پژوهشنامهٔ بیمه (صنعت بیمهٔ سابق)/سال‌بیست‌وششم/شمارهٔ ۲/تابستان ۱۳۹۰/شماره‌مسلسل۱۰۲/صفحات ۵۵ تا ۸۱

مدلی جهت دسته‌بندی ریسکی گروه‌های مشتریان بیمهٔ بدنهٔ اتومبیل براساس ریسک بااستفاده از تکنیک داده‌کاوی

(مورد مطالعه: بیمهٔ بدنهٔ اتومبیل در یک شرکت بیمه‌ای)

پیام حنفی‌زاده[۱]
ندا رستخیز پایدار[۲]

تاریخ دریافت مقاله: ۱۳۹۰/۰۴/۲۳
تاریخ پذیرش مقاله: ۱۳۹۰/۱۰/۱۲

چکیده

در کشور ما بیمهٔ اتومبیل از مهم‌ترین رشته‌های بیمه‌ای است که سهم عمده‌ای در پرتفوی صنعت بیمه دارد. در ایران نرخ حق‌بیمهٔ بدنهٔ اتومبیل باتوجه به تعرفهٔ بیمه مرکزی ج. ا. ا. تعیین می‌شود. هدف این تحقیق ارائه راهکار به شرکت‌های بیمه جهت تعیین نرخ حق‌بیمهٔ بدنهٔ اتومبیل باتوجه به سطح ریسک هریک از مشتریان و یاری‌کردن سازمان‌های بیمه در جهت پیشبرد اهداف و به‌کارگیری استراتژی‌های مناسب در خصوص هریک از دسته‌های مشتریان و بهبود موقعیت فعلی در بازار است. در مدل ارائه‌شده در ابتدا عوامل تأثیرگذار بر ریسک بیمه‌گذاران طی دو مرحله شناسایی شد. در مرحلهٔ اول هیجده فاکتور ریسک در چهار گروه شامل مشخصات جمعیت‌شناختی، مشخصات اتومبیل، مشخصات بیمه‌نامه و سابقهٔ راننده از بین مقالات علمی منتشرشده در مجلات معتبر در بازهٔ سال‌های ۲۰۰۹- ۲۰۰۰ استخراج شد و در مرحلهٔ دوم بااستفاده از نظرسنجی از خبرگان، فاکتورهای نهایی تعیین گردید. پس از بخش‌بندی مشتریان بااستفاده از شبکهٔ خودسازمانده، ویژگی‌های مشتریان در هریک از بخش‌ها شناسایی شد. شناسایی ویژگی‌های مشتریان، اولین گام برای تعریف حق‌بیمه‌های متفاوت براساس سطح ریسک هر مشتری است.

واژگان‌کلیدی: *بخش‌بندی مشتریان، شبکهٔ خودسازمانده، بیمهٔ بدنهٔ اتومبیل*

۱. استادیار دانشگاه علامه طباطبائی(نویسنده مسئول) (Email:Hanafizadeh@Gmail.com)
۲. کارشناس‌ارشد مدیریت فناوری اطلاعات، دانشگاه علامه طباطبائی (Email:Paydarneda@Gmail.com)

Archive of SID
www.SID.ir

۱. مقدمه

امروزه صنعت بیمه از ضروری‌ترین بخش‌های اقتصادی هر کشوری است که وظایف و کارکردهای مهمی برعهده‌دارد. رشد صنعت بیمه در هر کشوری بیانگر توسعه‌یافتگی و افزایش پس‌اندازهای مالی است. در ایران نیز بیمه از مهم‌ترین عوامل حفظ و تضمین سرمایه است (جعفری صمیمی و مرادی، ۱۳۸۷). مشتریان تفاوت‌های اساسی و زیادی با یکدیگر دارند. به‌طورکلی مشتریان از دو دیدگاه با یکدیگر متمایزند؛ ازلحاظ ارزش آنها نزد سازمان و ازلحاظ تفاوت در نیازهایشان (حسین‌زاده، ۱۳۸۶). شناسایی مشتریان مستلزم تحلیل مشتریان هدف و دسته‌بندی‌کردن مشتریان است که منجربه یافتن گروه‌هایی از مشتریان سودآور براساس ویژگی‌های آنها می‌شود. به‌طورکلی بخش‌بندی مشتریان شامل تقسیم‌کردن کلیهٔ مشتریان به گروه‌های اصلی که شباهت بیشتری به یکدیگر دارند و قراردادن آنها در یک گروه است (Ngai et al, 2009). طبقه‌بندی ریسک در حقیقت به معنای گروه‌بندی مشتریان با خصوصیات ریسک مشابه است که احتمال بروز خسارت‌های مشابهی دارند. طبقه‌بندی ریسک بیمه‌گذاران برمبنای ویژگی‌های قابل مشاهده می‌تواند به شرکت‌های بیمه جهت کاهش زیان، افزایش نرخ پوشش بیمه و جلوگیری از وقوع انتخاب نامساعد در بازار بیمه کمک شایانی کند (ماجد، ۱۳۸۷). نرخ حق‌بیمه در بسیاری از شرکت‌های بیمه‌ای در خارج از کشور باتوجه به متغیرهای گوناگون جمعیت‌شناختی، مشخصات اتومبیل و سابقه خسارت بیمه‌گذار محاسبه می‌شود. کریدلر[1]، عضو هیئت مدیرهٔ یکی از شرکت‌های بیمه‌ای در ایالت واشنگتن در گزارشی نرخ حق‌بیمه را متکی بر عواملی همچون سن بیمه‌گذار، جنسیت و وضعیت تأهل بیمه‌گذار، نوع خودرو، محل زندگی، سابقهٔ خسارت، میزان کیلومتر مصرفی و ... اعلام کرد. این در حالی است که در کشور ما نرخ حق‌بیمهٔ بدنهٔ

1. Kreidler, 2008

Archive of SID

www.SID.ir

اتومبیل باتوجه به تعرفهٔ بیمه مرکزی ج.ا.ا تعیین می‌شود. این امر سبب می‌شود که مشتریانی با ریسک پایین‌تر، خسارت‌های مالی مشتریانی با ریسک بالاتر را جبران کنند و لذا تفاوت آنچنانی بین مشتریان با ریسک بالا و پایین وجود نخواهد داشت.

۲. تعریف مسئله

فقدان سنجش‌های تعیین ریسک افراد در بیمهٔ اتومبیل علاوه‌بر ناکاراکردن قراردادهای بیمه، منتج به تعیین نرخ‌های غیرعادلانه نیز می‌شود، چرا که به جای فرد، اتومبیل بیمه می‌شود و این امر موجب شده تا اکثر شرکت‌های بیمه در زمینهٔ بیمهٔ اتومبیل متحمل زیان شوند. درصورتی‌که اکثر کشورهای توسعه‌یافته با بهره‌مندی از سیستم طبقه‌بندی ریسک افراد سعی در افزایش بهره‌وری و سوددهی صنعت بیمه خود داشته‌اند.

در فرآیند مدیریت ارتباط با مشتری[1] به‌طورکلی چهار بخش اصلی دیده شده است که اولین و مهم‌ترین گام آن شناخت مشتری است (Ngai et al, 2009). اولین گام جهت تعیین نرخ حق‌بیمه‌های متفاوت به ازای هر مشتری باتوجه به سطح ریسک آنها، شناخت ویژگی‌های هریک از مشتریان است.

در این تحقیق فاکتورهای مهم و تأثیرگذار بر ریسک مشتریان بیمهٔ اتومبیل در دو فاز جداگانه شناسایی و انتخاب شد. جهت انتخاب فاکتورها در اولین گام، مطالعه و بررسی بر روی مقالات علمی منتشرشده در بازه زمانی سال‌های ۲۰۰۹ - ۲۰۰۰ میلادی در بین مجلات معتبر[2] و گزارش‌های علمی صورت گرفت. فاکتورهای استخراج‌شده به‌طورکلی در چهار دستهٔ اصلی شامل مشخصات جمعیت‌شناختی، مشخصات اتومبیل، مشخصات بیمه‌نامه و سابقهٔ راننده طبقه‌بندی گردیدند. در فاز دوم پرسش‌نامه‌ای باتوجه به عوامل شناسایی‌شده در فاز اول طراحی و توسط خبرگانی تکمیل شد که از بین کارشناسان بخش خسارت مالی، رؤسا و مدیران بیمهٔ بدنهٔ اتومبیل در ده شرکت بیمه‌ای انتخاب شده بودند.

1. Customer Relationship Management (CRM)
2. Science Direct, Emerald, IEEE, Proquest

Archive of SID

www.SID.ir

درنهایت پس از تجزیه‌وتحلیل پرسش‌نامه‌ها بااستفاده از روش آزمون سنگریزه[1] فاکتورهای نهایی ریسک انتخاب گردید. در گام بعدی پس از شناسایی فاکتورهای نهایی ریسک، مشتریان بیمهٔ بدنهٔ اتومبیل بااستفاده از روش شبکه عصبی خودسازمانده[2] براساس فاکتورهای تعیین‌شدهٔ بخش‌بندی شدند.

۳. مروری بر ادبیات موضوع

۱-۳. داده‌کاوی و بخش‌بندی مشتریان

به‌دلیل رشد سریع تکنولوژی اطلاعات، حجم داده‌های ذخیره‌شده در پایگاه داده‌ها به سرعت در حال افزایش است. تجزیه‌وتحلیل داده‌های ذخیره‌شده و تبدیل‌کردن آنها به اطلاعات و دانش قابل استفاده برای سازمان به ابزارهای بسیار قدرتمند نیاز دارد. داده‌کاوی، ابزار ارزشمندی است که در سال‌های اخیر از آن به‌طور گسترده جهت استخراج اطلاعات، جستجو روابط و الگوها در بین حجم عظیم داده‌های استفاده شده است. براساس تحقیقات، تکنیک‌های داده‌کاوی مختلفی جهت گروه‌بندی مشتریان به‌کاررفته‌است.

در چرخه مدیریت ارتباط با مشتری، می‌توان بخش‌بندی مشتریان و تجزیه‌وتحلیل مشتریان را عناصر اصلی در گام شناسایی مشتری نام برد. تکنیک‌های داده‌کاوی به‌کاررفته در زمینه بخش‌بندی مشتریان را می‌توان به دو دستهٔ اصلی شامل دسته‌بندی مشتریان و خوشه‌بندی مشتریان تقسیم کرد. براساس تحقیق جای و همکارانش[3] مشخص شد که از بین سی‌وچهار تکنیک داده‌کاوی، شبکه‌های عصبی بیشترین کاربرد را داشته است. شبکه‌های عصبی که بااستفاده از مغز انسان شبیه‌سازی گردیده،

1. Scree Test
2. Self – Organizing Map
3. Ngai et al, 2009

Archive of SID

www.SID.ir

کاربردهای زیادی در زمینه دسته‌بندی، خوشه‌بندی و پیش‌بینی داشته است جدول ۱ مطالعات صورت‌گرفته در زمینه شناسایی مشتریان را نشان می‌دهد.

۲-۳. مروری بر شبکه‌های عصبی خودسازمانده

شبکه‌های عصبی خودسازمانده، شبکه‌های عصبی بدون نظارتی هستند که قابلیت ارائهٔ خروجی شبکه در قالب نقشه‌های گرافیکی گویا و قابل فهم برای مدیران سازمان‌ها را دارند. عدم حساسیت شبکه خودسازمانده به تعداد دادگان تعلیم و حساسیت کم این نوع شبکه‌ها به وجود نویز در دادگان تعلیم، توانایی نمایش روابط خطی و غیرخطی بین متغیرها و قدرت بالا در دسته‌بندی دادگان از دیگر مزیت‌های این شبکه‌هاست که در این تحقیق جهت بخش‌بندی مشتریان در بیمهٔ بدنهٔ اتومبیل با توجه به عوامل شناسایی گردیده در ریسک به عنوان یکی از پرکاربردترین تکنیک‌های بخش‌بندی در داده‌کاوی استفاده شد. ازآنجایی‌که این شبکه قادر است خروجی شبکه را در قالب نقشه‌های گرافیکی نمایش دهد، سرعت درک و تفسیر نتایج برای مدیران و کارشناسان راحت‌تر خواهد بود.

جدول ۱. تکنیک‌های مختلف داده‌کاوی به‌کاررفته در شناسایی مشتریان

ابعاد مدیریت ارتباط با مشتری	هدف مدیریت ارتباط با مشتری	نوع عملیات داده‌کاوی	تکنیک داده‌کاوی	محققان
شناسایی مشتری	بخش‌بندی	دسته‌بندی	درخت تصمیم‌گیری	Kim et al, 2006
			شبکه خودسازمان‌دهنده، درخت تصمیم‌گیری و مدل زنجیرهٔ مارکو	Ha et al, 2002
		خوشه‌بندی	K-Means	Dennis et al, 2001
			شبکهٔ خودسازمان‌دهنده و درخت تصمیم‌گیری، تحلیل پوششی داده	Lee & Park,2005
	تحلیل مشتریان هدف	رگرسیون	رگرسیون منطقی	Hwang et al, 2004
		دسته‌بندی	درخت تصمیم‌گیری	Chen et al. 2003; Wu et al, 2005;Yu et al, 2005
		خوشه‌بندی	شبکهٔ خودسازمان‌دهنده	Lee et al, 2004
		مصورسازی	نقشهٔ مشتری	Woo et al. 2005

(Ngai et al, 2009)

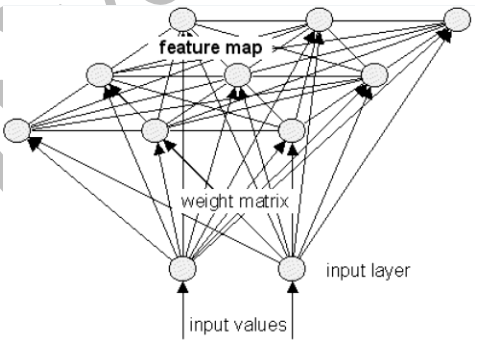
Archive of SID

www.SID.ir

اساس فلسفهٔ شبکه‌های خودسازمانده، نگاشت فضاهای با تعداد ابعاد بالا به فضایی دو یا سه بعدی است، به‌گونه‌ای که حداقل اطلاعات ازبین‌رفته و اطلاعات نهفته در ارتباط میان داده‌ها نیز قابل کشف و نمایش باشد. این روش توانایی نمایش همبستگی میان داده‌ها و اطلاعات و اثرات متقابل و هم‌زمان آنها بر یکدیگر را دارد (Vesanto, 1997).

شبکه‌های خودسازمانده از دو لایه مجزا تشکیل می‌شوند. یک لایهٔ ورودی و لایهٔ دیگر، لایهٔ نقشه نام دارد. لایهٔ نقشه مجموعه‌ای از نرون‌هاست که با قاعده در قالب یک شبکهٔ دو بعدی چیده‌شده و به لایهٔ ورودی کاملاً متصل است. هر نرون در لایهٔ نقشه مربوطبه یک بردار اطلاعات با ابعادی برابر ابعاد فضای مورد تحلیل است. به بیان‌دیگر هر نرون نمایندهٔ بخشی از فضای اطلاعات است. شکل۱ توپولوژی شبکه‌های خودسازمانده را نشان می‌دهد.

شکل۱. توپولوژی شبکه خودسازمانده



۱-۲-۳. آموزش شبکه‌های خودسازمانده

آموزش شبکه‌های خودسازمانده برمبنای الگوریتم یادگیری رقابتی و بدون ناظر (بدون استفاده از بردار هدف) است. در ابتدا بردار وزنی متناظر با هر نرون به‌طور تصادفی تولیدشده و ساختار اولیهٔ شبکه شکل می‌گیرد و سپس در طول فرآیند آموزش شبکه، بردار وزنی متناظر با هر نرون به‌گونه‌ای تنظیم می‌شود که بتواند

Archive of SID

www.SID.ir

قسمتی از اطلاعات فضای مورد تحلیل را پوشش‌دهد. بدین‌ترتیب، در قسمت‌هایی از فضا که چگالی داده‌ها بیشتر است، تعداد نرون بیشتری قرار خواهند گرفت و توپولوژی شبکه مطابق با توزیع مشترک مشخصه‌های فضا شکل می‌گیرد. الگوریتم آموزش شبکه‌های خودسازمانده چهار مرحله دارد (Kim et al, 2006):

– انتخاب پارامترهای نقشه مانند ابعاد و بردار وزن ابتدایی متناظر با هر نرون؛

– ارائه داده‌های مورد تحلیل به شبکه و یافتن بهترین نرون نظیر برای هر بردار داده ورودی؛

– به هنگام‌کردن بردار وزنی متناظر با هر نرون؛

– بررسی شرط خاتمهٔ الگوریتم.

چنانچه شرط بر قرار نباشد، الگوریتم از قدم دوم ادامه می‌یابد. ازآنجاکه الگوریتم آموزش شبکه‌های خودسازمانده برمبنای فاصلهٔ اقلیدسی بنا شده است، باید داده‌های هر بعد فضای مورد بررسی را مستقلاً نرمال استاندارد نمود. پس از پایان مرحلهٔ آموزش شبکه‌های خودسازمانده، نقشه‌ای از نرون‌ها به‌دست‌می‌آید که در حقیقت چکیده‌ای از فضای مورد تحلیل شبکه است. با ارائه هر بردار اطلاع جدید از فضای مورد تحلیل به شبکه، فاصلهٔ اقلیدسی بردار وزنی متناظر با هریک از نرون‌ها تا بردار ورودی، به‌دست‌آمده‌است؛ بنابراین مقدار تحریک هریک از نرون‌ها محاسبه و نرونی که بیش‌ترین مقدار تحریک را داشته باشد به‌عنوان نرون برنده انتخاب می‌شود.

۲-۲-۳. نمایش فضای مورد تحلیل بااستفاده از شبکه‌های خودسازمانده

پس از آموزش شبکه‌های خودسازمانده، به تعداد نرون‌های انتخاب‌شده برای شبکه، بردارهای وزنی بعدی به‌دست‌می‌آیند که هریک نمایانگر بخشی از فضای مورد تحلیل است. در صورت انتخاب تعداد مناسب نرون، ابعاد شبکه و درنهایت آموزش مناسب شبکه، نمایش بردارهای وزنی متناظر با نرون‌های هر نقشه می‌تواند به خوبی

Archive of SID

www.SID.ir

نمایانگر فضای مورد تحلیل باشد. بدین‌منظور متناظر با مقدار هر مشخصه در بردار وزنی یک بردار RGB[1] و در نتیجه یک رنگ درنظرگرفته‌می‌شود. به‌گونه‌ای که کلیه مقادیر بااستفاده از طیف رنگی، از آبی تیره (برای کمترین مقادیر) تا قرمز تیره (برای بیشترین مقادیر)، قابل نمایش باشند. بدین‌ترتیب به‌ازای هر مشخصه، رنگ هر نرون تعیین‌شده و نقشه متناظر با آن مشخصه به‌دست‌می‌آید. با به‌دست‌آمدن نقشه‌های مشخصات، بررسی ارتباط متقابل میان آنها (تست همبستگی) ممکن می‌شود. به‌عنوان‌مثال، هم‌رنگ‌بودن قسمت‌های متناظر دو نقشه، نشان‌دهندهٔ همبستگی مشخصه‌های نظیر با هر نقشه است. میزان همبستگی میان دو متغیر در قسمت‌های مختلف فضا نیز از شدت تفاوت یا تشابه رنگ میان نقشه‌های متناظر با هریک قابل بررسی است (Kreidler, 2008). شکل ۲، نشان‌دهندهٔ نمونه‌ای از کاربرد شبکه‌های خودسازمانده در تحلیل الگوهای پیچیده و نمایش هم‌زمان اثرات متغیرهای مختلف بر یکدیگر است. از مقایسهٔ نقشه‌ها با یکدیگر، موارد زیر قابل استنتاج است:

– متغیرهای V_2 و V_5 و همچنین V_1 و V_4، در تمام دامنهٔ تغییرات خود همبستگی معکوس دارند هر جا V_2 دارای رنگ قرمز است (مقادیر بالا به خود گرفته)، V_5 دارای رنگ آبی است (مقادیر پایین به خود گرفته است و بالعکس). گرچه شدت همبستگی V_2 و V_5 در تمام نقاط فضا تقریباً ثابت است، اما این مطلب در مورد متغیرهای V_1 و V_4 صادق نیست.

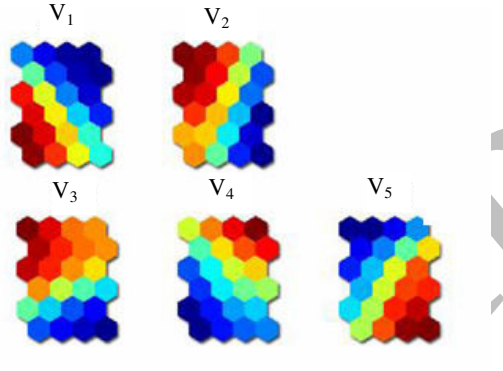
– متغیرهای V_3 و V_5 همبستگی معکوس دارند، اما شدت همبستگی در نقاط مختلف فضا متغیر و کمتر از شدت همبستگی V_2 و V_5 است.

۱. Red-Green-Blue: از فرمت‌های استاندارد تعریف رنگ‌هاست که هر رنگی را باتوجه به میزان شدت رنگ‌های اصلی و از ترکیب آن‌ها قابل حصول می‌سازد.

Archive of SID

www.SID.ir

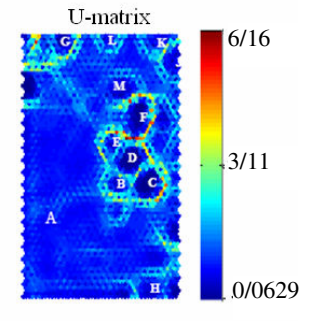
شکل ۲. کاربرد شبکه‌های خودسازمانده در تحلیل هم‌زمان روابط غیرخطی میان متغیرها



- ماتریس U [1]

از جمله خروجی‌های دیگر شبکه‌های خودسازمانده، ماتریس دسته‌بندی و متناظر با آن نقشهٔ دسته‌بندی است. درایه‌های این ماتریس، فاصله جبری نرون‌های همسایه را از یکدیگر نشان می‌دهند. بدیهی است انتظار داشته باشیم درصورتی‌که مشخصات دو قسمت از فضای مورد تحلیل به یکدیگر شبیه باشند، فاصله جبری بردارهای وزنی نرون‌های متناظر با آنها زیاد نباشد. یا به‌عبارت‌دیگر هر دو نرون، متعلق به خوشهٔ واحدی از فضای مورد تحلیل باشند. شکل ۳ یک ماتریس U را با تعدادی خوشه و زیرخوشه از یک فضای ۶۲ بعدی نشان می‌دهد.

شکل ۳. نمایشی از یک U_Matrix



1. U_Matrix

Archive of SID

www.SID.ir

۴. مدل پیشنهادی

در این تحقیق مشتریان بیمهٔ بدنهٔ اتومبیل در یک شرکت بیمه‌ای براساس ریسک آنها بخش‌بندی گردیدند. به‌طورکلی گام‌های مدل پیشنهادی شامل تعریف متغیرها، جمع‌آوری داده‌ها، پاکسازی داده‌ها، طراحی مدل و تجزیه‌وتحلیل داده‌هاست.

۴-۱. تعریف متغیرها

از عوامل مهم در ساخت یک مدل، انتخاب درست مشخصه‌هاست. جهت بخش‌بندی مشتریان براساس ریسک، اولین و مهم‌ترین گام اصلی، شناسایی عوامل ریسک است. شناسایی و انتخاب مشخصه‌ها در این تحقیق در دو فاز صورت گرفت.

- فاز اول) مطالعه و بررسی تحقیقات موجود

در این فاز، مقالات علمی معتبر در مجلات در بازه زمانی ۲۰۰۹-۲۰۰۰ بررسی شدند.

این مرحله از حساس‌ترین مراحل انجام این تحقیق است. در این قسمت هیجده مشخصه مطابق جدول ۲ استخراج گردید. سپس این مشخصه‌ها در چهار گروه مشخصات جمعیت‌شناختی، مشخصات خودرو، سابقهٔ راننده و مشخصات بیمه‌نامه برحسب نوع هریک از مشخصه‌ها تقسیم‌بندی گردیدند.

Archive of SID

www.SID.ir

جدول ۲. فاکتورهای استخراج‌شده ازادبیات تحقیق

عوامل / محققان	مشخصات جمعیت‌شناسی								مشخصات خودرو							سابقه راننده		مشخصات بیمه‌نامه
	سن	جنسیت	وضعیت تأهل	محل زندگی	نوع گواهینامه	سال اخذ گواهینامه (سابقه رانندگی)	میزان تحصیلات	میزان درآمد	نوع خودرو	رنگ خودرو	ظرفیت موتور	مدل خودرو	تجهیزات ایمنی (ABS)	میزان کارکرد بر حسب کیلومتر	ساخت داخل یا وارداتی	تعداد ادعای خسارت در سال قبل	سرعت رانندگی	میزان پوشش‌های بیمه‌ای
A	✓	✓		✓	✓	✓												
B	✓	✓				✓			✓					✓				
C	✓					✓												
D	✓	✓		✓					✓									
E	✓																	
F						✓												
G	✓	✓							✓									
H	✓				✓								✓	✓				
I	✓	✓				✓			✓			✓						
J	✓	✓		✓		✓												
K														✓				
L	✓	✓					✓											
M	✓	✓					✓	✓										
N	✓		✓															
O									✓					✓				
P	✓	✓	✓				✓		✓									
Q	✓	✓				✓												
R	✓	✓	✓	✓							✓	✓			✓	✓		
S	✓	✓		✓					✓									
T								✓									✓	
U	✓	✓												✓				
V										✓								
W	✓	✓							✓			✓						
X	✓					✓												
Y	✓	✓							✓			✓						
Z		✓																✓

(رستخیز پایدار، ۱۳۸۹)

Archive of SID

www.SID.ir

- فاز دوم) نظرسنجی از خبرگان

پس از استخراج مشخصه‌ها از مقالات و گزارشات علمی، پرسش‌نامه‌ای طراحی و توسط خبرگان تکمیل گردید. باتوجه به این موضوع که خبرگان باید از بین افرادی انتخاب گردند که در زمینهٔ خسارت‌های وارده به خودرو تجربه و دانش کافی داشته باشند، لذا کارشناسان خسارت مالی در بخش بیمهٔ بدنهٔ اتومبیل در شرکت‌های بیمه و رؤسا و مدیران آنها جهت نظرسنجی انتخاب گردیدند.

در کشور ما درحال‌حاضر هیجده شرکت بیمه در زمینهٔ بیمهٔ اتومبیل فعالیت دارند که تعداد کل کارشناسان، رؤسا و مدیران خسارت مالی در ده شرکت بیمه‌ای در تهران که سهم بازار آنها بیش از یک درصد است به‌عنوان جامعهٔ آماری انتخاب گردیدند. باتوجه به بررسی‌های صورت‌گرفته، تعداد کل کارشناسان خسارت مالی و رؤسا و مدیران واحد خسارت مالی در مجموع ۴۸۵ نفر در شهر تهران برآورد گردید.

جهت تعیین اندازهٔ نمونه باتوجه به اینکه جامعهٔ آماری، تعداد متناهی ولی بسیاری را شامل می‌گردد و باتوجه به روش نمونه‌برداری تصادفی، از روش محاسبهٔ حجم نمونه کوکران[1] استفاده شد. با استفاده از فرمول کوکران اندازهٔ نمونه ۱۱۴ نمونه انتخاب گردید. پرسش‌نامه‌ها به نسبت سهم بازار هریک از ده شرکت برگزیده توزیع گردید. از بین ۱۱۴ پرسش‌نامه توزیع گردیده در تهران، کارشناسان به ۸۷ پرسش‌نامه پاسخ دادند.

جهت انتخاب فاکتورهای نهایی و تحلیل نظرات خبرگان از روش آزمون سنگریزه استفاده شد. این روش را اولین‌بار کاتل[2] در سال ۱۹۶۶ به‌منظور تخمین گرافیکی فاکتورهای بااهمیت معرفی کرد. طبق نظرات کیسر[3] این روش از روش‌های بسیار

1. Cochran's Sample Size Formula
2. Cattell, 1966
3. Keiser

Archive of SID

www.SID.ir

پرکاربرد است که در اغلب نرم‌افزارهای آماری همچون نرم‌افزارهای SPSS و Minitab هم گنجانده شده است. آزمون سنگریزه در حقیقت مقادیر ویژه و ارزش هریک از فاکتورها را بر روی نمودار از صعودی به نزولی به تصویر می‌کشد. بدین‌ترتیب نقطهٔ شکست در انتخاب فاکتورها مشخص می‌گردد (Bener et al, 2006).

پس از تجزیه و تحلیل نتایج پرسش‌نامه علاوه‌بر هیجده فاکتور انتخاب‌شده، دو فاکتور شامل فاکتور سیستم خودرو (مانند پراید، زانتیا، پژو و...) و همچنین کاربری خودرو (مانند شخصی، تاکسی، اداری و...) باتوجه به نظر خبرگان به سایر عوامل اضافه گردید.

فاکتورهای نهایی انتخاب‌شده در این تحقیق مطابق جدول ۳ دسته‌بندی گردیدند.

جدول۳. فاکتورهای نهایی به‌دست‌آمده از نظرسنجی خبرگان

گروه عوامل	عوامل برگزیده
مشخصات جمعیت‌شناختی	سن
	جنسیت
	محل زندگی
	نوع گواهینامهٔ رانندگی
	سال اخذ گواهینامه (سابقه رانندگی)
مشخصات خودرو	سیستم خودرو
	کاربری خودرو
	نوع خودرو
	رنگ خودرو
	ظرفیت موتور
	مدل خودرو
	تجهیزات ایمنی (ABS)
	میزان کارکرد برحسب کیلومتر
سابقه رانندگی	تعداد ادعای خسارت در سال قبل
	سرعت رانندگی
مشخصات بیمه‌نامه	میزان پوشش‌های بیمه‌ای

Archive of SID

www.SID.ir

۱-۱-۴. روایی و پایایی پرسش‌نامه

مفهوم روایی یا اعتبار به این سؤال پاسخ می‌دهد که ابزار اندازه‌گیری تا چه حد خصیصه موردنظر را می‌سنجد. بدون آگاهی از اعتبار ابزار اندازه‌گیری نمی‌توان به دقت داده‌های حاصل از آن اطمینان داشت (ماجد، ۱۳۸۷). در این تحقیق سوالات پرسش‌نامه باتوجه به مطالعات و تحقیقات معتبر علمی طراحی گردید. در این تحقیق برای تعیین پایایی یا قابلیت اعتماد نتایج از روش آلفای کرونباخ استفاده شده است. این روش برای محاسبهٔ هماهنگی درونی ابزار اندازه‌گیری به‌کار می‌رود که فرمول آن عبارت است از:

$$\text{Alpha} = N/(N-1)[1-\sum \sigma^2(Y_j)/\sigma_x^2]$$

که در این فرمول N معرف تعداد زیرمجموعهٔ سؤال‌های پرسش‌نامه یا آزمون، $(Y_j)^2$ واریانس زیر آزمون jام و x^2 واریانس کل آزمون است.

داده‌های پرسش‌نامه بااستفاده از نرم‌افزار SPSS17.0 تحلیل شدند تا پایایی آن مشخص گردد. باتوجه به اینکه مقادیر بزرگ‌تر از ۰/۷ است، بنابراین آزمون از پایای قابل قبولی برخوردار است (الوانی و همکاران، ۱۳۸۳).

۲-۴. جمع‌آوری داده

در این تحقیق داده‌های موجود در پایگاه دادهٔ بیمهٔ بدنهٔ اتومبیل در یکی از شرکت‌های بیمهٔ داخلی استفاده شد. پنج فاکتور از فاکتورهای نهایی شامل سن، نوع گواهینامه و سال اخذ گواهینامه بیمه‌گذار و همچنین ظرفیت موتور و سرعت راننده به‌دلیل موجودنبودن اطلاعات در پایگاه داده در ابتدای امر حذف گردید. باتوجه به اینکه فراوانی سه نوع خودروی پراید، زانتیا و وانت مزدا با سایر انواع خودروهای بیمه‌شده در این شرکت تفاوت چشمگیری داشتند، لذا تنها سه نوع خودرو نامبرده در این تحقیق تجزیه‌وتحلیل شد. ازآنجاکه داده‌های مورد نیاز برای برخی از فاکتورها همچون میزان کارکرد (برحسب کیلومتر مصرفی) تنها در سال ۱۳۸۸ ورود

Archive of SID

www.SID.ir

اطلاعات گردیده بود، لذا تنها پایگاه دادهٔ بیمهٔ بدنهٔ اتومبیل در سال ۱۳۸۸ بررسی شد. همچنین به‌دلیل آنکه یکی از فاکتورهای برگزیده، سابقهٔ خسارت است لذا بیمه‌گذارانی مورد بررسی قرار گرفتند که سابقهٔ خسارت آنها در سال گذشته در پایگاه داده موجود بوده است. خسارت در بیمهٔ بدنه به‌طورکلی در دو دسته جداگانه خسارت‌های مربوطبه تصادفات و خسارت‌های مربوطبه سرقت قرار دارند. از آن جهت که این دو مسئله از یکدیگر کاملاً متفاوت‌اند و همچنین آمار مربوطبه سرقت‌های اعلام‌شده در بازه یک‌ساله بسیار پایین است لذا در این تحقیق داده‌های مربوطبه آمار سرقت‌های اعلام‌گردیده حذف شد. همچنین لازم به ذکر است که خودروهایی بررسی شدند که سال ساخت آنها در بازهٔ زمانی بین سال‌های ۱۳۸۸-۱۳۸۰ قرار داشت. تعداد رکوردهای موجود پس از اعمال تغییرات بیان گردیده در مجموع برابر ۳۴۷۲۶ رکورد است.

۳-۴. پیش‌پردازش داده

پیش‌پردازش داده‌ها از گام‌های مهم در فرآیند داده‌کاوی است که میزان دقت نتایج به‌دست‌آمده رابطهٔ قوی با نحوهٔ انجام آن دارد. برای پیش‌پردازش داده‌های تعلیم شبکه، فعالیت‌های ذیل انجام گرفت:

- حذف دادگان نامناسب؛
- تبدیل کلیهٔ مقادیر به مقادیر عددی؛
- نرمال‌سازی مقادیر.

باتوجه به اینکه مبنای آموزش شبکه‌های خودسازمانده براساس فاصلهٔ اقلیدسی بنا شده است، نرمال‌بودن دادگان تعلیم شبکه و وجود مقادیر خارج از الگو می‌تواند نقشه‌های خروجی را به‌شدت تحت تأثیر قرار دهد. به‌همین‌جهت نرمال‌سازی داده‌های ورودی، گام مهمی در آموزش شبکه‌های خودسازمانده است. در این تحقیق برای نرمال‌سازی دادگان تعلیم، روش نرمال‌سازی Z که کاربردهای فراوانی در آمار

Archive of SID

www.SID.ir

دارد، به‌کار رفت. در این روش نرمال‌سازی از طریق تقسیم اختلاف داده‌ها با میانگین بر انحراف معیار جامعه انجام می‌شود.

۴-۴. طراحی و ساخت مدل

در این تحقیق از شبکهٔ خودسازمانده به‌عنوان یکی از روش‌های پرکاربرد در تکنیک خوشه‌بندی در داده‌کاوی استفاده گردیده است. در ابتدا داده‌ها بااستفاده از تکنیک شبکه‌های خودسازمانده آموزش داده شد و نتایج حاصل از خروجی شبکه در قالب نقشه‌های گرافیکی تحلیل شد.

۴-۴-۱. شبکهٔ خودسازمانده

اساس شبکه‌های خودسازمانده نگاشت فضاهای با ابعاد بالا به فضای دو بعدی یا سه بعدی نقشهٔ خروجی است، به‌گونه‌ای که کم‌ترین میزان اطلاعات ازبین‌رفته و دانش نهفته در روابط میان دادگان تعلیم آشکار شود. داده‌های تعلیم شبکه از ۳۴۷۲۶ بردار یازدهٔ بعدی تشکیل شده است که هر بردار نمایندهٔ یک رکورد از رکوردهای پایگاه دادهٔ بیمهٔ بدنهٔ شرکت مورد مطالعه است. ابعاد این بردار برابر تعداد فاکتورهایی است که پس از نظرسنجی از خبرگان انتخاب گردیدند. رابطه زیر مجموعه دادگان تعلیم شبکه را تعریف می‌کند:

$$X_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6}, X_{i7}, X_{i8}, X_{i9}, X_{i10}, X_{i11}) \quad i: ۱ \text{ to } ۳۴۷۲۶$$

در مورد تعداد نرون‌های لایهٔ نقشه تحقیقات زیادی انجام گرفته است. کوهنن[1] مبدع این شبکه‌های خودسازمانده در کتابی با همین عنوان، فرمول $۵\sqrt{n}$ را برای تعداد نرون‌های لایهٔ نقشه توصیه می‌کند، که در آن n تعداد داده‌های آموزش است. در این تحقیق نیز ازهمین‌رابطه استفاده‌شده و باتوجه به اینکه تعداد داده‌های آموزش۳۴۷۲۶ نمونه بوده است، تعداد نرون‌های لایهٔ نقشه ۹۳۲ نرون انتخاب‌شده

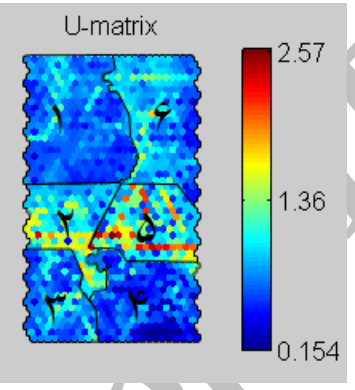
1. Kohonen

Archive of SID

www.SID.ir

است. شکل۴ ماتریس U حاصل از شبکهٔ خودسازماندهٔ آموزش‌دیده به‌وسیلهٔ داده‌های بیمهٔ بدنهٔ اتومبیل را نشان می‌دهد.

شکل ۴. ماتریس U شبکهٔ خودسازمانده پس از آموزش داده‌های شرکت بیمهٔ مورد مطالعه



ماتریس U به‌دست‌آمده، شش بخش یا شش خوشهٔ اطلاعاتی دارد و هریک از مشتریان بیمهٔ اتومبیل در یکی از این شش خوشه قرار می‌گیرند. مشتریان متعلق به هریک از این بخش‌ها از لحاظ مشخصه‌های بررسی‌شده، دارای شباهت نزدیک به یکدیگر بوده و اختلاف زیادی با مشخصه‌های مشتریان موجود در سایر بخش‌ها دارند. به‌همین‌جهت انتظار می‌رود که ریسک مشتریان قرارگرفته در یک گروه نیز مشابه یکدیگر بوده و با ریسک سایر مشتریان در گروه‌های دیگر متفاوت باشد.

جهت تعیین برچسب و نام هریک از خوشه‌ها تعدادی نمونه تصادفی از هر بخش انتخاب گردید و میانگین ضریب خسارت نمونه‌های انتخابی محاسبه شد. جدول ۴ نحوه برچسب‌زنی بخش‌های نقشهٔ خودسازمانده را نشان می‌دهد.

Archive of SID

www.SID.ir

جدول ۴. برچسب‌زنی گروه‌های مختلف مشتریان براساس ریسک

بخش	نمونه	ضریب خسارت	میانگین ضریب خسارت	میزان ریسک بخش
۱	۱	۰/۱۵	۰/۱۵	ریسک پایین
	۲	۰		
	۳	۰/۳۶		
	۴	۰		
	۵	۰/۲۴		
۲	۱	۰/۳۸	۰/۴	ریسک نسبتاً پایین
	۲	۰/۴۴		
	۳	۰/۵۴		
	۴	۰/۶۵		
	۵	۰		
۳	۱	۰	۰/۱۸	ریسک پایین
	۲	۰/۳		
	۳	۰		
	۴	۰/۳۶		
	۵	۰/۲۴		
۴	۱	۱/۴	۱/۷۲	ریسک بالا
	۲	۲/۲		
	۳	۰/۹		
	۴	۱/۶		
	۵	۲/۵		
۵	۱	۱	۱/۳۶	ریسک بالا
	۲	۱/۲		
	۳	۰/۹		
	۴	۱/۶		
	۵	۲/۱		
۶	۱	۰/۷۷	۰/۷۳	ریسک نسبتاً بالا
	۲	۰/۶۵		
	۳	۰/۸		
	۴	۰/۷۸		
	۵	۰/۶۶		

میانگین ضریب خسارت خوشهٔ یک و سه نزدیک به هم است، لذا هر دو خوشه در طبقهٔ مشتریان با ریسک پایین جای گرفتند. همچنین دو خوشهٔ چهار و پنج نیز

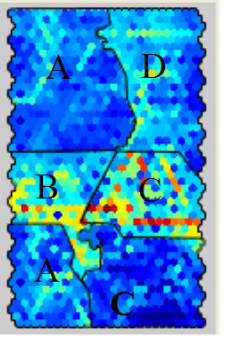
Archive of SID

www.SID.ir

میانگین ضریب خسارت نزدیک به هم دارند، لذا این دو خوشه هم باتوجه به آنکه میانگین ضریب خسارت آنها بالاتر از یک است - بدین معنا که خسارت واردشده به بیمه از سوی مشتریان این خوشه بیش از حق‌بیمهٔ پرداختی است- لذا این دو خوشه نیز در طبقه مشتریان با ریسک بالا جای گرفتند.

به‌طورکلی می‌توان ماتریس U را به چهار خوشه مطابق شکل ۵ تقسیم‌بندی کرد. مشتریان خوشهٔ A مشتریانی با ریسک پایین، مشتریان خوشهٔ B مشتریانی با ریسک نسبتاً پایین، مشتریان خوشهٔ C مشتریانی با ریسک بالا و مشتریان خوشهٔ D مشتریانی با ریسک نسبتاً بالا می‌باشند.

شکل ۵. تعیین برچسب هریک از خوشه‌های ماتریس U

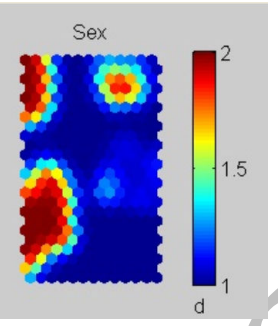


به‌طورکلی نقشه‌های رسم‌شده امکان سه نوع تحلیل مختلف را برای ما فراهم می‌کنند. ابتدا می‌توان در مورد هریک از متغیرها تنها بااستفاده از نقشهٔ رسم‌شده آن اطلاعاتی را استخراج کرد. به‌عنوان نمونه در نقشهٔ متغیر جنسیت بیمه‌گذار مطابق شکل۶ رنگ آبی تیره و قرمز تیره به‌ترتیب نمایانگر جنسیت مرد و زن است. همان‌گونه که در نقشه مشخص است فراوانی بیمه‌گذاران مرد نسبت به زن بیشتر است.

Archive of SID

www.SID.ir

شکل ۶. نقشه متغیر جنسیت بیمه‌گذار



از مقایسهٔ ماتریس U رسم‌شده با نقشهٔ تک‌تک متغیرها امکان تحلیل ویژگی‌های هریک از خوشه‌ها فراهم می‌گردد. در ادامه به‌عنوان نمونه خوشهٔ با ریسک بالا و خوشهٔ با ریسک نسبتاً پایین تحلیل می‌گردد:

- تحلیل خوشهٔ با ریسک بالا

- سیستم خودرو: باتوجه به طیف رنگی در این ناحیه متوجه می‌شویم که تعداد خودرو وانت در مقایسه با خودرو زانتیا و پراید بیشتر است. همچنین نسبت خودرو پراید در مقایسه با خودرو زانتیا بیشتر است.
- تجهیزات ایمنی (ترمز ضدقفل[1]): اکثر خودروهای این ناحیه فاقد سیستم ترمز ABS می‌باشند. برخی از خودروهای زانتیا و پراید در این ناحیه مجهز به این سیستم ترمزند.
- رنگ خودرو: تقریباً کلیه طیف‌های رنگی در این قسمت دیده می‌شود. خودروهای با رنگ‌های تیره، رنگی‌های تند (مانند نارنجی، زرد، آلبالویی و ...) و رنگ‌هایی از خانواده رنگ آبی در این قسمت قابل مشاهده است. رنگ‌های روشن (مانند سفید و کرم) در مقایسه با رنگ‌های تیره و سایر رنگ‌ها در این ناحیه کم‌تر است.

1. Anti-Lock Break System (ABS)

Archive of SID

www.SID.ir

- مدل خودرو: سال ساخت خودروهای این ناحیه غالباً پایین است.
- میزان کارکرد برحسب کیلومتر: میزان کارکرد بسیار بالا با رنگ قرمز تیره و کیلومترهای بالا با رنگ قرمز روشن مشخص گردیده است. اغلب خودروها در این ناحیه میزان کیلومتر مصرفی بالا دارند. در برخی از قسمت‌های نقشه هم میزان کیلومتر مصرفی پایین و در بعضی جاها حتی صفر کیلومتر هم دیده شده است.
- نوع خودرو: باتوجه به آنچه در نقشه دیده می‌شود در مقایسه با خودروهای وانت، تعداد کمی از خودروها در این قسمت از نوع سواری‌اند.
- کاربری خودرو: باتوجه به اینکه خودروهای این منطقه بیشتر از نوع وانت‌اند کاربری این نوع خودرو از نوع بارکش است. نکته قابل توجه در مورد خودروهای پراید در این ناحیه این است که اغلب خودروهای پرایدی که کاربری تاکسی یا آژانس دارند در این ناحیه قرار گرفتند.
- جنسیت بیمه‌گذار: جنسیت غالب بیمه‌گذاران این قسمت مرد است.
- شهر محل زندگی: در نقشهٔ این متغیر طیف رنگ‌های آبی پررنگ، آبی کم‌رنگ و زرد قابل مشاهده است که به‌ترتیب نمایانگر کلان شهر و شهرهای بزرگ و نسبتاً بزرگ است.
- میزان پوشش بیمه‌ای خریداری‌شده: میزان پوشش بیمه‌ای خریداری‌شده در این ناحیه بسیار پایین است.
- تعداد خسارت در سال قبل: بیشترین طیف رنگ قرمز در این ناحیه مشاهده می‌شود که نمایانگر میزان سابقهٔ خسارت‌های بالا (سابقهٔ خسارت ۲، ۳ و ۴) است. همچنین بسیاری از خودروهای این ناحیه دارای سابقهٔ خسارت یک می‌باشند.

Archive of SID

www.SID.ir

- تحلیل خوشه با ریسک نسبتاً پایین

- سیستم خودرو: هر سه نوع خودرو پراید، زانتیا و وانت در این ناحیه دیده می‌شود. تقریباً پراکندگی خودروی پراید و زانتیا در این بخش مشابه هم است. در مقایسه با این دو نوع خودرو، تعداد کم‌تری خودرو وانت در این ناحیه قرار گرفته است.
- تجهیزات ایمنی (ABS): خودروهایی که مجهز به سیستم ترمز ABS یا فاقد این سیستم‌اند، هر دو در این ناحیه قابل مشاهده‌اند.
- رنگ خودرو: کلیه طیف‌های رنگی در این ناحیه قابل مشاهده است.
- مدل خودرو: همه مدل‌های خودرو در این ناحیه وجود دارد. اغلب خودروهای این ناحیه مدل‌های ساخت بالا دارند.
- میزان کارکرد برحسب کیلومتر: طیف‌های رنگی آبی کم‌رنگ، نارنجی و زرد بیشتر در این ناحیه دیده می‌شود که نشان‌دهنده کارکردهای پایین و نسبتاً بالاست. همچنین در برخی قسمت‌ها خودروهای صفر کیلومتر هم دیده می‌شود.
- نوع خودرو: هر سه نوع خودرو در این ناحیه قرار دارد.
- کاربری خودرو: کاربری خودروهای این ناحیه شامل خودروهای شخصی و بارکش است.
- جنسیت بیمه‌گذار: هر دو نوع جنسیت در این ناحیه دیده می‌شود. باتوجه به نقشه می‌توان فهمید که تعداد بیمه‌گذاران زن در این ناحیه بیشتر از مردان است.
- شهر محل زندگی: بیشتر بیمه‌گذاران این ناحیه در شهرهای بزرگ و نسبتاً بزرگ قرار دارند.
- میزان پوشش بیمه‌ای خریداری‌شده: باتوجه به طیف رنگی آبی تیره و آبی روشن می‌توان فهمید که میزان پوشش بیمه‌ای خریداری‌شده در این ناحیه پایین است.

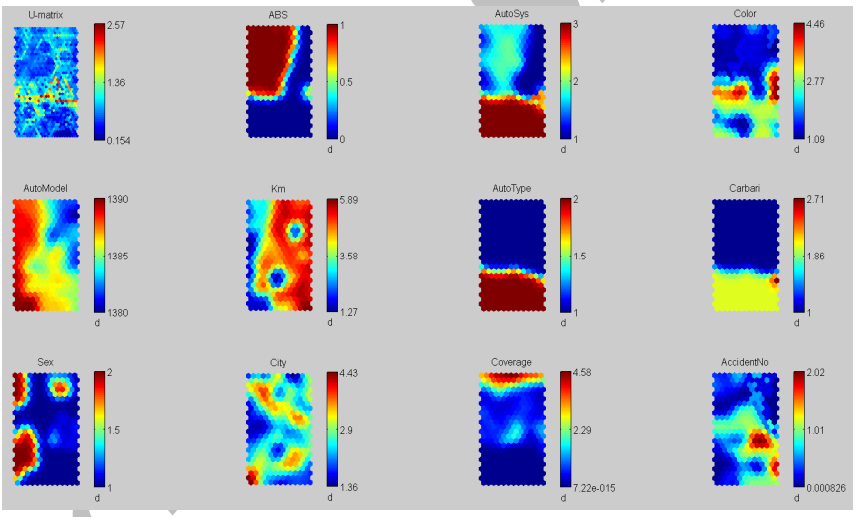
Archive of SID

www.SID.ir

- تعداد خسارت سال قبل: باتوجه به رنگ نقشه در این ناحیه می‌توان گفت که خودروهای با سابقه خسارت صفر و یک در این بخش دیده می‌شود. تعداد بسیار کمی هم خسارت بالای یک دیده می‌شود.

همچنین تحلیل دیگری که می‌توان به کمک نقشهٔ متغیرها انجام داد، یافتن روابط معنی‌دار و بی‌معنی در بین متغیرها از طریق مقایسهٔ نقشهٔ دو به دوی متغیرهاست. شکل ۷ نقشهٔ هم‌زمان کلیهٔ متغیرها و ماتریس U را نشان می‌دهد.

شکل ۷. نمایش هم‌زمان ماتریس U و سایر متغیرها



به‌عنوان نمونه از روی هم قراردادن نقشهٔ متغیرهای جنسیت و متغیر شهر محل زندگی می‌توان نتیجه گرفت که فراوانی بیمه‌گذاران زن در کلان‌شهرها و شهرهای بزرگ و نسبتاً بزرگ بیشتر از شهرهای کوچک است. باتوجه به نقشهٔ متغیر جنسیت و متغیر میزان پوشش اضافی خریداری‌شده، رابطه معنی‌داری بین متغیر میزان پوشش اضافی خریداری‌شده و جنسیت بیمه‌گذار نمی‌توان یافت. در رابطه با متغیر تعداد تصادف در سال قبل باید به این نکته اشاره کرد که آمار تصادفات در بین بیمه‌گذاران مرد در مقایسه با خانم‌ها بیشتر است.

www.SID.ir

۵. نتیجه‌گیری

در ایران نرخ حق‌بیمهٔ بدنهٔ اتومبیل باتوجه به تعرفهٔ بیمه مرکزی ج.ا.ا. تعیین می‌شود و این امر سبب می‌گردد که مشتریانی با ریسک پایین‌تر، خسارت‌های مالی مشتریانی با ریسک بالاتر را جبران کنند و لذا تفاوت آنچنانی بین مشتریان با ریسک بالا و پایین وجود نخواهد داشت.

در این تحقیق سعی گردید فاکتورهای مهم و اثرگذار بر ریسک مشتریان شناسایی گردیده و همچنین با کمک شبکه‌های عصبی خودسازمانده، مشتریان بیمهٔ بدنهٔ اتومبیل در یکی از شرکت‌های بیمهٔ خصوصی براساس ریسکشان خوشه‌بندی شوند. نتایج این تحقیق نشان داد که مشتریان برای شرکت‌های بیمه از جنبه‌های مختلف با یکدیگر متفاوت‌اند؛ به‌عنوان‌مثال ریسک انواع خودرو، رنگ خودرو و جنسیت بیمه‌گذار و بسیاری از عوامل دیگر سبب می‌شود که مشتریان از یکدیگر متمایز باشند. تجزیه‌وتحلیل داده‌ها اهمیت شاخص‌های نهایی انتخاب‌شده را به‌خوبی نشان داد. نکته قابل توجه در مورد شاخص‌های انتخابی این است که کلیهٔ شاخص‌های ریسکی که با مطالعهٔ تحقیقات گذشته و همچنین نظرسنجی از خبرگان تعیین گردید همگی اهمیت زیادی دارند. به‌عنوان نمونه باتوجه به اینکه این تحقیق تنها بر روی سه نوع خودرو تمرکز نمود در بین این سه نوع خودرو، ریسک خودرو وانت از دو خودرو دیگر و ریسک خودرو پراید از خودرو زانتیا بیشتر است. که از بین خودروهای زانتیا و پراید با استناد به نتایج تجزیه‌وتحلیل داده‌ها می‌توان بیان کرد که ریسک خودروهایی که مجهز به سیستم ترمز ABS می‌باشند در مقابل سایر خودروها کم‌تر است. نتایج همچنین نشان داد که ریسک بیمه‌گذاران مرد در برابر بیمه‌گذاران زن در کشور ما بالاتر است.

Archive of SID

www.SID.ir

۶. محدودیت‌های پژوهش و پیشنهاد برای پژوهش‌های آتی

- در این تحقیق باتوجه به اینکه فراوانی سه نوع خودرو پراید، زانتیا و وانت از سایر خودروها در شرکت مورد مطالعه بیشتر بود، تنها بر این سه نوع خودرو تمرکز کردیم. پیشنهاد می‌شود تحقیقات آتی بر روی انواع خودروهایی انجام شود که فراوانی نسبتاً مشابهی با یکدیگر داشته باشند؛ چرا که این امر بر نتایج تحلیل تأثیرگذار است.
- تحلیل بر روی تک‌تک انواع خودروها می‌تواند نتایج دقیق‌تر و بهتری را برای تاثیرگذاری سایر متغیرها در مورد آن خودرو خاص نشان دهد.
- برخی از فاکتورهای مهم و اثرگذار همچون سن بیمه‌گذار، سال اخذ گواهینامه و نوع گواهینامه رانندگی به‌دلیل موجودنبودن اطلاعات آن در پایگاه داده شرکت مورد مطالعه حذف گردید. باتوجه به آنچه تحقیقات در مورد اثرگذاری و اهمیت این فاکتورها نشان می‌دهد، پیشنهاد می‌گردد در زمان صدور بیمه‌نامه برخی از این اطلاعات از مشتریان دریافت گردد تا بتوان با تحلیل آنها به نتایج مهمی دست یافت که در سودآوری شرکت مؤثر واقع گردد.

منابع

۱. الوانی، سیدمهدی، آذر، عادل و دانایی‌فرد، عادل ۱۳۸۳، *روش‌شناسی پژوهش کمّی در مدیریت: رویکردی جامع*، ناشر اشراقی.

۲. جعفری صمیمی، احمد و مرادی، محمدعلی ۱۳۸۷، خصوصی‌سازی و بیمهٔ اتومبیل در ایران، *همایش بین‌المللی صنعت بیمه، چالش‌ها و فرصت‌ها*.

۳. حسین‌زاده، لیلا ۱۳۸۶، *دسته‌بندی مشتریان هدف در صنعت بیمه بااستفاده از داده‌کاوی*، پایان‌نامه کارشناسی‌ارشد، دانشگاه تربیت مدرس.

۴. رستخیز پایدار، ندا ۱۳۸۹، *بخش‌بندی مشتریان بیمهٔ بدنهٔ اتومبیل بااستفاده از داده‌کاوی، مورد مطالعه، بیمهٔ ملت*، پایان‌نامه کارشناسی‌ارشد، دانشگاه علامه طباطبائی.

۵. سرمد، زهره، بازرگان، عباس و حجازی، الهی ۱۳۸۳، *روش‌های تحقیق در علوم رفتاری*،

Archive of SID

www.SID.ir

نشر آگاه تهران، چ ۱۰.

۶. ماجد، وحید ۱۳۸۷، راهکارهای طبقه‌بندی ریسکی بیمه‌گذاران در بازار بیمه: شواهدی از بازار بیمهٔ بدنهٔ اتومبیل در ایران، *همایش بین المللی صنعت بیمه، چالش‌ها و فرصت‌ها*.

7. Bener, A, Ghaffar, A, Azab, A, Sankaran-Kutty, M, Toth, F & Lovasz, G 2006, 'The impact of four-wheel drives on road traffic disability and deaths compared to passenger cars', *Journal of College of Physicians and Surgeons Pakistan*, vol. 16, no. 4, pp. 257-60

8. Cattell, RB 1966, *The scree test for the number of factors*, Multivariate Behavioral Research. 1, pp. 245-76.

9. Chen, YL, Hsu, CL & Chou, SC 2003, 'Constructing a multi-valued and multilabeled decision tree', *Expert Systems with Applications*, vol. 25, pp. 199–209.

10. Chou, JS & Hilty, RD 2003, *Proceedings of IPC annual metting*, Exotic Interconnections and Advanced Process Capability, Minneapolis, MN, September 28- October 2.

11. Dennis, C, Marsland, D & Cockett, T 2001, 'Data mining for shopping centrescustomer knowledge management framework', *Journal of Knowledge Management*, vol.5,pp.368-74.

12. Dickson, PR 1982, 'Person–situation: segmentation's missing link', *Journal of Marketing*, vol.46,no.4.pp.56-64.

13. Ha, SH, Bae, SM & Park, SC 2002, 'Customer's time-variant purchase behavior and corresponding marketing strategies: An online retailer's case', *Computers and Industrial Engineering*, vol. 43, pp. 801–20.

14. Hwang, H., Jung, T & Suh, E 2004, An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry', *Expert Systems with Applications*, vol.26, no. 2, pp. 181–8.

15. Kim, SY, Jung, TS, Suh, EH, & Hwang, HS 2006, Customer segmentation and strategy development based on customer lifetime value: a case study', *Expert Systems with Applications*, vol. 31, no. 1, pp. 101-7.

16. Kohonen, T, *Automatic formation of topological maps of patterns in a self-organizing system*, Oja, E & Simula, O (Eds), Proceedings Of SCIA Scand, Conference on Pattern Recognition, Los Alamitos, CA. IEEE Computer Soc. Press, pp. 182-5.

17. Kohonen, T 2001, *Self-organizing maps*, Springer Series in Information Sciences, 30, Springer, Berlin,Heidelberg New York, Th^{3}ed.

Archive of SID

www.SID.ir

18. Kreidler, M 2008, *Guide to auto insurance*, Washington State Office of the Insurance Commissioner, viwed 2010/4/24 <http://www.insurance.wa.gov>.

19. Lee, JH & Park, SC 2005, 'Intelligent profitable customers segmentation system based on business intelligence tools', Expert Systems with Applications, vol. 29,pp. 145–52.

20. Lee, SC, Suh, YH, Kim, JK & Lee, KJ 2004, 'A cross-national market segmentation of online game industry using SOM', *Expert Systems with Applications*, vol. 27,pp 559–70.

21. Leung, Y, Zhang, J & Xu, Z 2000, 'Clustering by scale-space filtering', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, np.12, pp. 1396–410.

22. Li, Y & Bai, Y 2008, 'Development of crash-severity-index models for the measurement of work zone risk levels', *Accident Analysis and Prevention*, vol. 40, pp. 1724-31.

23. Ngai, EWT Xiu, L Chau, Dck 2009, 'Application of data mining techniques in customer relationship management: a literature review and classification', *Expert Systems with Applications*, vol. 36, pp. 2592-602.

24. Vesanto, J & Alhoniemi, E 2000, 'Clustering of the self-organizing map', *IEEE Transactions on Neural Networks*, vol. 11, no.3, pp. 586-600.

25. Vesanto, J 1997, Data mining techniques based on the Self-Organizing map', Helsinki University of Technology Web Log Classification. *IEEE Intelligent Systems*,vol. 20, pp. 55–9.

26. Woo, JY, Bae, SM & Park, SC 2005, 'Visualization method for customer targeting using customer map', *Expert Systems with Applications*, vol. 28, pp. 763–72.

27. Wu, CH, Kao, SC, Su, YY & Wu, CC 2005, 'Targeting customers via discovery knowledge for the insurance industry', *Expert Systems with Applications*, vol, 29. pp. 291-99.

28. Yu, JX, Ou, Y, Zhang, C & Zhang, S 2005, 'Identifying interesting visitors through web log classification', *IEEE Intelligent Systems*,vol. 20, pp. 55–9.

Archive of SID

www.SID.ir