سنجش میزان توجه نسبت به تابلو «ایستادن ممنوع» و مؤلفه‌های مؤثر بر آن

امیرمسعود رحیمی*، استادیار گروه عمران، دانشکده مهندسی، دانشگاه زنجان، ایران
مجتبی کاظمی، کارشناس ارشد عمران گرایش راه و ترابری، دانشکده عمران، واحد رودسر و املش دانشگاه آزاد ، ایران
*amrahimi@znu.ac.ir

تاریخ پذیرش: تیر ۹۲ تاریخ دریافت مقاله: مرداد ۹۱

چکیده

ترافیک مجموعه‌ای از روابط و کنش‌های متقابلی است که میان کنشگران اجتماعی، خارج از محل سکونت‌شان حادث می‌شود. فرهنگ ترافیکی، عموماً در چارچوب آیین‌نامه و قوانین راهنمایی و رانندگی نمایان می‌شود که خود در برگیرنده‌ی تصویر و نوشتار است. یکی از این عناصر فیزیکی، تابلو انتظامی «ایستادن ممنوع» است که در این مقاله با تکمیل تعداد ۵۲۷ برگ پرسشنامه، به بررسی مقدار و عوامل مؤثر بر میزان توجه نسبت به آن اقدام شده است. تعداد متغیرهای مستقل بررسی شده، ۱۶ متغیر بوده که در نهایت با استفاده از الگوریتم کارت (شامل فرآیند پاک‌سازی جینی و هرس)، تعداد ۳ متغیر به عنوان متغیرهایی با توان بیشتر، انتخاب شده‌اند. سه متغیر تحصیلات، زمان رانندگی و تعداد ساعت‌های رانندگی در این تحقیق از اولویت بیشتری نسبت به سایر متغیرها برخوردار هستند.

کلید واژه : تابلو ایستادن ممنوع، الگوریتم کارت، پرسش گری.

۱ – مقدمه

اگر بخواهیم تعریف جامعه شناختی از ترافیک ارایه کنیم، شاید یکی از تعاریف این باشد: «ترافیک مجموع روابط و کنش‌های متقابلی است که خارج از محل سکونت میان کنش‌گران اجتماعی وجود دارد.» این کنش‌ها اساساً در شرایط و موقعیت‌های خاصی میان متغیرهایی چون انسان (سواره، پیاده، سرنشین و پلیس)، محیط طبیعی (شب و روز،کوهستانی، جلگه، دشت و ...)، محیط مصنوعی (جاده، خیابان،کوچه، میدان، چهار راه و ...)، اتومبیل (انواع اتومبیل) و حیوان (انواع حیوان) هنگام رفت و آمد و جا به جایی انسان و کالا برقرار است. فرهنگ ترافیکی اساساً در قالب آیین‌نامه و قوانین راهنمایی و رانندگی تبلور می یابد که خود حاوی اشکال (تصاویر) و نوشتارها می باشد. پرسش و پاسخ ها، تصاویر و علائم خبری، تصاویر و علائم اخطار و احتیاط، مهم‌ترین عناصر فرهنگ ترافیکی به دو گونه یاد شده است. در کنار قوانین و مقررات یاد شده، رفتارها و کنش‌های عموم بازیگران ترافیک، تکمیل کننده فرهنگ مکتوب هستند. بدین ترتیب، افراد جامعه به دو صورت، یعنی خواندن (کسب اطلاعات و یا شناخت هنجارهای قانونی راهنمایی و رانندگی از طریق مطالعه آیین نامه‌ها) و نیز مشاهده رفتارهای مردم با فرهنگ ترافیک آشنا می شوند (عبدالرحمانی، ۱۳۸۵).

در بخش دوم از تحقیق به بررسی برخی از مطالعات پرداخته شده که توسط افراد داخلی و خارجی، با هدف‌های مختلف و روش‌های متفاوت انجام شده است. پس از آن به تفسیر و تشریح کامل مدل استفاده شده در تحقیق پرداخته شده است. در بخش چهارم، نتایج حاصل از فرآیند مدل‌سازی، طبق شیوه وصف شده در بخش سوم، به صورت کامل آمده است و در انتها نیز بحث درباره‌ی نتایج به دست آمده به زبان نوشتار در آمده است.

۲ – پیشینه تحقیق

در پژوهشی در ایران به مطالعه ی طرح جانمایی علائم و تجهیزات ایمنی و ارتباط آن با کاهش تصادفات جاده‌ای پرداخته شده است. نتایج آن تحقیق برای محورهای زنجان- میانه و بهبهان- خیرآباد منجر به ترتیب کاهش ۲۶.۷ درصد و ۴۷ درصد تصادفات همسنگ در خسارات مالی و متوسط نسبت منفعت به هزینه‌ی حاصل از آن مطالعات به ترتیب برابر ۱.۴۷ و ۲.۴۵ شده است. در آن تحقیق برای دست‌یابی به نتایج از آزمون K و خی دو استفاده شده است (پورعبدل، ۱۳۸۵).

تحقیق دیگری که روی رانندگان ایرانی انجام شده،

نشان می‌دهد که جنسیّت رانندگان و میزان تحصیلات بر توجه رانندگان نسبت به علائم متغیّر خبری بی‌تأثیر بوده و رانندگان مرد و زن نیز به یک میزان به این نوع علائم توجه نشان می دهند. در آن تحقیق از روش های AHP، Kalman، MI و Cross استفاده شده است (زنگنه‌پور و همکاران. ۱۳۹۰).

هاشم ال مدنی بر اساس تحقیقی در پنج کشور بحرین، قطر، امارات، کویت و عمان، با استفاده از تکمیل پرسش نامه و به کار بستن روش چند متغیره دریافت که رانندگان جوان با سطح تجربه کم، درآمد پایین و داشتن سطح تحصیلات کمتر نسبت به رانندگان میانسال با تجربه ی رانندگی بیشتر، درآمد بالا و مدارک تحصیلی بالاتر، دقت کمتری نسبت به علائم خواهند داشت. نتیجه دیگر تحقیق ایشان آن بوده است که رانندگان مجرد و متأهل به یک نسبت علائم را درک می‌نمایند. بر اساس نتایج این تحقیق، رانندگان پنج کشور فوق‌الذکر، به طور صحیح ۵۰ ٪ تا ۶۰ ٪ علائم را درک می‌کنند (Al-Madani , Al-Jahani, ۲۰۰۲a). نتایج تحقیق دیگر نشان‌دهنده آن بوده که رانندگان جوان (زیر ۲۴ سال) نسبت به افراد مُسن درک پایین‌تری از علائم داشته‌اند، از این‌رو یک شخص میانسال (۲۴– ۳۵ سال) نسبت به یک شخص مُسن بهتر خواهد بود. این موضوع نیز حائز اهمیت می‌باشد که تجربه ی سال‌ها رانندگی نمی‌تواند در بهبود بخشیدن برای درک از علائم مؤثر واقع گردد (Al-Madani , Al-Jahani, ۲۰۰۲b).

تحقیقی دیگر نمونه‌ای ۱۰۹ عددی از رانندگان دارای گواهینامه را پس از تکمیل پرسشنامه، جهت دنبال نمودن تأثیرات عوامل رانندگی و ویژگی‌های طرح تابلو بر درک علائم ترافیکی مورد بررسی قرار داد. آن تحقیق با هدف موضوعات تسخیرشده ویژگی‌های شخصی، دسته‌بندی ویژگی‌های علائم، و نمره‌ی درک علائم طراحی شده بود. سال‌های رانندگی با گواهینامه و سطح تحصیلات، به صورت معناداری پیش بینی کننده ی درک تابلو قلمداد شدند. بر خلاف انتظار، پارامتر سن راننده، سال‌های رانندگی، ساعات رانندگی، زمان‌های گذشته‌ی رانندگی، فراوانی رانندگی و تجربه رانندگی در محیط نا آشنا تأثیری بر عملکرد درک نداشته‌اند. تابلوی آشنا با امتیاز درک برای رانندگان دارای گواهینامه در ارتباط بوده است، در حالیکه پیوستگی، سادگی و معناداری برای تابلو از چنین حالتی برخوردار نبوده‌اند. نتایج آن تحقیق خط مشی برای طراحی علائم ترافیکی مأنوس با استفاده در آینده فراهم نموده است (W.Y.Ng, H.S.Chan ۲۰۰۸).

تحقیق دیگری رانندگان مُسن را در حیطه‌ی کاری خود قرار داده است. حدود ۲۵ ٪ رانندگان مُسن و با تجربه با خواندن علائم مشکل داشته و عمده مشکلات دیگر همچون مکان تابلو، اندازه و وضوح حروف و وضوح پیام ارائه شده توسط تابلو رؤیت شده می‌باشد. فاصله ی دید خوب و قابل فهم علائم برای رانندگان مُسن ۶۵ تا ۷۰ درصد فاصله‌ای است که رانندگان جوان می‌توانند علائم را ببینند. طی اظهارات رانندگان در آن تحقیق، اغلب مواقع خواندن و تفسیر علائم جاده‌ای به شکلی سریع و کارآمد بسیار مشکل است و این شرایط، به خصوص در وضعیّت روشنایی کم چند برابر می‌گردد. عدم توجه رانندگان مُسن که بسیار زیاد است، از اصلی‌ترین عوامل تصادفات این گروه از رانندگان به شمار می‌رود. در کنار این موضوع مهم، اشتباه در تغییر خط یا دور زدن را نیز می‌توان اضافه نمود. رانندگان مُسن هنگامی از دیدن و خواندن علائم احساس راحتی می‌کنند که علائم روشن‌تر باشند، هرچند که تأثیرگذاری با درخشش بیشتر زیاد می‌شود. مشخص شده که افزایش سرعت، رانندگی طولانی مدت و سن بالا تأثیرگذاری مهمی در کاهش کارایی رانندگان مُسن ندارد (Bayam, Liebowitz, Agresti. ۲۰۰۵).

در تحقیق دیگری تأثیر عوامل پایه بر درک روان‌شناسی مورد مطالعه قرار گرفته است. برخی از این عوامل عبارتند از خصوصیات فیزیکی علائم، زاویه دید، کیفیت و چگونگی اطلاعات علائم، خصوصیات روحی و روانی راننده، خصوصیات فیزیولوژی و فیزیکی راننده. طی نتایج این تحقیق می‌توان بیان داشت که ۲۵ ٪ تصادفات ترافیکی از تضاد بین انسان و محیط جاده نشأت می گیرد. همچنین اینگونه بیان شده که ٪۹۰ اطلاعات رانندگی به واسطه ی درک دیداری علائم حاصل می‌گردد (Wang Fang , ۲۰۰۳).

۳ – روش‌شناسی

در این تحقیق از روش رگرسیون درختی و طبقه‌بندی استفاده شده‌است. طبقه بندی و رگرسیون دو مسأله مهم در علم آمار می باشند. الگوریتم طبقه بندی و درخت رگرسیون در برگیرنده سه وظیفه مهم است. اولین وظیفه این است که چگونه در هر مرحله داده‌ها را بخش‌بندی نماید. دومین وظیفه آن است که چه زمانی بخش بندی را متوقف نماید. آخرین وظیفه، چگونگی پیش‌بینی مقدار y برای هر x در یک بخش بندی (قسمت) است (Loh. ۲۰۰۸). همچنین این روش به سه دلیل عمده، نوع جذاب و مخصوص مدل‌ها می باشد. اول آن‌که نشان‌دهنده نتایج مدل به صورت آسان برای درک و تلفیق (شبیه‌سازی) توسط انسان است. دومین دلیل این است که درخت تصمیم مدلی ناپارامتریک می باشد، مداخله توسط کاربر نیاز ندارد و بسیار مناسب برای جستجوی دانش اکتشافی است. سوم این که الگوریتم قابل درجه‌بندی است، به مفهوم دیگر کارآیی درجه‌بندی مطلوب با افزایش اندازه نمونه آموزشی ارتباط دارد. این حالت برای درخت تصمیم مدل‌های ساخته شده وجود دارد و نیز صحت درخت تصمیم، همسنگ یا برتر از دیگر مدل‌ها می‌باشد (Dobra. ۲۰۰۲).

الگوریتم کارت(CART)

مدل کارت، ابزاری قدرتمند در تعیین مهم‌ترین متغیرهای

مستقل و حل مسائل دسته‌بندی و پیش‌بینی است. هدف تهیه درختی است که بتوان به وسیله آن متغیر وابسته یا همان کلاس را برای یک رکورد جدید پیش‌بینی و تعیین نمود. روش CART شاخه‌های خود را به صورت دوتایی و تنها بر اساس یک فیلد (متغیر مستقل) ایجاد می‌کند. یعنی هر گروه غیر برگ آن، به دو گروه دیگر تفکیک می‌گردد.

گوناگـــونی، معیاری است که برای ارزیابی شاخه‌ها به کار می‌رود. برای محاسبه گوناگونی در یک مجموعه از رکوردها روش‌های بسیاری وجود دارد که در تمامی آنها گوناگونی زیاد عبارت است از مجموعه‌هایی که از کلاس‌های گوناگون در خود داشته باشند. گوناگونی کم هم عبارت است از مجموعه‌هایی که اعضای یک کلاس در آن بر سایر کلاس‌ها غلبه کند. بهترین نحوه ایجاد شاخه آن است که گوناگونی در مجموعه‌ها را تا حد امکان کم کند. در مرحله بعد دو شاخه وجود دارد که هرکدام دارای یک سری رکورد می‌باشند (هر یک از رکوردهای گره بالاتر در یکی از شاخه‌ها قرار گرفته است). حال برای هر شاخه مثل قبل عمل می‌گردد. برای جداسازی هر گره به دو زیر گره، شـــاخص‌های مختلفی وجود دارد که معروف‌تـــرین آن برای داده‌های اسمی، شاخص جینی است که به شکل رابطه (۱) و (۲) تعریف می‌شود:

معادله ۱

$$P(j|m) = \frac{P(j,m)}{P(m)} \quad , P(j,m) = \frac{\pi(j) N_j(m)}{N_j} \quad , P(m) = \sum_{j=1}^{J} P(j,m)$$

معادله ۲

$$Gini(m) = 1 - \sum_{j=1}^{J} P^2(j|m)$$

که در آن J تعداد دسته‌ها یا همان متغیرهای هدف، (j)، احتمال اولیه مربوط به دسته j و توسط تصمیم‌گیرنده مشخص می‌شود.(N_{jm}) تعداد مشاهدات مربوط به دسته j در گره m، N_j تعداد کل مشاهدات مربوط به کلاس j در گره ریشه، ($P_{j|m}$)، احتمال قرارگیری مشاهدات مربوط به دسته j در گره m و (Ginim)، که همان شاخص جینی است، معرف عدم خلوص یا ناهمگنی در گره m است. به این معنی که مثلاً اگر همه مشاهدات در یک گره از یک دسته باشند،(Ginim)، برابر صفر و مُبَین کمترین ناخالصی و به عبارت دیگر بیشترین خلوص در گره است و برعکس، بیشترین مقدار(Ginim)، زمانی حاصل می‌شود که از همه مشاهدات به یک نسبت در گره وجود داشته باشند. شاخص جینی در هر گره برای تمام متغیرها محاسبه شده و متغیری به عنوان متغیر جداکننده انتخاب می‌شود که کمترین مقدار برای جینی از آن به دست آید. احتمال اولیه، مبین سهم هر یک از دسته‌ها در جامع مرجع است. رشد درخت بر اساس شاخص جینی از همان گره ریشه، که اولین گره بوده و در برگیرنده تمام مشاهدات است، آغاز شده و برای هر درختی که ایجاد میشود (شریعت مهیمنی، توکلی کاشانی، ۱۳۸۹).

رگرسیون درختی

رگرسیـــون درختـــی قابلیت طبقه‌بندی را ندارد. در عوض این بردار پاســـخ Y می‌باشد که نشـــان‌دهنده مقادیر پاسخ برای هر مشاهده در ماتریس متغیر X است. از آن جایی که درخت رگرسیـــــون پیش اختصـــاص طبقه‌بندی انجام نمی‌دهد، قواعد جداسازی طبقه‌بندی مشابه جینی یا دوتایی کاربردی نخواهند بود. جداسازی در رگرسیون درختی مطابق با الگوریتم حداقل مربع باقیمانده با دلالت بر آن که مجموع واریانس‌های مورد انتظار برای دو نتیجه‌گیری گره ها باید حداقل شده باشد، ساخته می شود.

معادله ۳ :

$$\underset{x_j \le x_j^R, j=1,2,\dots,M}{\arg\min} \left[P_l Var(Y_l) + P_r Var(Y_r)\right]$$

که در معادله(۳)، Var(Yl) و Var(Yr)، بردارهای پاسخ برای متنـــاظر بودن گره های فرزند چـپ و راست می‌باشد. $x_j \le x_j^R, j = 1, \dots, M$ بهینه پرسش های جداسازی که هر کدام رضایتمندی شرایط فرمول (۳) است.

الگوریتم حداقل مربع باقیمانده برابر با قاعده جداسازی جینی است. تابع ناپاکی جینی (۲) در توصیف کامل نکته واریانسها ساده است. اگر به مقادیر طبقه K مقدار ۱ و به مقادیر سایر طبقات عدد صفر را اختصاص دهند، بدین‌گونه واریانس نمونه این مقادیر برابر $p(k|t)[1-p(k|t)]$ خواهد بود. در مجموع توسط شماره (تعداد) طبقات K، میتوان معیار سنجش ناپاکی i(t) ذیل را به دست آورد:

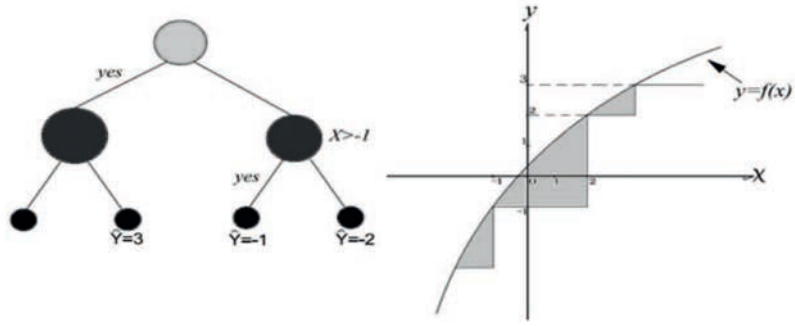
معادله ۴:

$$i(t) = 1 - \sum_{k=1}^{K} p^2(k|t)$$

در بالا این نکته به نام درخت حداکثر ساخته شده بود و به این معنا می‌باشد که جداسازی برای مشاهدات گذشته در نمونه آموزشی ساخته شده‌است. درخت حداکثر، شاید حاصل بسیار بزرگی باشد، به ویژه در نمونه رگرسیون درختی، زمانی که احتمال هر مقدار پاسخ نتیجه‌ای در یک گره مجزا باشد (Timofeev. ۲۰۰۴).

ارزیابی درخت ایجاد شده

برای ارزیابی درخت ایجاد شده توسط روش CART یا هر روش دیگری معیارهایی وجود دارند. از مهم‌ترین و اصلی‌ترین این معیارها نرخ خطا در درخت می باشد. رشد درخت بر اساس شاخص جینی از همان گره ریشه که اولین گره بوده و در برگیرنده تمام مشاهدات است، آغاز شده و برای هر درختی که ایجاد می‌شود، هزینه دسته‌بندی اشتباه آن (که می‌توان از آن به عنوان شاخص خوبی برازش یاد کرد) طبق رابطه (۵) محاسبه می‌شود:



شکل ۱: نمونه‌ای از درخت رگرسیونی

معادله ۵:

$$\text{misclassification cost} = \sum_{t=1}^{T} P(t)\left[1 - \sum_{j=1}^{J} P^{2}(j|t)\right]$$

که در آن $P(t)$، سهم مشاهدات موجود در گره نهایی t از کل مشاهدات بوده و T، تعداد گره‌های نهایی است. رابطه (۵) نمایانگر آن دسته از داده‌هایی است که به اشتباه در دسته‌های غیر مرتبط با خود، دسته‌بندی شده‌اند.

برای محاسبه نرخ خطا در درخت، ابتدا می‌بایست نرخ خطا در هر شاخه به دست آید. نرخ خطا در هر برگ عبارت است از نسبت تعداد رکوردهایی که کلاس یا دسته آنها درست پیش‌بینی نشده است. جهت برآورد نرخ خطای کل درخت، مجموع وزنی نرخ خطاهای برگ ها به دست آورده می شود (وزن هر برگ در واقع نسبت جمعیّت آن برگ به کل جمعیّت رکوردها می‌باشد). کیفیّت درخت حاصله نیز مهم خواهد بود. به منظور جلوگیری از تولید قانون های بی کیفیّت در بعضی از شاخه‌ها، قطع (هرس) صورت می گیرد. این کار با آن‌که نرخ خطا را افزایش می‌دهد ولی از ایجاد بعضی قانون‌های ناکارآ جلوگیری می‌نماید. همچنین باید به این نکته توجه داشت که باید قطع کردن به نحوی صورت گیرد که خطا از مقدار معینی بیشتر نشود (Tavakoli Kashani, Shariat Mohaymany, ۲۰۱۱) و (محجوبی، شهیدی، ۱۳۸۷). در تحقیق کنونی از نرم افزار PASW جهت دست‌یابی به بهترین درخت ایجاد شده بر اساس روش CART استفاده شده است.

٤ - تحلیل داده‌ها و بحث درباره‌ی نتایج آن

متغیرهای مستقل برای تحقیق، عبارتند از سن، جنس، وضعیت تأهل (مجرد یا متأهل)، میزان تحصیلات (بی‌سواد، زیر دیپلم، دیپلم یا فوق دیپلم، لیسانس، فوق لیسانس و بالاتر)، شغل (کارمند ادارات دولتی، پرسنل نیروهای مسلح، کارمند ادارات خصوصی، معلم یا استاد، محصل یا دانشجو، راننده، خدماتی، خانه‌دار، کارگر، بازنشسته، بیکار، آزاد)، سرعت رانندگی، تعداد ساعت رانندگی، بیشترین زمان رانندگی (اول صبح، صبح تا ظهر، بعد از ظهر، غروب، اول شب، آخر شب، نیمه شب، فرقی نمی‌کند)، تعداد تصادفات در ۵ سال گذشته، تعداد سال دریافت گواهینامه (کمتر از ۱ سال، ۱ تا ۵ سال، ۵ تا ۱۰ سال، بیشتر از ۱۰ سال)، تعداد سال تجربه ی رانندگی (کمتر از ۱ سال، ۱ تا ۵ سال، ۵ تا ۱۰ سال، بیشتر از ۱۰ سال)، نوع گواهینامه، استفاده از عینک به دستور پزشک (بلی، خیر)، نوع وسیله نقلیه (سواری شخصی، تاکسی و مسافربر شخصی، وانت بار و کامیونت، مینی‌بوس، اتوبوس، کامیون، تریلی، هیچ کدام)، حدود درآمد ماهیانه (کمتر از ۳۰۰ هزار تومان، ۳۰۰ تا ۴۵۰ هزار تومان، ۴۵۰ تا ۶۰۰ هزار تومان، ۶۰۰ تا ۸۰۰ هزار تومان، ۸۰۰ هزار تومان به بالا) و این که از بیماری خاصی رنج می‌برید (بله یا خیر).

تحلیل توصیفی برای تحقیق حاضر، در قالب جدول (۱) و برای متغیرهای جنسیت و سن افراد آمده است. بانک اطلاعاتی تحقیق، بر روش پرسش نام‌های و شیوه‌ی پرسش‌گری تصادفی، متکی می باشد. تعداد نمونه های مورد نیاز در تحقیق، با استفاده از فرمول کوکران، ۳۸۴ برگه پرسشنامه ارزیابی شد. تعداد پرسشنامه های تکمیل شده در تحقیق، ۵۲۷ برگه پرسشنامه بوده است.

شکل ۲ : تابلو انتظامی ایستادن ممنوع

جدول ۱ : تحلیل توصیفی بانک اطلاعاتی بر اساس سن و جنسیت رانندگان

سن و جنسیت		جنسیت		مجموع
		زن	مرد	
سن	زیر ۱۸ سال	٪۰	٪۰.۲	٪۰.۲
	۱۸ تا ۲۰	٪۹	٪۳.۴	٪۴.۴
	۲۱ تا ۲۹	٪۴.۶	٪۳۲.۳	٪۳۶.۸
	۳۰ تا ۳۹	٪۳	٪۲۳.۵	٪۲۶.۶
	۴۰ تا ۴۹	٪۰.۶	٪۱۹.۷	٪۲۰.۳
	۵۰ تا ۵۹	٪۰.۴	٪۸.۹	٪۹.۳
	۶۰ تا ۶۵	٪۰	٪۱.۹	٪۱.۹
	۶۵ سال به بالا	٪۰	٪۰.۶	٪۰.۶
مجموع		٪۹.۵	٪۹۰.۵	٪۱۰۰

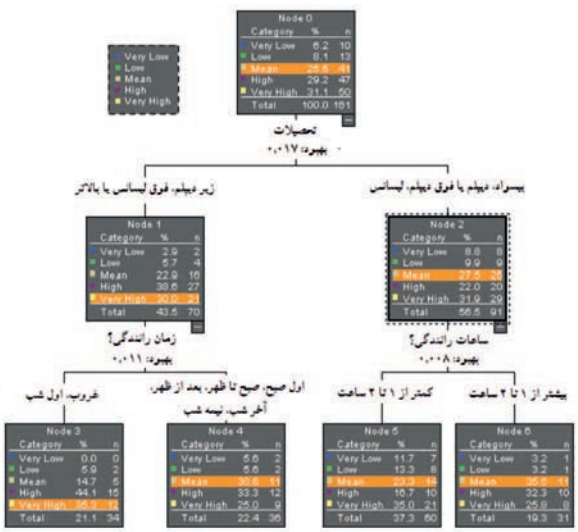
پس از تحلیل توصیفی بانک اطلاعاتی تحقیق، به فرآیند مدلسازی با استفاده از نرم افزار پرداخته شده است. پیش از انجام هرگونه عملی، ۷۰ ٪ داده‌ها به عنوان داده‌ی آموزشی و ۳۰ ٪ به عنوان داده‌ی آموزشی در نظر گرفته شده تا بتوان نسبت به صحت نتایج به دست آمده، اطمینان حاصل شود. فـــراوانی حداکثـــر برای گره‌ی والـــدین و فرزنـدان در نرم افزار که نشان دهنـده ی حضور جمعیت در هر گره می باشد، به ترتیب ۱۰۰ و ۵۰ بوده است. برای هرس درخت حاصله از «ضریب جینی» استفاده شده است. پس از تصحیح و آماده نمودن گزینه های نرم افزاری، فرآیند مدل سازی آغاز و در انتها خروجی آن به صورت درختی در شکل (۲) برای داده‌های آموزشی و شکل (۳) برای داده‌های آزمایشی آمده است. در جدول (۲) احتمالات اولیه برای توجه نسبت به تابلو مورد بحث آمده است. این مقادیر از داده‌های آموزشی به‌دست آمده‌اند.

جدول ۲: احتمال اولیه میزان توجه رانندگان برای داده‌های آموزشی نسبت به تابلو « ایستادن ممنوع»

میزان توجه	بسیار کم	کم	متوسط	زیاد	بسیار زیاد
احتمال اولیه	۰،۰۴۱	۰،۱۰۷	۰،۳۲۸	۰،۲۶۵	۰،۲۶۰

شکل ۳ : خروجی درخت رگرسیونی و تصمیم‌گیری برای داده‌های آموزشی

شکل ۴ : خروجی درخت رگرسیونی و تصمیم‌گیری برای داده‌های آزمایشی



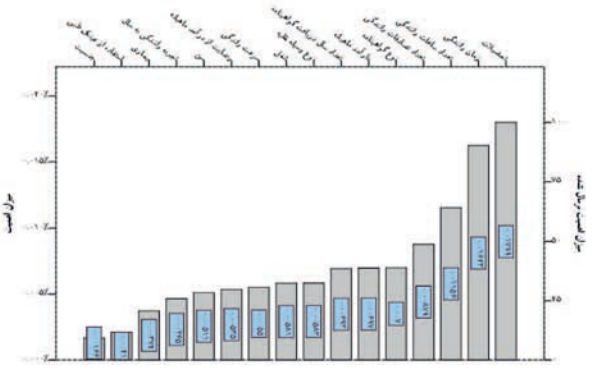
با نگاه به نتایج به دست آمده از مدل‌سازی در شکل‌های (۲) و (۳) می توان بیان نمود که میان مدل ایجـــاد شده بر اساس ۳۰ ٪ نمـــونه از داده‌های آزمایـــشی، با نتایج کســـب شده از ۷۰ ٪ داده‌های آموزشی، تشابه نزدیک برقرار است و این موضوع نشان‌دهنده‌ی صحت و اعتبار مدل می‌باشد.

مطابق با نتایج به دست آمده و دقت در نمودارهای ترسیمی، می‌توان ادعا نمود که میزان توجه رانندگان به تابلو راهنمایی و رانندگی «ایستادن ممنوع»، در حد «متوسط» بوده است. اولین پارامتری که به روی این میزان توجه تأثیر گذارده، سطح تحصیلات رانندگان است. در حقیقت می‌توان با استفاده از نتایج

حاصل شده اذعان داشت که سطح تحصیلات، با توجه به زمان و میزان رانندگی افراد، می‌تواند باعث تغییر در میزان سطح توجه افراد نسبت به این تابلو خاص شود. به شکل (۲) و (۳) رجوع شود.

۵ - نتیجه‌گیری

در این تحقیق شانزده متغیر مستقل که هر کدام به نوبه‌ی خود می‌توانند بر میزان توجه نسبت به تابلوی «ایستادن ممنوع» تأثیرگذار باشند، مورد ارزیابی قرار گرفته‌اند. پس از انجام فرآیند مدل‌سازی توسط الگوریتم CART برخی از متغیرهای با سطح اهمیت بالا به عنوان پارامترهای بارز، استخراج شده‌اند و مابقی متغیرها که نقش کمرنگ‌تری ایفا نموده‌اند به منظور جلوگیری از حجیم شدن درخت رگرسیونی و طبقه‌بندی و تفسیر منطقی نتایج، از نمایش در شکل نهایی کنار گذاشته شده‌اند (مطابق با خاصیت جینی). پارامترهایی که از مقدار اهمیت بیشتری برخوردار شده‌اند عبارتند از تحصیلات، زمان و تعداد ساعات رانندگی در طول روز که هر کدام به اندازه‌ی خود بر میزان توجه اثرگذار هستند. بر اساس روش‌های پاکسازی درخت در مدل CART، متغیرهای با اهمیت کمتر از فرآیند مدل‌سازی حذف شده‌اند. حذف متغیرها به مفهوم آن نخواهد بود که این متغیرها کاملاً بی تأثیر هستند، بلکه منظور آن است که این متغیرها نسبت به متغیرهای سطح تحصیلات، زمان و میزان ساعت‌های رانندگی تأثیر کمتری داشته‌اند (شکل ۵ را مشاهده فرمایید).



شکل ۵ : میزان اهمیت متغیرهای مستقل تاثیرگذار بر توجه نسبت به تابلو «ایستادن ممنوع»

٦ - منابع و مراجع

۱ - جامعه شناسی ترافیک، ۱۳۸۵، رضا عبدالرحمانی، معاونت آموزش دانشگاه علوم انتظامی ناجا، تهران.

۲ - پورعبدل، ن، ارزیابی میزان تأثیر طرح مطالعات جانمایی علائم و تجهیزات ایمنی با رویکرد کاهش تصادفات جاده‌ای، پایان‌نامه کارشناسی ارشد، ۱۳۸۵، دانشگاه علم و صنعت ایران.

۳ - زنگنه پور، س. لسانی، ا. مشیری، ب. رحیمی کیان، ا، عملکرد رانندگان در مواجه با تابلوهای متغیّر خبری و تأثیر عوامل مختلف آن، ۱۳۹۰، اولین همایش ملّی ترافیک، ایمنی و راهکارهای اجرایی ارتقای آن در محورهای درون شهری، برون شهری و روستایی، دانشگاه باهنر کرمان، بهار ۹۰.

۴ - افشین شریعت مهیمنی، علی توکلی کاشانی. تعیین شدّت مصدومیّت ناشی از تصادف ها در راه های دوخطه برون شهری با استفاده از مدل های داده کاوی. پژوهش نامه حمل و نقل، سال هفتم، شماره دوم، تابستان ۱۳۸۹، صفحات: ۱۵۳-۱۶۵.

۵ - جواد محجوبی، اسدا... شهیدی. پیش‌بینی پارامترهای امواج ناشی از باد در بندر امیرآباد به کمک درخت‌های تصمیم رگرسیونی. چهارمین کنگره ملّی عمران، دانشگاه تهران، اردیبهشت ۱۳۸۷.

6- Al-Madani. H , Al-Janahi. A.R, (2002), " Role of drivers' personal characteristics in understanding traffic sign symbols", Accident Analysis and Prevention 34, pp. 185–196.

7- Al- Madani. H, Al-Janahi. A.R, (2002), "Assessment of drivers' comprehension of traffic signs based on their traffic, personal and social characteristics ", Transportation Research Part F 5, pp. 63–76.

8- W.Y.Ng, A., H.S.Chan, A. (2008). The effects of driver factors and sign design features on the comprehensibility of traffic signs, Journal of Safety Research 39, June, pp. 321–328.

9- Bayam, E., Liebowitz, J., Agresti, W. (2005). Older drivers and accidents: A meta analysis and data mining application on traffic accident data, Expert Systems with Applications 29, pp. 598–629.

10- Wang Fang (2003). "Study on signs comprehension and driving safety Based on drivers' psychology perception", Research Institute of Highway,MOC No. 8 Xitucheng Rd., Haidian District, Beijing, pp. 298-304

11- Loh, W-Y. (2008). Classification and Regression Tree Methods, In Encyclopedia of Statistical in Quality and Reliability, Wiley, pp. 315- 323.

12- Dobra, A. (2002). Classification and Regression Tree Construction, Thesis Proposal, Department of Computer Science Cornell University, Ithaca, NY.

13- Timofeev, R. (2004). Classification and Regression Trees (CART) Theory and Applications, A Master Thesis, Center of Applied Statistics and Economics, Hambodelt University, Berlin.

14- Tavakoli Kashani, A., Shariat Mohaymany, A. (2011). Analysis of traffic injury severity on two-lane, two-way rural roads based on classification tree models. Safety Science 49. pp. 1314- 1320.