

مدل سازی نقش عامل انسانی در تصادفات با استفاده از شبکه عصبی مصنوعی

شهریار افندی زاده^۱، مهدی نبی زاده^۲، نازنین جوادی درخانه^۳
۱- دانشیار، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران
۲- کارشناس ارشد مهندسی عمران گرایش برنامه ریزی حمل و نقل، مدرس دانشگاه

چکیده

هدف این مقاله ارائه مدلی مبتنی بر شبکه های عصبی با دقت بالا برای ساخت مدل پیش بینی تصادفات براساس نقش عامل انسانی می باشد. با عنایت به این موضوع که در اکثر مطالعات به نقش جاده و وسیله نقلیه بیش از موارد دیگر نگریسته شده است، در این تحقیق با استفاده از تخلفات ترافیکی رانندگان که از مهمترین نشانه های رفتار عامل انسانی می باشد، اقدام به طراحی مدل گردیده است.

جهت انتخاب متغیرهای موثر در پروسه طراحی مدل، از بین تخلفات رانندگی ثبت شده در بزرگراه های برون شهری تحقیق، تخلفات حادثه ساز انتخاب شده اند، از طرف دیگر با بررسی تصادفات در فاصله زمانی مورد نظر، تصادفات براساس شدت (خسارتی، جرحی و فوتی) استخراج و با استفاده از ضرایب معادل سازی مناسب، تبدیل به شدت تصادف گردیدند، سپس با استفاده از مدل سازی به روش شبکه عصبی، شبکه هایی با ساختار متفاوت ساخته شده و در نهایت مدل بهینه با ارزیابی پارامترهای شبکه انتخاب و بوسیله داده های جدید که در آموزش مدل وارد نگردیده، مورد اعتبار سنجی قرار گرفت و نتایج پارامترهای شبکه بهینه نشان از دقت بالای شبکه عصبی در ساخت مدل فوق دارد.

کلید واژه: تصادفات، نقش عامل انسانی، تخلفات ترافیکی رانندگان، مدل شبکه عصبی مصنوعی

۱- مقدمه

کارشناسان در مبحث مدیریت ترافیک باید از عملکرد ایمن معابر مطلع باشند. بدین منظور روش قابل اطمینانی برای پیش بینی میزان تصادفات مورد نیاز است. در همین راستا محققین مدل هایی ارائه کردند که با کمک داده های تصادفات، تجزیه و تحلیل آن و قضاوت کارشناسی می توان تعداد تصادفات را پیش بینی نمود. برای به دست آوردن نتایج واقعی و قابل اطمینان از مدل، به اطلاعات مناسب تصادفات، ترافیک و طرح هندسی نیاز است. توسعه اینگونه مدل ها برای بررسی تأثیر عوامل موردنظر بر روی تصادفات، مهندسين را قادر می سازد تا اقدامات لازم را در جهت شناسایی و کاهش تعداد و شدت تصادفات و یا کاهش احتمال وقوع آن ها اعمال نمایند. انسان به عنوان استفاده کننده از سیستم حمل و نقل، نقش اصلی را در ایمنی و تصادفات ایفا می کند و در رخداد هر تصادفی می توان تأثیر عامل انسانی به صورت های مختلف را مشاهده نمود، در بیش از ۹۵ درصد از تصادفات عامل انسانی به صورت عامل مستقیم و یا همراه با عوامل دیگر نقش دارد. استفاده از مدل های تخمین در برآورد تصادفات در

تصادفات و تلفات انسانی از جمله عوارض ناخوشایند توسعه ناموزون تکنولوژی حمل و نقل می باشند. لازمه موفقیت سیاست های ایمنی و تضمین کننده کارایی اقدامات پیشگیرانه، بکارگیری نتایج مطالعات و تحقیقات علمی در خصوص ایمنی راه می باشد، از جمله ابزارهای مؤثر و مفید در این زمینه مدل ها می باشند چرا که مدل از یک سو ابزاری است جهت توصیف، توضیح و یا پیش بینی یک پدیده معین و از سوی دیگر می تواند جهت ترکیب و یکپارچه سازی اثرات متغیرهای تصمیم گیری بکار گرفته شود، آشکار است که چنین شناختی، منتهی به امکان پذیر شدن تدوین برنامه های ایمن سازی تردد در جهت کاهش تبعات ناشی از تصادفات می شود. هزینه های ناشی از تصادفات، مسئولان و مدیران شهری را به فکر ارائه راهکارهای حل این معضل پرداخت.

m_a_dehshiri@yahoo.com
shokrgozar_mahdi90@yahoo.com
moghadas@aut.ac.ir

حوزه حمل و نقل درون شهری و برون شهری، به دلیل افزایش سالانه مرگ و میر ناشی از حوادث ترافیکی در کشورهای در حال توسعه در سال های اخیر مرسوم شده است. با توجه به متفاوت بودن شرایط مربوط به رانندگان، وسایل نقلیه، جاده و محیط در کشورهای مختلف لازم است که در کشور ما نیز برای انواع و اقسام راه های موجود در کشور مدل های پیش بینی تصادفات توسعه داده شود تا نقش عوامل مختلف در بروز تصادفات مشخص شده و راهکارهای مناسب برای کاهش این تصادفات ارائه شود.

۲- مروری بر مطالعات گذشته

تقریباً پس از دهه هفتاد میلادی، تحقیقات مربوط به تصادفات جاده ای از کارهای علمی و تحلیل های منسجم اثر گرفت. محققان به این واقعیت پی بردند که تصادفات نتیجه فقط یک علت نبوده بلکه پیامد زنجیره ای از دلایلند، این زنجیره شامل علل و عوامل جاده ای، انسانی، وسیله نقلیه و محیط اطراف است که هر یک از آن ها نیز مشتمل بر اجزای متعدد دیگری هستند. بنابراین یافتن علل واقعی تصادفات از میان سلسله عوامل جمع آوری شده، مثل یافتن حلقه های ضعیف این زنجیره خواهد بود و این اولین قدم در راه ایمن سازی حمل و نقل جاده ای مبتنی بر تحقیقات علمی است.

تحقیقاتی برای بیان ارتباط تصادفات جاده ای و تخلفات رانندگی صورت گرفته اند، برای نمونه مک دونالد در سال ۲۰۰۴ ثابت کرد که رانندگانی که تحت تأثیر موادمخدر هستند بیشترین تخلفات رانندگی را مرتکب میشوند که منجر به تصادف میگردد و افرادی که تحت درمان ترک اعتیاد هستند تخلفات کمتری انجام می دهند [۱].

در سال ۲۰۰۴ میلادی نیز کاستلا و پرز رابطه میان الگوهای رفتاری شخصیتی مختلف و ارتکاب به تخلف را دریافتند که به وسیله تست روان کاوی و تست عکس العمل رانندگان نسبت به مجازات اعمال شده بدست آمد، بدین صورت که رانندگانی که در هنگام ارتکاب به تخلف با مجازات و تنبیه کمتری مواجه می شوند تمایل بیشتری به تخلف از قوانین راهنمایی و رانندگی دارند و این مورد در خانم ها کمتر از آقایان مشهود می باشد [۲].

مطالعات در مورد آنالیز تصادفات بیشتر بر مدل های آماری مانند مدل های رگرسیون خطی، مدل رگرسیون لگاریتمی اعم از لگاریتم طبیعی و پواسون و مدل های دو جمله ای منفی تکیه داشته اند و در این مدل ها روابط بین تصادفات و ترافیک عبوری، مشخصات هندسی و فاکتور های محیطی بررسی شده است. در برخی از این تحقیقات از مدل های رگرسیون چند متغیره خطی در پیش بینی تصادفات استفاده گردیده که فرض شده است تعداد تصادفات و یا نرخ تصادفات و یا نرخ تصادفات یک تابع خطی از خصوصیات مختلف محیطی ترافیک مثل جریان

عبوری، آب و هوا، طرح هندسی، وضعیت رانندگی و ... می باشد. از جمله این تحقیقات می توان به مطالعات لاندی، ایوی اشاره کرد. ایوان و اوامر برای پیش بینی تعداد تصادفات ترافیکی از مدل رگرسیون پواسون استفاده کرده و به این نتیجه رسیده اند که مدل پواسون نسبت به مدل رگرسیون خطی ترجیح داده می شود و بکارگیری مدل پواسون در وضعیت های همگن و زمانی که داده های تصادف پراکندگی زیادی ندارند، درست می باشد [۳].

در یکی از تحقیقاتی که توسط جونز و جانسن و منرینگ که در منطقه سیاتل آمریکا جهت تجزیه و تحلیل تصادفات با استفاده از مدل پواسون انجام شده است از بانک اطلاعاتی تصادفات بعنوان مجموعه اطلاعاتی و تعیین متغیرها استفاده شده است و متغیرهایی بکار رفته که عوامل مزبور بیانگر تغییرات حجم ترافیک، الگوی سفر، تاثیرات متقابل اندازه گیری نشده و طرح هندسی و نظایر آن می باشد [۴].

میاو و لوم و همچنین جوشو و گاربر در تحقیق خود در زمینه مدل پواسون و کمترین مربعات، نامناسب بودن روش های کمترین مربعات را جهت مدلسازی تصادفات نشان دادند و بکارگیری توزیع پواسون را توصیه نمودند، اما توزیع پواسون یک محدودیت مهمی دارد، به عبارت دیگر در این توزیع میانگین و واریانس بایستی برابر باشند [۵].

عبدل اتی و همکاران از یک شبکه عصبی احتمالی برای پیش بینی تصادفات در کریدور بین شهری شماره ۱ در اورلاندو استفاده کرده اند. آنالیز نتایج نشان داده اند که حداقل ۷۰ درصد از تصادفات می تواند یا مدل شبکه عصبی احتمالی به طور صحیح پیش بینی شود [۶].

چانگ در مدل سازی تصادفات جاده ها در کشور تایوان از دو مدل شبکه عصبی مصنوعی و دو جمله ای منفی استفاده کرده است و با مقایسه کارایی دو مدل، به این نتیجه رسیده است که مدل شبکه های عصبی مصنوعی یک روش جایگزین محکم و قوی برای آنالیز تصادفات می باشد [۷].

۳- روش تحقیق

۴-۱- متغیر وابسته

در اکثر مدل های پیش بینی تصادفات از متغیرهایی نظیر تعداد تصادفات رخ داده و یا نرخ تصادف استفاده می شود و تمایزی در نوع تصادفات مشاهده نمی گردد، لذا در این تحقیق این نقص برطرف شده و معادل سازی انواع تصادفات انجام شده است.

شدت تصادف بعنوان متغیر وابسته در این تحقیق استفاده می شود، این متغیر بیانگر شاخصی برای تصادف های روی داده در محورهای مورد مطالعه می باشد، با توجه به اینکه تصادفات از نظر شدت به سه نوع خسارتی، جرحی و فوتی تقسیم می شوند، لذا شدت تصادف بوسیله رابطه (۱) معرفی و تعریف می شود.

$$AS = [DA + (a \times IA) + (b \times FA)] / L \quad (1)$$

AS شدت تصادف^۱، DA : تصادف خسارتی^۲، IA : تصادف جرحی^۳، FA : تصادف فوتی^۴
a و b : وزن نسبی هزینه تصادفات جرحی و فوتی نسبت به تصادف خسارتی
طول محور به کیلومتر :

در این تحقیق ضرایب معادل سازی متغیر وابسته براساس نسبت هزینه تصادفات فوتی، جرحی و خسارتی می باشد، که از پروژه برآورد هزینه های تصادفات (درون شهری و برون شهری) در راه های ایران، استفاده می شود [۹]. رابطه شدت تصادف براساس تحقیق، فوه، بصورت زیر در می آید.

$$AS = [DA + (6 \times IA) + (111 \times FA)] / L \quad (2)$$

مبنای زمانی رابطه (۲) بصورت ماهانه می باشد، یعنی شدت تصادف براساس تعداد تصادفات فوتی، جرحی، خسارتی ماهانه در محور مورد نظر محاسبه و تقسیم آن بر طول محور بدست می آید.

۴-۲- متغیرهای مستقل

با عنایت به اینکه موضوع تحقیق، طراحی مدل پیش بینی شدت تصادف براساس تخلفات ترافیکی رانندگان می باشد، از بین تخلفات شش نوع تخلف رانندگی را که براساس آمار و اطلاعات موجود علت اصلی وقوع اکثر تصادفات و یا باعث افزایش شدت تصادف بوده، شامل سرعت غیرمجاز، سبقت غیرمجاز، عدم رعایت فاصله طولی، انحراف به چپ، عدم رعایت حق تقدم، عدم استفاده

با توجه به اینکه اکثر بهینه سازی های مهندسی بر پایه روش های محاسبات خطی عددی می باشند که این مستلزم پیدا کردن جواب بهینه در بین یک سری جواب می باشد، استفاده از الگوریتم های فرا ابتکاری می تواند تا حدودی یک شبیه سازی از دنیای واقعی در حل این گونه مسائل باشد، در واقع روش های فرا ابتکاری قابلیت زیادی در حل مسائل کاربردی از نوع بهینه سازی ترکیبی دارند، در این تحقیق از بین مدل های فرا ابتکاری از روش شبکه عصبی مصنوعی برای ساخت مدل استفاده شده است. شبکه های عصبی توان بالقوه ای برای حل مسائلی دارد که شبیه سازی آن ها از طریق منطقی، تکنیک های تحلیلی سیستم های خبره و تکنولوژی های استاندارد نرم افزاری مشکل است. این تکنیک قادر است درمواقعی که داده ها در شرایط عدم اطمینان اعم از آنکه داده ها فازی باشند و یا به طور ناقص و توأم با نویز دریافت شده باشند جواب منطقی ارائه دهد. بنا بر نظر اندیشمندان در حل مسائلی که فاکتورهای انسانی دخیل هستند، ارائه متدولوژی های مبتنی برانسان (دراین جا سلول های عصبی انسان) احتمالاً منجر به راه حل های بسیار کارساز می شود [۸]. عوامل فوق همگی بیانگر این نکته هستند که شبکه های عصبی گزینه مناسبی برای حل مسائل حمل و نقل می باشند، درعمل نیز به کارگیری شبکه های عصبی در حل مسائل حمل و نقلی امروزه بسیار گسترده شده است. روش تحقیق به این صورت است که پس از بررسی تحقیقات گذشته در زمینه مدل ها و متغیرهای مؤثر بر تصادفات، منطقه مطالعاتی تعیین و اطلاعات مورد نیاز برای دوره مشخص جمع آوری شده است. سپس با ساخت مدل شبکه عصبی و توسعه آن، مدل شبکه عصبی بهینه انتخاب شده و صحت و اعتبار مدل با انجام اعتبارسنجی مورد ارزیابی قرار گرفت.

۴- متغیرهای مؤثر در مدل

با عنایت به هدف تحقیق که بررسی نقش عامل انسانی با استفاده از تخلفات ترافیکی رانندگان در تصادفات جاده ای می باشد، اقدام به ساخت مدل پیش بینی تصادف با استفاده از شبکه های عصبی مصنوعی می گردد. در طراحی این مدل نیازمند معرفی کامل متغیر وابسته و نیز متغیرهای مستقل بوده که در واقع از ارتباط بین متغیرهای مستقل و وابسته هدف مورد نظر که ساخت مدل می باشد، حاصل گردد.

- 1- Accident Severity
- 2- Property Damage Accident
- 3- Injury Accident
- 4- Fatal Accident

می باشند به جهت اینکه همه متغیرها بر یک مبنای واحد باشند، تعداد تخلف در هر کیلومتر محور مورد نظر از تقسیم تعداد تخلف ماهیانه محور مورد نظر بر طول محور بدست می آید. در جدول ۱ متغیرهای موثر در مدل سازی نشان داده شده اند.

از کمریند ایمنی انتخاب می شوند. متغیرهای مستقل در واقع تعداد ثبت شده هر کدام از تخلفات رانندگی ذکر شده در بالا در محورهای مورد مطالعه در هر کیلومتر در بازه زمانی یک ماهه می باشند، یعنی در واقع با توجه به اینکه طول محورها متفاوت

ردیف	نوع متغیر	نام متغیر (در هر ماه در هر کیلومتر محور مورد نظر)	نشانه متغیر
۱	متغیر وابسته	شدت تصادف	AS
۲	متغیرهای مستقل	تعداد تخلف سرعت غیر مجاز	T ₁
۳		تعداد تخلف انحراف به چپ	T ₂
۴		تعداد تخلف سبقت غیر مجاز	T ₃
۵		تعداد تخلف عدم رعایت حق تقدم	T ₄
۶		تعداد تخلف عدم رعایت فاصله طولی	T ₅
۷		تعداد تخلف عدم استفاده از کمریند ایمنی	T ₆

جدول ۱: متغیرهای موثر در مدل سازی

۵- اطلاعات مورد استفاده در تحقیق

در نهایت از مجموعه یاد شده یک بانک اطلاعاتی شامل اطلاعات تصادفات و همچنین اطلاعات تخلفات رانندگی مربوط به ۴ محور بزرگراه بین شهری در نرم افزار MATLAB تشکیل و برای مدل سازی مورد استفاده قرار می گیرد. این بانک اطلاعاتی شامل اطلاعات تصادفات و تخلفات رانندگی ۵۲ ماه استان گیلان در حد فاصل سال های ۱۳۸۸ الی ۱۳۹۳ بوده است که اطلاعات ۴۴ ماه طی سال های ۱۳۸۸ الی ۱۳۹۲ بعنوان داده های ساخت مدل و اطلاعات ۸ ماهه سال ۱۳۹۳ برای تست مدل مورد استفاده قرار گرفته است. تصویری از این بانک اطلاعاتی در جدول ۲ نشان داده شده است.

در این تحقیق نیاز به محورهایی با خصوصیات مورد نظر و برخوردار از بانک اطلاعات تصادفات می باشد. پس از انجام بررسی های لازم و امکان دست یابی به اطلاعات مشخص شد که تعدادی از محورهای استان گیلان نیازهای تحقیق را برآورده می کنند. از طرف دیگر با توجه به اینکه اطلاعات تخلفات ترافیکی ثبت شده در جاده ها در بانک اطلاعاتی تصادفات وجود ندارد، بنابراین نیازمند افزودن چنین اطلاعاتی به بانک اطلاعات تحقیق بود. در نتیجه اقدام به تهیه اطلاعات مربوط به تخلفات ترافیکی رانندگی ثبت شده در جاده های ناحیه مطالعاتی از طریق پلیس راه گردید.

ردیف	تعداد تخلفات ماهیانه در هر کیلومتر محور						تعداد تصادفات (ماهیانه)			شدت تصادف
	عدم استفاده از کمریند ایمنی	عدم رعایت فاصله طولی	عدم رعایت حق تقدم	سبقت غیر مجاز	انحراف به چپ	سرعت غیر مجاز	فوتی	جرحی	خسارتی	
۱	۳۷/۸۹	۳۷/۸۹	۰/۸۴	۷/۱۱	۶/۰۲	۳/۷۶	۷/۲	۲۴	۵۰	۰
۲	۱۸/۶۷	۰/۹۱	۰/۶	۲/۰۹	۱/۲۵	۳/۷۶	۱۲/۵۳	۱۰۷	۶	۲
۳	۵۷/۰۳	۰/۳۸	۰/۲۲	۴۶/۵۶	۴/۷۵	۱۰۰/۶۳	۲۲/۵۰	۶۶	۳۵	۴
۴	۵۷۶۸	۰/۲۱	۰/۲۶	۲۴/۷۹	۳/۲۹	۶۳/۱۵	۱۵/۴۱	۳۲	۴۵	۲
۵	۲/۶۴	۰/۵۱	۰/۳۳	۴/۸۹	۰/۸۰	۳۳/۹۱	۷/۲۹	۴۳	۲۹	۱
۶	۱۹/۹۵	۱/۱۱	۰/۳۶	۱/۸۲	۱/۴۷	۴۷/۳۸	۹/۸۴	۱۱۲	۵۳	۱
۷	۳۹/۶۰	۰/۷۱	۰/۶۷	۲/۳۵	۱/۷۶	۳۲/۵۱	۷/۳۵	۱۰۴	۵	۰

جدول ۲: تصویری از بانک اطلاعاتی مورد استفاده در طراحی مدل

۵-۱- نرمال سازی داده ها

در این تحقیق رابطه شماره ۳ برای تبدیل اعداد بین صفر تا یک مورد استفاده قرار گرفته است. شرط استفاده از این فرمول ها آن است که داده های جمع آوری شده در دامنه اعداد مثبت تعریف شده باشند. آمارهای مربوط به مقادیر واقعی و نرمال شده در جدول ۳ آورده شده است.

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \quad (3)$$

x_{new} : داده جدید x_{min} : کوچکترین مقدار داده
 x_{max} : بزرگترین مقدار داده x_{old} : داده قدیمی

از آن جا که تمامی پارامترها در مدل های تصمیم گیری از ابعاد یکسانی برخوردار نیستند، برای یکسان سازی آن ها از روش های نرمال سازی یا استاندارد سازی داده ها استفاده می شود. روش های متعددی جهت نرمال سازی داده های ورودی و خروجی وجود دارد که متداول ترین آن ها داده های جمع آوری شده را به اعداد بین ۰ تا ۱ و یا بین -۱ تا ۱ تبدیل می نمایند.

ردیف	عنوان (در هر کیلومتر)	مقادیر واقعی			مقادیر نرمال شده		
		حداقل	حداکثر	انحراف معیار	حداقل	حداکثر	انحراف معیار
۱	شاخص خطر راه	۲/۰۹	۳۰/۷۸	۵/۶۹	۰	۱	۰/۱۹۶
۲	تخلف سرعت غیرمجاز	۰/۸۴	۱۰۲/۲۵	۲۱/۶۹	۰	۱	۰/۲۱۴
۳	تخلف سبقت غیرمجاز	۰/۱۶	۴۶/۵۶	۹/۶۲	۰	۱	۰/۱۹۷
۴	تخلف انحراف به چپ	۰	۱۵	۲/۹۵	۰	۱	۰/۲۰۵
۵	تخلف عدم رعایت فاصله طولی	۰	۱/۹۴	۰/۳۶	۰	۱	۰/۲۲۳
۶	تخلف عدم رعایت حق تقدم	۰	۲/۲۵	۰/۴۵	۰	۱	۰/۱۸۰
۷	تخلف عدم استفاده کمربند ایمنی	۰	۱۰۲/۵۳	۲۰/۴۷	۰	۱	۰/۱۹۹

جدول ۳: آمارهای مربوط به مقادیر واقعی و نرمال شده پارامترهای بانک اطلاعاتی

۶- مدل شبکه عصبی مصنوعی

آخرین لایه شبکه اعصاب می باشد، این لایه به دلیل تولید خروجی، جزء مهمترین قسمت شبکه بشمار می رود، در واقع خروجی های شبکه در این لایه تولید می شوند و با خروجی های مطلوب مقایسه می شوند. در لایه خروجی نیز به تعداد متغیرهای خروجی نرون خروجی وجود دارد که همان عصب می باشد در این تحقیق خروجی دارای یک پارامتر و به تبع آن لایه خروجی شبکه نیز دارای یک عصب است. تنها مرحله ای که تعداد عصب آن از اهمیت خاصی برخوردار است و از ابتدا مشخص نیست، لایه پنهان است، لایه پنهان هسته اصلی محاسباتی را تشکیل می دهند و این لایه است که بوسیله نرون ها، ورودی ها را از لایه ورودی دریافت کرده و سپس مقدار خروجی را در لایه خروجی تولید می کنند.

برای مدلسازی از نرم افزار MATLAB به خاطر توانایی بالا در ساخت شبکه های عصبی و همچنین انعطاف پذیر بودن و راحتی کار استفاده شده است. رایج ترین نوع ساختار شبکه عصبی از سه لایه ورودی، پنهان و خروجی تشکیل شده است. لایه ورودی در شبکه عصبی لایه ای است که در آن داده ها جهت آموزش شبکه اعمال می شود. این لایه ها در واقع به عنوان دروازه های شبکه هستند. ساختار شبکه باید طوری باشد که در این لایه ورودی به اندازه تعداد متغیرها جهت حل مساله، عصب وجود داشته باشد. در لایه ورودی هر داده دارای ۶ پارامتر و در نتیجه لایه ورودی نیز دارای ۶ عصب می باشد. لایه خروجی

۶-۱- ساختار شبکه عصبی مصنوعی

لایه پنهان از جدول ۴ استفاده شده که ساختار شبکه های استفاده شده در این جدول آورده شده است. در ابتدا تعداد عصب ها در لایه پنهان ۵ در نظر گرفته شده و پارامترهای شبکه محاسبه شده و در مراحل بعد نیز با افزایش تعداد عصب در لایه پنهان پارامترهای مربوط به شبکه محاسبه شده تا شبکه عصبی بهینه مشخص گردد. در این تحقیق از ۲۰۸ داده جمع آوری شده، ۱۷۶ داده برای ساخت مدل شبکه عصبی و ۳۲ داده برای ارزیابی و اعتبارسنجی مدل مورد استفاده قرار می گیرد.

جهت رسیدن به شبکه ای با بهترین عملکرد، باید تعداد بهینه نرون ها در لایه پنهان را با تغییر تعداد نرون ها در این لایه، جستجو کرد، در واقع باید در ساخت مدل شبکه عصبی تعداد عصب لایه پنهان را با سعی و خطا تغییر داد و بر اساس تعداد عصب مختلف پارامترهای شبکه را مورد ارزیابی قرار داد تا زمانی که به نتیجه مطلوب که همان شبکه عصبی بهینه با میانگین مربعات خطای کم و ضریب رگرسیون بالا می باشد، حاصل شود. به منظور ارزیابی پارامترهای شبکه با تعداد عصب متفاوت در

کد شبکه	۱	۲	۳	۴	۵	۶
ساختار شبکه	۶-۵-۱-۱	۶-۱۰-۱-۱	۶-۱۶-۱-۱	۶-۲۰-۱-۱	۶-۲۵-۱-۱	۶-۳۰-۱-۱

جدول ۴: حالت های مختلف ساختار شبکه عصبی

۶-۲- ساخت شبکه های عصبی مختلف

داده ها که به هر یک از این سه دسته اختصاص داده می شود در پارامترهای شبکه تاثیر گذار می باشند. بدین منظور داده ها در چهار حالت با درصد های مختلف به شرح جدول ۵ در شبکه اعمال شده است که تفکیک داده ها در هر کدام از این حالت ها بصورت تصادفی انجام شده است.

یکی از مشکلات رایج هنگام آموزش مدل های پیش بینی، تطبیق بیش از حد مدل با خصوصیات داده های آموزشی فعلی یا به عبارتی پیش برازش^۱ است که این مسأله به پیش بینی نادرست و با دقت کمتر از حد انتظار برای مقادیر جدید منجر می شود. برای جلوگیری از چنین حالتی با استفاده از روش اعتبارسنجی تقاطعی^۲، داده ها به سه بخش داده های آموزشی^۳، داده های آزمایشی^۴ و داده های اعتبارسنجی^۵ تقسیم می شوند. مقداری از

کد نمونه	اندازه نمونه					
	آموزش		اعتبارسنجی		آزمایش	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
A	۱۵۸	۹۰	۹	۵	۹	۵
B	۱۴۰	۸۰	۱۸	۱۰	۱۸	۱۰
C	۱۲۴	۷۰	۲۶	۱۵	۲۶	۱۵
D	۱۰۶	۶۰	۳۵	۲۰	۳۵	۲۰

جدول ۵: تفکیک داده های ورودی به شبکه عصبی

- 1- Over fitting
- 2- Crossed Validation
- 3- Training Data
- 4- Testing Data
- 5- Validation Data

در هر مرحله از اجرا، شبکه حالت پیش آموزش پیدا نموده و بر این اساس پارامتر شبکه برای تکرار بیشتر بهبود نمی یابد که تعداد کل تکرار و همچنین تعداد تکرار تا توقف الگوریتم در حالت های مختلف در جدول ۶ آورده شده است. نمودار خطای شبکه برای داده های آموزشی، اعتبار سنجی و آزمایشی برای حالتی که تعداد تکرار ۱۰ بوده و الگوریتم در مرحله ۴ از تکرارها متوقف شده، در شکل ۱ آورده شده است. نکته قابل توجه دیگر ارتباط بین ضریب همبستگی و میانگین مربعات خطاست که برابر تحلیل ریاضی آن ها هر دو در خلاف جهت یکدیگر تغییر می کنند. بررسی اعداد جدول ۶ نشان می دهد که با توجه به اینکه در شبکه ساخته شده شماره ۵ و ۶ تعداد عصب های لایه پنهان افزوده شده است ولی پارامترهای شبکه شامل میانگین مربعات خطا و ضریب همبستگی بهبود نیافته اند، بنابراین مدل ساخته شده با تعداد ۲۰ عصب در لایه پنهان یعنی شبکه شماره ۴ به عنوان مدل بهینه انتخاب می شود.

تابع انتقال جهت آموزش شبکه تابع سیگموئید می باشد، الگوریتم آموزش شبکه، الگوریتم لوببرگ- مارکوات که بصورت TRAINLM، در نرم افزار مطلب بوده، استفاده می شود. در نهایت برای حالت های مختلف شبکه، با داشتن تعداد نرون متفاوت در لایه پنهان و تفکیک داده های متفاوت، شبکه عصبی مورد اجرا قرار گرفته و برای ۵ بار اجرای های متفاوت، میانگین پارامترهای شبکه (میزان همبستگی و میانگین مربعات خطا) که عمومی ترین پارامترهای ارزیابی شبکه های عصبی به شمار می روند، به شرح جدول ۶ محاسبه شده است.

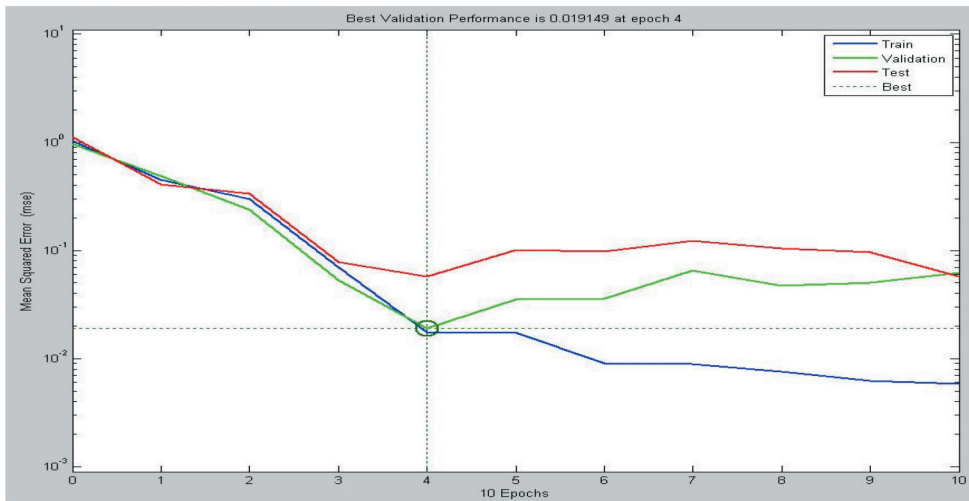
همان طوری که در جدول ۶ مشاهده می شود با افزایش تعداد اعصاب در لایه پنهان پارامترهای مربوط به شبکه بهبود می یابند. همچنین با افزایش تعداد عصب در لایه پنهان تعداد تکرار تا رسیدن به مرحله پیش آموزش نیز کاهش می یابد به این مفهوم که با افزایش تعداد عصب تکرارهای محاسباتی کمتری در رسیدن به جواب نهایی لازم است. نکته قابل توجه در اجرای شبکه عصبی آن است که پس از تکرار مشخصی

کد شبکه	کد نمونه	میزان همبستگی (R) ($\times 10^{-1}$)			میزان میانگین مربعات خطا (MSE) ($\times 10^{-2}$)			کل تکرار	توقف تکرار
		آموزش	اعتبارسنجی	آزمایش	آموزش	اعتبارسنجی	آزمایش		
۱	A	۷/۹۲	۷/۲۲	۸/۸۷	۲/۲۸	۲/۹۵	۲/۲۹	۱۹	۱۳
	B	۸/۲۷	۶/۹۱	۷/۹۵	۱/۹۱	۳/۴۴	۲/۲۳	۱۰	۴
	C	۷/۸۲	۷/۹۵	۶/۳۷	۲/۴۵	۲/۵۹	۲/۷۲	۱۷	۱۱
	D	۶/۹۴	۶/۲۹	۶/۸۷	۲/۳۴	۱/۶۵	۲/۲۴	۱۲	۶
۲	A	۷/۸۷	۸/۲۱	۷/۴۵	۲/۴۷	۳/۱۲	۱/۸۶	۱۱	۵
	B	۹/۰۳	۸/۹۵	۸/۸۷	۱/۸۲	۱/۵۷	۲/۲۹	۹	۳
	C	۸/۸۱	۷/۸۹	۸/۳۰	۲/۷۳	۲/۲۴	۲/۱۷	۱۱	۵
	D	۷/۸۲	۸/۲۳	۸/۳۱	۲/۹۱	۱/۴۰	۲/۷۶	۱۳	۷
۳	A	۹/۴۱	۸/۳۳	۹/۲۱	۱/۰۹	۱/۲۳	۱/۴۱	۹	۳
	B	۹/۲۱	۹/۵۲	۸/۸۷	-۱/۸۹	-۱/۹۲	۱/۱۲	۱۰	۴
	C	۹/۰۹	۸/۹۵	۹/۱۴	۱/۳۱	-۱/۵۵	۱/۰۳	۹	۳
	D	۸/۲۹	۹/۴۷	۹/۴۵	۱/۳۵	۱/۰۲	۱/۳۶	۱۰	۴
۴	A	۹/۶۹	۹/۸۳	۹/۷۱	-۱/۲۱	-۱/۲۴	-۱/۱۹	۹	۳
	B	۹/۳۶	۹/۷۹	۹/۷۴	-۱/۳۵	-۱/۲۲	-۱/۲۹	۸	۲
	C	۹/۶۶	۹/۵۵	۹/۶۴	-۱/۳۳	-۱/۲۴	-۱/۳۳	۹	۳
	D	۹/۲۴	۹/۶۷	۹/۶۵	-۱/۴۱	-۱/۴۹	-۱/۳۲	۸	۲
۵	A	۹/۰۶	۹/۴۹	۸/۹۵	-۱/۶۲	-۱/۲۳	-۱/۴۱	۱۰	۴
	B	۸/۹۷	۹/۶۲	۸/۷۰	-۱/۷۹	-۱/۵۹	-۱/۴۲	۱۰	۴
	C	۹/۴۴	۸/۵۵	۹/۵۴	۱/۷۳	۱/۲۴	۱/۰۳	۱۱	۵
	D	۹/۰۲	۸/۵۷	۹/۰۶	۱/۹۱	-۱/۴۵	۱/۵۲	۱۰	۴
۶	A	۸/۴۹	۸/۳۳	۹/۲۱	۱/۰۹	۱/۲۳	۱/۴۱	۹	۳
	B	۸/۳۶	۹/۵۲	۹/۱۴	۱/۵۵	-۱/۹۲	۱/۱۴	۱۰	۴
	C	۹/۲۶	۸/۷۹	۸/۴۴	۱/۷۱	-۱/۵۲	۱/۰۷	۱۰	۴
	D	۹/۳۱	۸/۹۲	۸/۶۵	۱/۳۹	۱/۰۵	۱/۳۲	۹	۳
کل داده ها در مدل بهینه		۹/۵۲			+۱/۳۸۱			۱۰	۴

جدول ۶: میزان همبستگی و میانگین مربعات خطا برای مدل های متفاوت شبکه عصبی با داده های متفاوت

میانگین مربعات خطا 0.00381 می باشد. نتایج نشان می دهد که با ضریب همبستگی محاسبه شده 90% درصد خروجی مدل یاد شده مطابقت مناسبی با داده های واقعی دارد.

پس از تعیین مدل بهینه شبکه عصبی، کل داده های ساخت شبکه یعنی 176 داده در مدل بهینه قرار گرفته که حاصل پارامترهای شبکه برای کل داده ها در سطر آخر جدول 6 قابل ملاحظه می باشد که در این حالت ضریب همبستگی (R^2) برابر با 90% درصد ($R=0.952$) و میزان

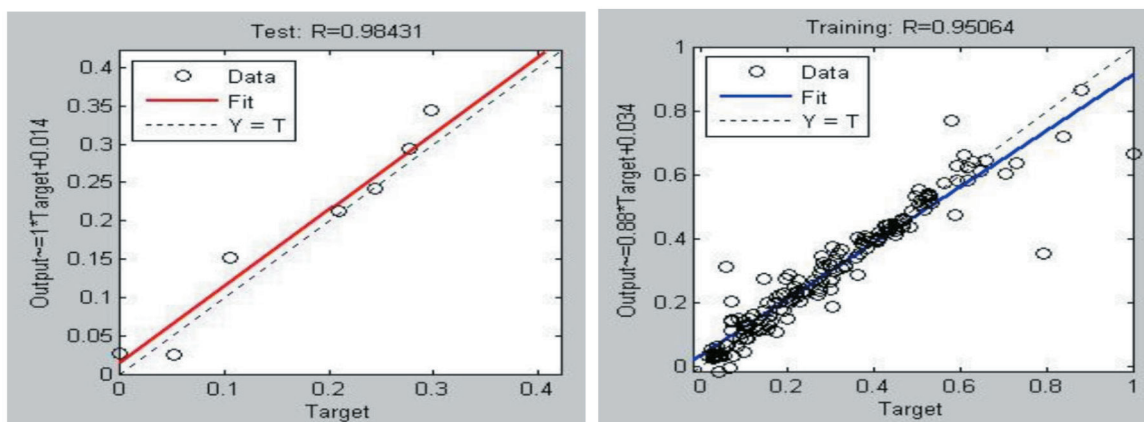


شکل ۱: خطای شبکه برای داده های آموزش، اعتبارسنجی و آزمایش

به مدل شبکه عصبی می باشد به صورت شکل های ۲ و ۳ ارزیابی شده است. همان گونه که در اشکال مزبور مشخص است، در معادلات رگرسیون داده های خروجی واقعی و خروجی شبکه، خروجی مدل شبکه عصبی و هدف^۱ (خروجی واقعی) در حد قابل قبولی منطبق با یکدیگر می باشد.

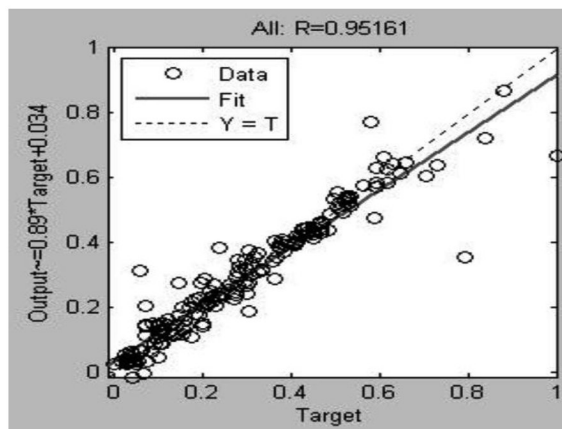
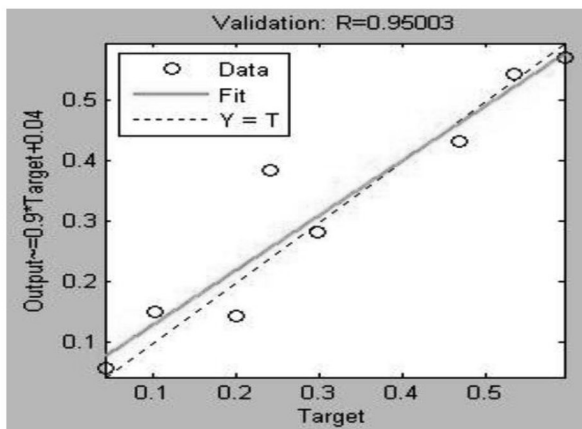
۳-۶- بررسی تطابق خروجی مدل شبکه عصبی با مشاهدات

به منظور ارزیابی بیشتر مدل شبکه عصبی، تطابق خروجی مدل با داده های واقعی در حالات مختلف داده های آموزشی، داده های اعتبارسنجی، داده های آزمایشی و کل داده ها مورد بررسی قرار می گیرد. تطابق داده های خروجی شبکه و خروجی واقعی که حاصل از گزارش ارزیابی شده در بخش رگرسیون مربوط



شکل ۲: تطابق خروجی های مدل و داده های واقعی در داده های آموزشی و آزمایشی

1- Target

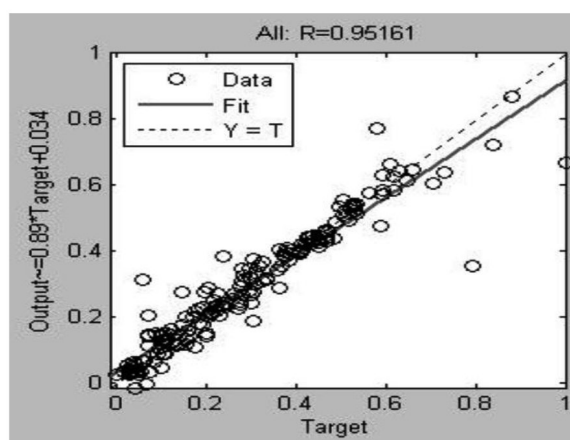
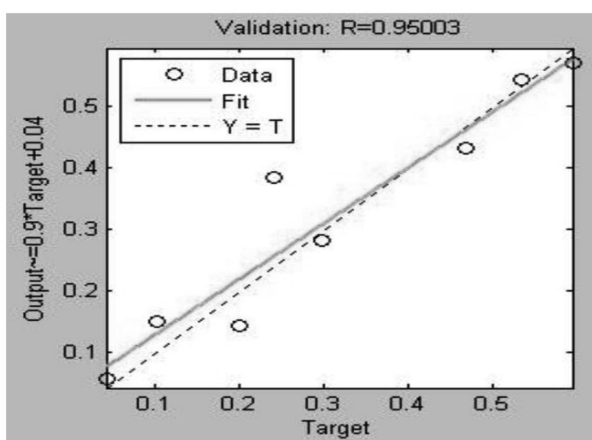


شکل ۳: تطابق خروجی های مدل و داده های واقعی در داده های اعتبارسنجی و کل داده ها

۴-۶- اعتبارسنجی مدل شبکه عصبی

در این قسمت، اعتبارسنجی مدل با استفاده از داده های جدید که در مدل وارد نشده بودند انجام می شود، یعنی از اطلاعات ۸ ماه اول سال ۹۳ یعنی ۳۲ داده آماری باقی مانده که از ابتدا برای این منظور کنار گذاشته شده بودند استفاده می شود و این داده ها وارد مدل بهینه می گردند، میانگین مربعات خطا در این حالت برابر 0.44900 و ضریب همبستگی (R^2) به میزان ۹۳ درصد نشان دهنده آن می باشد که مدل شبکه عصبی ساخته شده با واقعیت تطابق لازم را دارد (شکل ۴).

در کلیه تصاویر فوق ملاحظه می گردد شیب خط برازش در همه موارد بیش از 0.85 و عرض از مبدا نیز برای همه موارد ناچیز می باشد و از این رو می توان نتیجه گرفت که خروجی شبکه با خروجی واقعی تطابق لازم را دارد.



شکل ۴: تطابق خروجی های مدل و داده های واقعی در داده های تست

۵- نتیجه گیری

در این تحقیق پروسه ساخت مدل شبکه عصبی مصنوعی برای بررسی نقش عامل انسانی در سوانح رانندگی انجام و با توجه به بررسی های صورت گرفته در این تحقیق، نتایج زیر حاصل می شود.

۱- شبکه های عصبی پتانسیل لازم جهت رقابت با متدهای کلاسیک مدل سازی (روش های آماری) را دارا هستند. مخصوصا این ویژگی که این نوع مدل سازی نیاز به هیچ پیش فرضی برای تابع هدف ندارد باعث می شود، این متد را به عنوان ابزاری برای مدل سازی فرآیند هایی که هیچ توجیهی برای رابطه ریاضی بین متغیرها و هدف وجود ندارد، پذیرفته شود.

۲- ساخت مدل شبکه عصبی نیازمند تعریف تابع مشخصی نیست، بلکه شبکه عصبی با تعدادی داده و نمونه آموزش دیده و فرا می گیرد. بنابراین از جهت ساخت، مدل سازی با شبکه عصبی راحت است.

۳- شبکه عصبی ابزاری مناسب برای زمینه ای است که جوابها مهمتر از درک آن هستند. از طرفی برای بکارگیری مدل می بایست مقادیر به نرم افزار وارد و اجرا شده تا خروجی مناسب به دست آید.

۴- نتایج ارزیابی مدل شبکه عصبی ساخته شده به منظور پیش بینی تصادف براساس تخلفات حادثه ساز رانندگان، نشان از توانمندی بالای شبکه های عصبی در فرآیند مدل سازی دارند.

۶- مراجع

1. Accident Analysis and Prevention, 2004, Collisions and traffic violations of alcohol, cannabis and cocaine abuse clients before and after treatment. 36 (5) , PP.795-800.

2. Accident Analysis and Prevention, 2004, Sensitivity to punishment and sensitivity to reward and traffic violations, 36 (6), PP. 947-952.

3. Lee J. and Mannering F, 1999, Analysis Of Roadside Accident Frequency and Severity and Roadside safety Management, Washington State Department of Transportation , Technical Monitor.

4. Accident Analysis And Prevention, 1991, Analysis of the Frequency And Duration Of Freeway Accidents in Seattle, 23(4), PP. 239-255.

5. Accident Analysis and Prevention , 1993, Modeling Vehicle - Accidents and Environmental Factors on rural accident frequencies", 27(3) , PP. 371-389.

6. Journal of Safety Research, 2005, Identification Crash Propensity Using Traffic speed Condition, 36(1) , PP. 97- 108.

7. Safety Science, 2005, Analysis of highway accident frequencies :Negative binomial regression versus artificial neural network , 43(8), PP.541-557.

۸- مبانی شبکه های عصبی، ۱۳۸۹، محمدباقر منهج، دانشگاه صنعتی امیرکبیر، تهران.

۹- برآورد هزینه تصادفات در راه های ایران، ۱۳۸۹، پژوهشکده حمل و نقل وزارت راه وشهرسازی.