

شناسایی عوامل کلیدی و بررسی دقت پیش بینی مدل های داده کاوی در تصادفات ترافیکی محور زنجان-تبریز

بهزاد مسلم*^۱، فرزاد موحدی سبحانی^۲، عباس سقایی^۳

۱- کارشناسی ارشد، رشته مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران، تهران، ایران

۲- استادیار گروه مهندسی صنایع دانشکده فنی و مهندسی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران، تهران، ایران

۳- دانشیار گروه مهندسی صنایع دانشکده فنی و مهندسی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران، تهران، ایران

behzad.moslem@srbiau.ac.ir

چکیده

این پژوهش با هدف شناسایی و بیان عوامل مهم و موثر بر تصادفات و کشف عملکرد تکنیک های داده کاوی در پیش بینی تصادفات ترافیکی در محور زنجان-تبریز مورد بررسی قرار گرفت. برای دستیابی به اهداف، مدل های انتخاب خصیصه مبتنی بر همبستگی، مبتنی بر یکپارچگی و مدل ارزیابی زیر مجموعه فیلتر شده برای انتخاب خصیصه بر اساس ۳ روش جستجوی حریمانه، روش جستجوی تابو و روش جستجوی ژنتیک جهت شناسایی و انتخاب بهترین خصیصه ها استفاده گردید. سپس از تکنیک های مختلف داده کاوی جهت بررسی دقت مدل ها در پیش بینی تصادفات استفاده گردید. به همین منظور ۶ مدل داده کاوی، درخت تصمیم گیری جی ۴۸، قوانین استنتاجی (پارت)، بیزین ساده، پرسپترون چندلایه، لجستیک، و کلاس بندی از طریق رگرسیون با استفاده از نرم افزار داده کاوی وکا مورد استفاده قرار گرفت. برای آزمون داده ها برای هر مدل از سه روش داده های تمرینی، اعتبارسنجی متقاطع ۱۰ گانه و تقسیم درصدی با ۷۰ درصد آموزش و ۳۰ درصد آزمون استفاده گردید. پس از انجام محاسبات و ارزیابی های صورت گرفته نتایج نشان دادند که خصیصه های عامل تصادف، نوع برخورد، مانع دید، موقعیت تصادف، هندسه محل، عامل انسانی، علل اولیه، علل قبلی، علل مستقیم و تاریخ از مهمترین عوامل برای پیش بینی تصادفات ترافیکی در این محور هستند. همچنین نتایج نشان دادند که مدل های کلاس بندی از طریق رگرسیون، بیزین ساده و پرسپترون بهترین عملکردها را در پیش بینی تصادفات داشتند.

واژگان کلیدی: پیش بینی، تصادفات ترافیکی، داده کاوی، مدل سازی

۱- مقدمه

امروزه با افزایش روز افزون جمعیت در ایران و جهان، شاهد افزایش نیاز به تردد و جابه جایی افراد از مکانی به مکان دیگر هستیم. این تردها نیازمند ایجاد بستری ایمن و مناسب جهت حفظ جان انسان ها می باشد. طبق گزارش سازمان بهداشت جهانی، تصادفات جاده ای یکی از بزرگترین عوامل مرگ و میر در جهان به شمار می آیند. در سال های اخیر به دلیل افزایش قابلیت جمع آوری و ذخیره داده ها، تحلیل و استخراج اطلاعات مفید از آن ها بدون استفاده از ابزارهای تحلیل پیشرفته داده ها به یک چالش تبدیل می شود. برای کمک به مدیران در انجام تصمیم گیری ها، مجموعه بزرگی از داده ها جمع آوری می شوند. شناسایی دلایل تصادفات از این مجموعه داده ها، کاری پیچیده است؛ چرا که رابطه پیچیده ای میان عوامل آن وجود دارد. اینجا جایی است که داده کاوی مورد توجه قرار می گیرد. داده کاوی، یک تکنیک نوظهور در تجزیه و تحلیل داده می باشد که فرآیند استخراج اطلاعات (قوانین و الگوها) ضمنی، غیربدیهی، از قبل ناشناخته و بالقوه مفید از داده ها را بر عهده دارد. پژوهش حاضر شکاف های تحقیقاتی در زمینه دقت و صحت روش های مورد استفاده و عدم مقایسه روش های مختلف به منظور انتخاب روشی مناسب تر جهت انجام کلاس بندی و کشف الگوهای مناسب از آن ها، و عدم ارائه قوانینی جامع و قوانین پنهان در حوزه تصادفات جاده ای با