کلاستر

کلاستر	Ni·Sev مشاهده شده	Ni·Sev تخمین از مدل۱	Ni·Sev تخمین از مدل۲	باقیمانده ۱ (مدل۱- مشاهده)	باقیمانده ۲ (مدل۲- مشاهده)
۱	۱۱۳۱	۱۱۱٤/۰۸	۱۰۹۲/۵	۱٦/۹۲	۳۸/۵
۲	٦۷۲	٦٤٦/٤٦	٦٦٤/٤۷	۲۵/۵٤	۷/۵۳
۳	۲۲۰٤/۷۲	۲۱٦۹/۹۱	۲۱٦٤/٦	۳٤/۸۱	٤۰/۱۲
٤	۱۵۳۲/۹٦	۱۵۱۷/۹٦	۱۵۱۰/۳	۱۵	۲۲/٦٦
۵	۱۹۲۸/۳٦	۱۸۹۲/۲۲	۱۸۸۱/۷	۳٦/۱٤	٤٦/٦٦
٦	۳٦۷/۵	۳۵۰/٤۵	۳٦۲/۵	۱۷/۰۵	۵
۷	٦٦۳	٦۳۵/۱۱	٦٤۵/۳٤	۲۷/۸۹	۱۷/٦٦
۸	۹٦۳/۸۲	۹۳۹/٦۵	۹۵٤/۱۹	۲٤/۱۷	۹/٦۳
جمع باقیمانده ها				۱۹۷/۵۳	۱۸۷/۷٦

- Delen, D., Sharda, R., Bessonov, M., (2006), "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks", Accident Analysis & Prevention 38(3), pp. 434–444.

-Mario De Luca, (2015), "A Comparison Between Prediction Power of Artificial Neural Networks and Multivariate Analysis in Road Safety Management", Jurnal of Transport, Vol. 32,(4), pp. 379–385.

-Chang, L.-Y., (2005), "Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network", Safety Science 43(8), pp. 541–557.

۹- پی‌نوشت‌ها

1. Multivariate Analysis
2. Artificial Neural Network
3. Analysis of Variance

۱۰- مراجع

-Abdel-Aty, M., Abdelwahab, H., (2004), "Predicting injury severity levels in traffic crashes: a modeling comparison", Journal of Transportation Engineering 130(2), pp. 204–210.

-Chiou, Y. C., (2006), "An artificial neural network-based expert system for the appraisal of two-car crash accidents", Accident Analysis & Prevention 38(4), pp. 777–785.

-Čokorilo, O., De Luca, M., Dell'Acqua, G., (2014), "Aircraft safety analysis using clustering algorithms", Journal of Risk Research 17(10), pp. 1325–1340.

-De Luca, M., Dell'Acqua, G., (2012), "Freeway safety management: case studies in Italy", Transport 27(3), pp. 320–326.

-Dell'Acqua, G., Russo, F., Biancardo, S. A., (2013), "Risk-type density diagrams by crash type on two-lane rural roads", Journal of Risk Research 16(10), pp.1297–1314.

-Highway Safety Manual, (2009), 1st edition. 1057 p. Available from Internet: http://www.este.civ.uth.gr/apodeltiosi/HSM.pdf.

-Vujanić, M., Lipovac, K., Jovanović, D., Pešić, D., Antić, B., (2013), "Bottom-up and top-down approach for defining road safety strategy–case study:city of Belgrade", International Journal for Traffic and Transport Engineering 3(2), pp.185–203.

-Žilionienė, D., De Luca, M.; Dell'Acqua, G., Lamberti, R., Biancardo, S. A., Russo, F., (2014), "Evaluating freeway traffic noise using artificial neural network", in 9th International Conference on Environmental Engineering, pp. 22–24.

Prediction of the Road Accidents Severity Using Artificial Neural Network and Comparing with Multivariate Analysis Method

Mohammadreza keymanesh, Assistant Professor, Civil Engineering Faculty, Payame Noor University, Tehran, Iran.
Nassir Baradaran Rahmanian, Ph.D. Student, Civil Engineering Faculty, Payame Noor University, Tehran, Iran.

E-mail: mrkeymanesh@pnu.ac.ir

Received: September 2020-Accepted: January 2021

ABSTRACT
Prediction of the road accidents severity based on scientific methods can be considered as an important tool in road safety management and transportation engineering. This paper presents a methodology to identify and predict the severity of accidents on highways. The methodology is based on an experimental investigation carried out on an important corridor and arterial way located in khorasan-Razavi province. The study analyses accidents occurring between 1392 and 1395 and the data concerning the accidents that were acquired from police reports. The geometric data were acquired from field study and the related office. The data, organized and stored in a specific designed Geographic Information System (GIS), were processed using a series of statistical procedures; in particular, the results took out the following two models: Model 1 was produced by Multivariate Analysis (MVA) and the Model 2 was obtained using the Artificial Neural Network (ANN) technique. Comparing the two models, it emerged that Model 2 is better than Model 1 because the total sum of the residual is lower. However, Model 1 is more efficient in estimating the black spots with a large number of accidents.

Keywords: Prediction, Accidents Severity, Artificial Neural Network, Multivariate Analysis