دانشگاه پیام نور

دانشکده فنی مهندسی

دانشگاه پیام نور مرکز تهران

پایان‌نامه

برای دریافت درجه کارشناسی ارشد

رشته مدیریت فناوری اطلاعات

گروه مهندسی کامپیوتر و فناوری اطلاعات

پیش‌بینی میزان ریسک مشتریان هدف در صنعت بیمه با استفاده از روش‌های داده‌کاوی (مورد کاوی بیمه بدنه شرکت بیمه ایران)

سامرند خالهء

استاد راهنما:

دکتر نسترن حاجی حیدری

استاد مشاور:

دکتر احمد فراهی

اردیبهشت ۹۱


بسم الله الرحمن الرحيم

تقدیم به:

پدر و مادر گرانقدر و مهربانم که

همیشه شمع راه من بوده اند و

خواهر دوست داشتنیم

که حامی همیشگی من است.

تشکر و قدردانی:

قبل از هرچیز پروردگار بزرگ را شاکرم که به من نیرو داد تا در این راه قدم گذارم.

لازم می‌دانم از زحمات استاد عزیز و گرانقدر سرکار خانم دکتر نسترن حاجی حیدری نهایت تشکر و قدردانی را به عمل آورم. بی‌شک رهنمودهای ایشان، مهم‌ترین عامل موفقیت این پایان‌نامه بوده است.

همچنین از حمایت‌ها و مشاوره استاد گرامی جناب آقای دکتر احمد فراهی در تمام مراحل این تحقیق کمال تشکر را دارم.

جا دارد از زحمات استاد ارجمند و دوست عزیز، دکتر مهدی فسنقری که نخستین الهام بخش من در زمینه فعالیت در شاخه داده‌کاوی بودند کمال تشکر را داشته باشم.

چکیده:

پایه گذاری سیستمی که ریسک مشتریان را کنترل می‌کند، یک بخش مهم در مدیریت علمی یک شرکت بیمه تلقی می‌گردد. با توجه به اهمیت ریسک بیمه های خودرو در این تحقیق، تکنیک داده‌کاوی برای تحلیل ریسک مشتریان در یک شرکت بیمه ایرانی مورد استفاده قرار می‌گیرد. هدف آنست که نهایتاً مدل تصمیم‌گیری ارائه گردد تا مشتریان قبل از بیمه نمودن شناسایی شوند. روش تحقیق بر اساس یک فرایند استاندارد داده‌کاوی می‌باشد بدین صورت که داده‌های مشتریان سابق جمع‌آوری و پالایش شده و متغیرهای موثر بر رفتار ریسک مشتریان شناسایی گردیده‌اند. سپس بر پایه اینکه مشتریان ریسکی داشته‌اند یا خیر در دو طبقه خوب و بد قرار داده شده‌اند و برای هر مشتری یک برچسب کلاس در نظر گرفته شده است. بعد از پیش پردازش و تغییر روی رکوردها (مشتریان) و فیلدها (ویژگی‌های مشتریان)، تکنیک C5 بر روی داده‌های نهایی اجرا گردید و الگوهای پنهان به صورت یک درخت در قوانین داده‌کاوی ارائه گردید. داده‌های نهایی همچنین با چند الگوریتم دیگر مانند درخت‌های تصمیم سنتی، شبکه‌های عصبی، رگرسیون لجستیک، شبکه‌های بیزین، ماشین بردار پشتیبان نیز مدل‌سازی گردید. نتایج نشان داده الگوریتم نتایج بهتری را برای طبقه بندی مشتریان نسبت به سایر الگوریتم‌ها داشته است. همچنین مدل ترکیبی طبقه مشتریان را بهتر از سایر تکنیک‌ها پیش‌بینی نموده است.

کلمات کلیدی:

بیمه بدنه اتومبیل، طبقه‌بندی ریسک مشتریان، طبقه بندی، داده‌کاوی، الگوریتم C5

فهرست مطالب

عنوان صفحه

فهرست اشکال

فهرست جداول

فصل اول

مقدمه

۱- مقدمه

۱-۱- مقدمه

همانند سایر بخش‌های اقتصادی، صنعت بیمه نیز شاهد تغییرات بسیاری در عرصه فناوری اطلاعات در طی سال‌های گذشته بوده است. پیشرفت در زمینه های سخت افزار، نرم افزار و شبکه های فناوری اطلاعات فواید بسیاری را از قبیل کاهش هزینه‌ها، کاهش زمان پردازش داده‌ها، افزایش قابلیت‌های سودآوری، همزمان با رو یارویی با چالش‌های جدید در فضایی که هر روز شاهد افزایش میزان رقابت در آن بوده‌ایم به همراه داشته است. ابداعات و نوآوری‌های فنی از قبیل استخراج و نیز نگهداری داده‌ها به طور قابل ملاحظه ای موجبات کاهش هزینه‌های نگهداری، دست یا بی و نیز پردازش داده‌ها را فر اهم آورده است. بسیاری از سؤالات تجاری که سابقاً به علت کمبود داده او امکانات پردازشگری غیرممکن، غیرعملی و یا غیرقابل حل می‌نمودند، در حال حاضر با استفاده از تکنیک داده کاوی قابل بررسی و پاسخ گویی هستند. (حسین زاده، ۱۳۸٦)

۱-۲- تعریف مسئله و سؤالات اصلی تحقیق

پیچیدگی محیطی، شدت رقابت، رواج تکنولوژی‌های نو و پیشرفته، توسعه فناوری اطلاعات و ارتباطات، شیوه های نوین عرضه کالاها و خدمات، مسایل زیست محیطی و سمت گیری سازمان‌ها از دارایی‌های مشهود به نامشهود و... از عوامل عمده ای است که موجب شده است سازمان‌ها و بنگاه‌های اقتصادی در دوران حیات خود با ریسک های بسیار متعدد و خطرات زیاد و حتی پیش بینی نشده مواجه شوند. به همین جهت به منظور کاهش ریسک و جبران زیان‌های ناشی از آن امروزه

در ادبیات علمی انواع مدیریت ریسک نظیر مدیریت ریسک بنگاه، مدیریت ریسک کسب و کار و مدیریت ریسک استراتژیک مطرح شده و هریک جایگاهی خاص دارند. بدیهی است هر سازمان باتوجه به ماهیت کار خود، ریسک های گوناگونی را تجربه می کند و در شرایط متحول امروز، اساساً موفقیت هر بنگاه به تسلط آن بر ریسک‌ها و نوع مدیریتی است که بر انواع ریسک‌ها اعمال می کند. مدیریت ریسک زمانی معنا و مفهوم می یابد که شرایط با احتمال متحمل شدن زیان و عدم اطمینان مواجه شود. این نوع مدیریت شامل حوزه‌های گسترده‌ای است که مسایل مالی، عملیاتی، تجاری، استراتژیک و حوزه وسیع‌تری به نام حوادث خطرآفرین را دربرمی گیرد. درمجموع مدیریت ریسک فرایند سنجش یا ارزیابی ریسک و سپس طرح استراتژی‌هایی برای اداره ریسک است. بدیهی است باتوجه به شرایط پیچیده و رقابتی کسب و کار در عصر امروز، مدیریت ریسک بیش از گذشته اهمیت خود را بازیافته و مدیران برای بقای بنگاههای خود و کاهش زیان ناگزیرند به آن روی آورده و متعهد به اجرای آن باشند. ریسک دلیل وجود بیمه است و بدون ریسک درواقع بیمه مفهوم خود را از دست می دهد. کار بیمه گری با ریسک و ریسک‌پذیری و کاهش ریسک و محاسبه ریسک سروکار دارد..

یک از چالش‌های اصلی شرکت‌های بیمه کنترل ریسک برای مشتریان به خصوص مشتریان بیمه بدنه می‌باشد و شناسایی و طبقه‌بندی مشتریان کم ریسک و پر ریسک می‌تواند در مدیریت ریسک مشتریان و سیاست‌های مرتبط مانند روی‌گردانی و کاهش ریسک و تعدیل نرخ حق بیمه موثر باشد. استفاده از داده کاوی در این زمینه می‌تواند بسیار راهگشا باشد و این تحقیق به بررسی موضوع طبقه بندی مشتریان به لحاظ ریسک خسارت بیمه بدنه خودرو در یک شرکت بیمه کشور می‌پردازد.

به دلیل نفوذ زیاد بیمه‌های اتومبیل و نقشی که این نوع بیمه در بین دیگر رشته‌های بیمه در ایران دارد[1]، بیمه بدنه اتومبیل برای این مطاله انتخاب شد. ولی متأسفانه ابزارهای جدیدی مانند داده‌کاوی در صنعت بیمه ایران تاکنون جایگاه اصلی خود را پیدا نکرده است. اما به علت تغییر محیط به لحاظ قانونی (تبدیل نظام تعرفه‌ای حق بیمه به نظام نرخ‌گذاری آزاد) استفاده از فناوری‌های جدید برای رقابت در یک بازار کاملاً رقابتی از سوی شرکت‌های بیمه به نظر اجتناب ناپذیر می‌باشد. در حال حاضر از ابتدا تمامی بیمه‌گذاران حق بیمه واحدی پرداخت می‌کنند که این امر باعث می‌شود بیمه-گذاران با ریسک پایین‌تر نسبت به بیمه‌گذاران با ریسک بالاتر حق بیمه بیشتری پرداخت نمایند که به نظر منصفانه نمی‌رسد. با دسترسی به ابزارهای تحلیل و طبقه‌بندی ریسک می‌توان این مشکل را برطرف ساخت. با استفاده از داده‌کاوی و تجزیه تحلیل داده‌ها، می‌توان این بیمه‌گذاران را به لحاظ ریسک طبقه‌بندی نمود و با استفاده از این طبقه‌بندی می‌توان طبقه بیمه‌گذاران آتی را پیش‌بینی کرد و در میزان حق دریافتی از بیمه‌گذاران مختلف تعدیل ایجاد نمود و با ایجاد یک سیستم نرخ‌گذاری مبتنی بر ریسک بیمه‌گذاران، میزان رضایت بیمه‌گذاران را افزایش داده و از طرف دیگر بر سودآوری شرکت‌های بیمه تأثیر مثبتی گذاشت.

در این راستا سؤالات زیر مطرح اند:

سؤال اصلی:

[1] طبق آمار سال ۸۸ بیمه مرکزی ج.ا. ایران سهم بیمه بدنه اتومبیل از نظر میزان حق بیمه ۱۵،۱۱ درصد است که در بین سایر رشته‌های بیمه‌ای جایگاه سوم را داراست.

۱) چگونه می‌توان در یک شرکت بیمه، مشتریان را به لحاظ ریسک خسارتی با استفاده از الگوریتم‌های داده‌کاوی طبقه بندی نمود؟

سؤالات فرعی:

۱) آیا الگوریتم C5، طبقه بندی صحیح‌تری را نسبت به سایر الگوریتم‌های طبقه بندی داده کاوی ارائه خواهد کرد؟.

۲) آیا شاخص «نوع خودرو» مهم‌ترین شاخص در طبقه بندی مشتریان بیمه خودرو می‌باشد؟

۳) آیا مدل ترکیبی الگوریتم‌های داده کاوی نسبت به سایر الگوریتم‌ها صحت بالاتری دارد؟

۱-۳- اهداف تحقیق

۱) تعیین مهم‌ترین متغیر های دخیل در پیش بینی رفتار بیمه گذاران بیمه بدنه اتومبیل.

۲) جستجوی الگوهای موجود بین داده های بیمه گذاران سابق و ارائه چارچوب مناسبی برای طبقه بندی بیمه گذاران بیمه بدنه اتومبیل با استفاده از داده‌کاوی.

۳) مقایسه نتایج پیش بینی الگوریتم‌های مختلف طبقه بندی در داده‌کاوی مشتریان بیمه بدنه خودرو.

٤) افزایش سودآوری صنعت بیمه در زمینه طبقه بندی بیمه گذاران برای تعیین حق بیمه.

٥) کاهش خطاهای انسانی و جلوگیری از فرایند قضاوتی کارشناسان خبره به وسیله بکارگیری الگوهای کشف شده این تحقیق در یک سیستم هوشمند (یادگیری ماشین).

٦) - تعیین چارچوب ریسک برای طبقات مختلف بیمه‌گذاران و در نتیجه پرداخت میزان عادلانه حق بیمه بر مبنای میزان ریسک هر بیمه‌گذار.

۱-٤- روش تحقیق

روش انجام تحقیق از طریق مطالعه و بررسی کتب، مقالات، پایان‌نامه‌های انجام شده داخلی و خارجی، پروژه های تحقیقاتی صورت گرفته و اینترنت می‌باشد.

۱-۵- مراحل انجام تحقیق

۱- جمع آوری داده از پایگاه‌داده بیمه‌گذاران فعلی بانک اطلاعاتی بیمه بدنه خودرو و تقسیم آنها به دو دسته داده های آزمایشی و داده های آموزشی؛

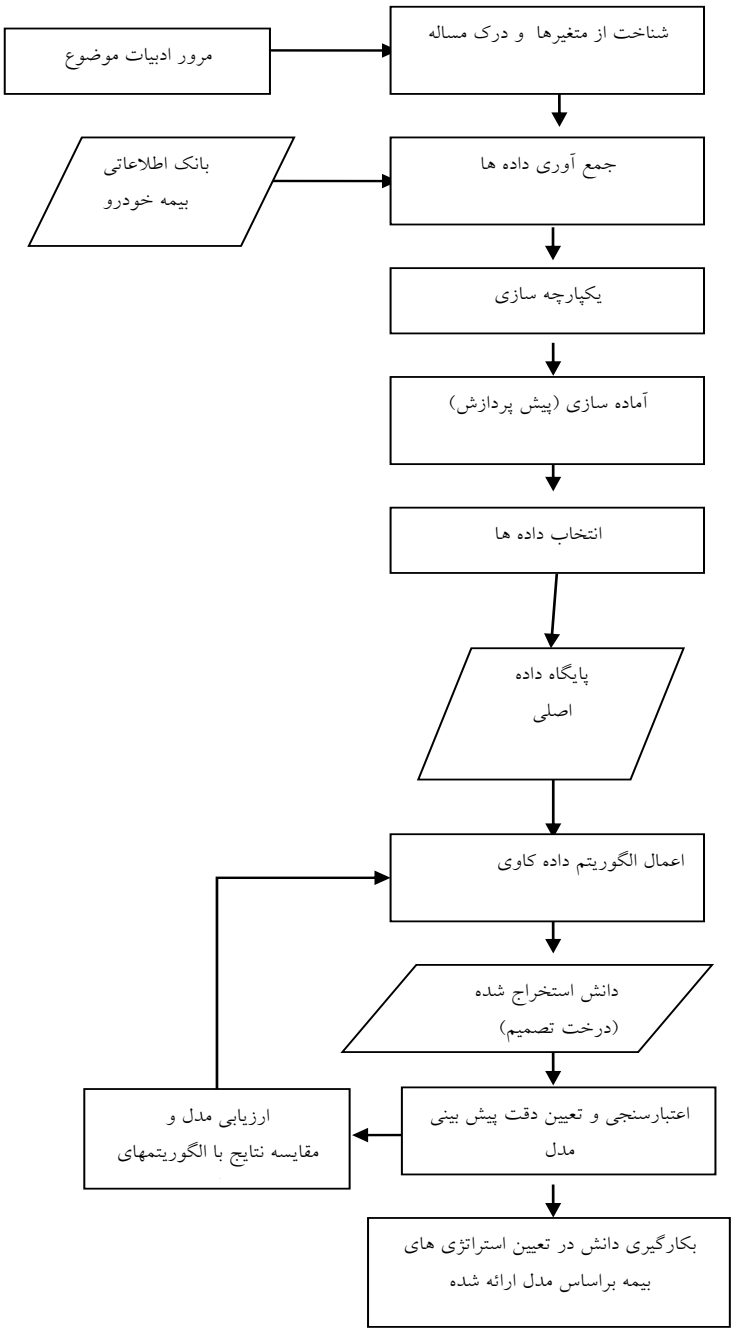
۲- تعیین شاخص‌هایی برای تعریف طبقات ریسک بیمه گذاران؛

۳- استخراج الگوها با استفاده از داده‌های آزمایشی با تکنیک درخت تصمیم و مقایسه نتایج با استفاده از روش‌های دیگر؛

٤- ارائه الگوی کشف شده از طبقه بندی بیمه‌گذاران؛

۵- اعتبارسنجی مدل با استفاده از مجموعه داده های آزمایشی.

در خصوص مراحل انجام تحقیق و مدل اجرایی در فصل سوم به تفصیل اشاره شده است. در این فصل مدل اجرایی تحقیق را می‌توان در قالب شکل ۱-۱ نمایش داد.



شکل ۱-۱ مدل اجرایی تحقیق