

Customer Credit Clustering for Presenting Appropriate Facilities

Amir Afsar¹, Rahmat Houshdar Mahjoub^{2*}, Behroz Minaie³

1. Assistant Professor, Department of Industrial Management, Qom University, Qom, Iran
2. M.A. in Information Technology, Qom University, Qom, Iran
3. Assistant Professor, Department of Commuter Engineering, Science & Technology University, Tehran, Iran

Receive: 23/12/2012

Accept: 13/7/2013

Credit institutions to provide variety of granted facilities to their customers, need to conduct comprehensive studies in order to know qualitative and quantitative aspects of their applicants. By this way, they can accomplish a complete evaluation of repay ability measure and calculate the refund facilities probability and finance services (generally called as validation). The purpose of this study was ranking the customer groups and specifying the best part of them until brokerage firms can do their credit allocation process mechanically. Here, after the preprocessing of the data, we converted them into RFM model. Then SOM neural network as one of the clustering algorithms changed the customers into 10 clusters. Using the proposed model, the clusters were ranked. The top clusters were identified and facilities grant operations were done to the members of these clusters. Finally, three clusters 5, 1 and 7 were defines as top clusters, which are the target customers. Coefficient facilities granted to the top three clusters are 0.271, 0.173 and 0.556, respectively.

Keywords: Credit risk management, Customer validation, RFM, SOM neural network

* Corresponding Author's E-mail: houshdar.mahjoob@gmail.com

خوشه‌بندی اعتباری مشتریان برای ارائه تسهیلات متناسب

امیرافسر^۱، رحمت هوشدار مهجوب^{۲*}، بهروز مینایی^۳

۱- استادیار گروه مدیریت، دانشگاه قم، قم، ایران

۲- دانشجوی کارشناسی ارشد فناوری اطلاعات گرایش تجارت الکترونیک، دانشگاه قم، قم، ایران

۳- استادیار دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت، تهران، ایران

دریافت: ۹۱/۱۰/۳

پذیرش: ۹۲/۴/۲۲

چکیده

مؤسسات اعتباری برای در اختیار قرار دادن انواع تسهیلات اعطایی به مشتریان خود، نیاز به انجام بررسی‌های کاملی به منظور شناخت متقاضیان از ابعاد کیفی و کمی دارند تا از این راه، ارزیابی کاملی از سنجش توان بازپرداخت و محاسبه احتمال عدم بازپرداخت تسهیلات و خدمات تأمین مالی از سوی آنان به عمل آید، این بررسی‌ها را به طور عام اعتبارسنجی گویند. هدف از انجام این تحقیق رتبه‌بندی گروه-های مشتریان و تعیین بخش‌های برتر از آنها می‌باشد تا با استفاده از آن شرکت کارگزاری بتواند عملیات تخصیص اعتبار را به نحوی مکانیزه انجام دهد. در اینجا پس از پیش پردازش اولیه از داده‌ها، آنها به شکل مدل RFM^۱ پردازش می‌شوند. سپس با استفاده از شبکه عصبی SOM^۲ به عنوان یکی از الگوریتم‌های خوشه‌بندی، مشتریان به ۱۰ خوشه تبدیل خواهند شد. در ادامه با استفاده از مدل پیشنهادی، خوشه‌ها رتبه‌بندی می‌شوند. خوشه‌های برتر شناسایی و عملیات اعطای تسهیلات برای اعضای این خوشه‌ها انجام می‌شود. در نهایت سه خوشه ۵، ۱ و ۷ به عنوان خوشه‌های برتر تعیین شدند که به عنوان مشتریان هدف می‌باشند. ضریب تسهیلات اعطایی به این سه خوشه برتر به ترتیب ۰/۲۷۱، ۰/۱۷۳ و ۰/۵۵۶ می‌باشد.

کلیدواژه‌ها: مدیریت ریسک اعتباری، اعتبارسنجی مشتریان، RFM، شبکه عصبی SOM



۱- مقدمه

در بازار رقابتی امروز در بسیاری از صنایع، شرکتها در جهت تحقق کامل استراتژی های کسب و کار مشتری‌گرای خود هستند و برای حفظ مزیت رقابتی و سطح سود خود و درآمد-زایی بیشتر، تکیه بر حفظ مشتریان کنونی خود دارند. ارائه انواع روشهای تبلیغاتی و ارائه تسهیلات ویژه‌ای به مشتریان خاص بخصوص درصنایعی که از نظر حجم کاری وسعت بیشتری دارند بیشتر نمود پیدا می‌کند.

مدیریت ارتباط با مشتری، بر مبادله ارزش بین مشتری و شرکتها بنا شده و بر ارزش ایجاد شده در این ارتباط تاکید می‌کند بنابراین، تلاش شرکتها برای توسعه ارتباط بلندمدت با مشتریان بر مبنای ایجاد ارزش برای هر دو طرف از اهداف اصلی مدیریت ارتباط با مشتری می‌باشد. تقسیم بازار از جمله مباحثی است که در متون جدید بازاریابی مورد تأکید قرار گرفته است. به‌کارگیری روش‌های یکسان در بازاریابی، بدون توجه به تفاوت‌های موجود در بازار و مشتریان، رضایت همه خریداران را نمی‌توان جلب کرد. [۱، ص ۶۰-۶۲]

برای اینکه CRM بتواند موفق باشد مستلزم آن است که سازمان ارتباطی انعطاف‌پذیر با مشتریان خود داشته باشد. موسساتی که به درستی به ارزشیابی مشتریان خود می‌پردازند، خدمات سفارش‌سازی شده‌ای را برای مشتریان متنوع ارائه می‌دهند این کار باعث افزایش سود-آوری آنها می‌شود چون از این راه می‌توان مشتریان پر ارزش را حفظ کرد. [۲، صص ۱-۲].

شرکت‌های زیادی به دنبال اندازه‌گیری ارزش مشتریان و استفاده از این اطلاعات در مدیریت حفظ مشتریان و افزایش پتانسیل سودآوری آنها هستند [۳]. در سال‌های اخیر به دلیل رقابت شدید در صنعت اعتباری، اعتبارسنجی به یکی از موضوعات حیاتی و چالشی تبدیل شده است. اغلب مدیران اعتبارسنجی براساس تجربیات خود، به ارزیابی اعتبار متقاضیان خود می‌پردازند که این باعث افزایش ریسک اعتباری می‌شود [۴، صص ۲-۴].

بدون شک ارزیابی مشتریان برای اعطای اعتبار موضوع مهمی در حوزه خرید اعتباری سهام به حساب می‌آید. وقتی مشتریان هدف به درستی شناسایی نمی‌شود هم ممکن است باعث ریزش گروه پر سوده شود هم باعث می‌شود که بعضی از مشتریان که به اشتباه به آنها تسهیلات بیشتر داده شده نتوانند بدهی های خود را بازپرداخت کنند، در نتیجه از طرفی سود دهی موسسه مالی را کاهش می‌دهد و از طرف دیگر امکان شکست اقتصادی را برای سازمان

مالی وام دهنده افزایش می دهد. بنابراین بهبود در امر تصمیم گیری در مورد اعطای تسهیلات به مشتریان شرکت کارگزاری و درجه بندی اعتباری یکی از مسایل مرتبط با مدیریت ریسک اعتباری شرکت های کارگزاری می باشد. به عبارتی مساله رتبه بندی اعتباری و تخصیص اعتباری به فراخور اعتبار، گریبانگیر بسیاری از مراکز تصمیم گیری است. لذا استفاده از مدل های مناسب جهت تخصیص بهینه اعتبار و توزیع اعتبار خرید میان مشتریانی که از اعتبار بالایی برخوردارند اهمیت بسزایی دارد.

روش های سنتی تصمیم گیری در مورد اعطای اعتبار به متقاضیان ، همانند آنچه که اکنون در کشور ما انجام می گیرد که برپایه قضاوت شخصی استوار است، دیگر جوابگو نخواهد بود. شرکت ها می توانند با مدل های داده کاوی مشتریان همگن با ارزش یکسان را در خوشه های کوچکتر مشابه قرار دهند تا به تقسیم بندی بازار بپردازند [۱][۵] و از این راه تمرکز و تلاش خود را به شناسایی فرصت ها و ارائه خدمات به شیوه سفارشی مبذول دارد [۲، صص ۱-۳]. مزایای اعتبارسنجی شامل کاهش هزینه های تحلیل اعتبار، امکان اتخاذ سریعتر تصمیم های اعتباری، نظارت دقیق تر بر حساب های موجود و ... می باشد [۶، ص ۵].

این تحقیق به این صورت سازماندهی شده است : در بخش دوم به بررسی ادبیات موضوع تحقیق یعنی اعتبارسنجی پرداخته شده و کارهای که در گذشته در این زمینه انجام شده است بیان می شود. بخش سوم اختصاص به روش تحقیق دارد. بخش چهارم نحوه آماده سازی داده ها و انجام روند اعتبارسنجی مشتریان شرح داده شده است و در انتها در قسمت پنجم نتیجه گیری گفته می شود .

۲- پیشینه تحقیق

اعتبارسنجی یعنی سنجیدن ظرفیت افراد در استفاده از منابع از جمله تسهیلات اعطایی می باشد. شیوه های اعتبارسنجی مشتریان، اساس رتبه بندی یا رتبه سنجی مشتریان است و هدف اصلی از آن، فراهم آوردن اطلاعات لازم برای اعطای تسهیلات به مشتریان در بانکها و مؤسسات تأمین مالی، بازار بورس و برای سرمایه گذاران در بازارهای سرمایه، می باشد. با شیوه های اعتبارسنجی می توان ریسک سرمایه گذاری برای سرمایه گذاران را پیش بینی و برآورد کرد. مؤسسات اعتباری در گذشته به علت محدودیت در منابع و دسترسی به اطلاعات مشتریان



از روش‌های ساده کیفی برای سنجش اعتبار آن‌ها استفاده می‌کردند. روش‌های کمی بعد از مدت زمان نسبتاً طولانی وارد عرصه تصمیم‌گیری شدند [۷].

تحقیقاتی در این زمینه انجام شده که در این میان می‌توان به مطالعه «فیشر» (۱۹۳۶) به عنوان اولین سیستم ارزیابی تقاضای اعتبار و مطالعه «دوراند» (۱۹۴۱) - که با استفاده از «تحلیل ممیزی» و با استفاده از نتایج فیشر صورت گرفت - به عنوان اولین سیستم‌های امتیازدهی اعتباری در زمان حاضر اشاره کرد.

موسسات اعتباری باید با توجه به پیچیدگی فعالیت‌ها و محیط اقتصادی، مدل‌هایی مناسب جهت ارزیابی امتیازدهی اعتباری مشتریان انتخاب کنند. در خصوص مشتریان بزرگ معیار 5C برای تصمیم‌گیری اعتباری استفاده شده، این معیار با کمی تغییر می‌تواند برای تمام مشتریان بکار رود، البته معیارهای دیگری نیز وجود دارد. معیارهای مدل 5C شامل ویژگی‌های فردی^۲، ظرفیت^۳، سرمایه^۴، وثیقه^۵ و شرایط اقتصادی^۶ می‌باشند [۸].

موسسات اعتباری تلاش می‌کنند تا از روش‌ها و ابزار مختلف برای تصمیم‌گیری خودکار بهره‌گیرند تا فرایند ارزیابی مشتریان را بهبود بخشند و مدیران را در پیش‌بینی و تصمیم‌گیری یاری دهند. در گذشته روش‌های کلاسیک مختلفی نظیر تجزیه و تحلیل ممیزی خطی، رگرسیون خطی، درخت باینری، روش حداقل برای اعتبارسنجی استفاده شده است [۴، صص ۲-۵] [۹]. علاوه بر این چند روش برجسته در هوش مصنوعی همچون سیستم‌های خبره، سیستم فازی، شبکه‌های عصبی و الگوریتم ژنتیک نیز برای اعتبارسنجی مورد استفاده قرار گرفته است، در این میان شبکه عصبی بهترین توان را داشته است [۴، صص ۲]. اما این مورد برای همه مواقع ممکن است عمومیت نداشته باشد.

روش‌های داده‌کاوی می‌تواند به شرکت‌ها برای کسب‌وکار مهم خود کمک کند تا به سرعت مشتریان خود را از داده‌های پیچیده بشناسند و تفکیک کنند و برای درک و تعامل با آن‌ها با استفاده از تاکتیک‌های مناسب به افزایش ارزش مشتری و بهبود مزیت رقابتی خود نسبت به شرکت‌های دیگر استفاده کنند [۲، صص ۷-۲].

این روشن است که افراد مختلف نیازهای متفاوتی دارند. برای پاسخگویی به نیازهای مختلف، آن‌ها را به بخش‌های کوچکتر تقسیم می‌کنیم تا برنامه‌های متناسب آن‌ها طرح ریزی

شود [۱۰]. شناخت درست از مشتری اجازه می‌دهد تا شرکت‌ها به ارائه تعرفه خاص مطابق با ویژگی‌های بیرونی و ذاتی از مشتریان با کشف مشتریان غیر معمولی و یا خاص و یا تفکیک شده بپردازند [۱۱].

وسعت مطالعات خارج از کشور در زمینه اعتبارسنجی مبتنی بر روش‌های پارامتریک و ناپارامتریک تقریباً زیاد بوده و شامل موارد زیر است: رگرسیون لجستیک، مدل پروبیت و لوجیت، تحلیل مولفه‌های اساسی و برخی روش‌های انتخاب ویژگی‌ها، تحلیل تمایزی، روش Naive Bayes، روش SOM، روش MARS، مدل‌های مختلف درخت تصمیم‌گیری، شبکه‌های عصبی، ماشین بردار پشتیبان، سیستم‌های خبره، منطق فازی، مدل‌های ترکیبی طبقه‌بندی، الگوریتم پس‌انتشار شبکه عصبی، Boosting, Bagging، نزدیک‌ترین همسایه و الگوریتم ژنتیک و ترکیب برخی موارد فوق با هم [۶، صص ۲-۵].

فرنس کیس مدل‌های رتبه‌بندی اعتباری را در دو گروه ارزیابی پارامتری و درخت دسته‌بندی، تقسیم می‌کند. ایده جداسازی گروه‌ها در یک جمعیت به وسیله فیشر در سال ۱۹۳۶ مطرح شد. التمن در سال ۱۹۳۸ اولین سیستم ارزیابی تقاضانامه‌های اعتباری را با بکارگیری ۵ معیار توسعه داد. دوران در سال ۱۹۴۱ یک پروژه تحقیقاتی برای اداره ملی تحقیقات اقتصادی آمریکا انجام داد. ایشان بنیانگذار سیستم‌های رتبه‌بندی اعتبار به روش کنونی می‌باشند. در سال ۱۹۶۳ مایرزو هنرجی آنالیز ممیزی چند متغیره برای رتبه‌بندی اعتباری مشتریان ارائه نمودند. مور و کلن در سال ۱۹۶۸ برای رتبه‌بندی شرکت‌ها با استفاده از روش آنالیز ممیزی چند متغیره ارائه کردند. در سال ۱۹۷۷ التمن - هالدمن و نارایانان نسل دوم این مدل را معرفی کردند. در سال ۱۹۸۰ رگرسیون لجستیک و برنامه ریزی خطی برای ارزیابی اعتبار مشتریان بکار گرفته شد.

دسای و همکاران در سال ۱۹۹۸ پژوهشی انجام دادند که طی آن با استفاده از روش‌های نظیر تجزیه و تحلیل ممیزی خطی و رگرسیون لجستیک اقدام به طبقه‌بندی مشتریان وام‌های بین‌المللی در سه کشور آمریکا، آلمان و استرالیا به سه طبقه خوب، ضعیف و بد نمودند. محققان البته در پایان پژوهش به این نتیجه رسیدند که طبقه‌بندی به دو دسته خوب و بد ارجحیت دارد و در ادامه با استفاده از متغیرهای تعیین شده و مجموع داده‌های جمع‌آوری شده از شبکه‌های عصبی مصنوعی به منظور طبقه‌بندی مشتریان خوب و بد بهره بردند. محققان در



این پژوهش با استفاده از اطلاعات مربوط به اعتبارات اتخاذ شده در سه کشور ذکر شده، حد فاصل سال‌های ۱۹۸۸ تا ۱۹۹۱ به فهرست های پیش‌بینی کننده اعتبار رسیدند که این فهرست عبارت است از تعداد کارت های اعتباری اصلی، مالکیت منزل، حقوق و سایر درآمدها، مدت سابقه در شغل فعلی، پرداخت ماهانه متناسب با درآمد، بد حسابی مشتری در ۱۲ ماه گذشته، تعداد حساب‌های فعال و موارد دیگر.

پژوهشی توسط «وست» در سال ۲۰۰۰ انجام شد که در آن با استفاده از شبکه عصبی و روش‌های تجزیه و تحلیل عددی، مشتریان به دو دسته خوب و بد تقسیم شدند. در این پژوهش از معیارهای مختلفی چون طول عمر حساب، تاریخ اعتبار، طبقه بندی شغلی، دارایی های شخصی و ... برای توسعه سیستم اعتبارسنجی استفاده شد.

یکی از روش‌های مطرح در تحلیل ارزشیابی مشتری، الگوی RFM است که توسط هیوز در سال ۱۹۹۴ ارائه شد. در این الگو، تفاوت مشتریان با استفاده از سه متغیر تأخر (تازگی آخرین خرید)، تکرار خرید و ارزش پولی خرید اندازه‌گیری می‌شود [۱۲]. بررسی‌ها نشان می‌دهد که هر چه R و F بیشتر باشد، احتمال آنکه تراکنش جدیدی با مشتری صورت بگیرد، بیشتر خواهد بود و هرچه M بیشتر باشد احتمال بازگشت مشتری برای خرید بیشتر است [۱۳]. مطالعات نوئل^۱ (۱۹۹۷) نشان داده است که متغیر های RFM برای رتبه‌بندی مشتریان بسیار کارا هستند. یه و همکاران برای انتخاب روش بازاریابی مستقیم از متغیرهای RFM استفاده کردند که با اضافه کردن دو متغیر زمان اولین خرید و احتمال ریزش، الگوی RFM را بسط می‌دهد [۱۴، صص ۳-۸]. این الگو کاربردهای گوناگون دیگری نیز داشته است. جانکرا از این الگو برای بخش‌بندی مشتریان به منظور تعیین سیاست‌های بهینه بازاریابی استفاده نمود [۱۵]. در مطالعه‌ای دیگری از این مدل جهت رتبه‌بندی مشتریان استفاده شده است [۱۶، صص ۴-۶]. برخی نیز از این مدل جهت محاسبه ارزش دوره عمر مشتری استفاده کرده‌اند [۱۷].

هسیه (۲۰۰۴) در تحقیق خود یکپارچه‌سازی داده‌کاوی و مدل امتیازدهی رفتاری را برای مدیریت مشتریان کارت اعتباری بانک پیشنهاد داد. از شبکه عصبی SOM برای شناسایی گروه‌های مشتریان براساس رفتار بازپرداخت و تاخر، فرکانس و رتبه بندی گزارش های پولی استفاده کرد. مشتریان بانک را به سه گروه عمده سودآور طبقه‌بندی کردند. سپس ویژگی‌های

گروهی از مشتریان با استفاده از قوانین انجمنی اپریوری^۴ تعیین شد. این مطالعه نشان می‌دهد که شناسایی از راه یک مدل امتیازدهی رفتاری ویژگی‌های مفید مشتری را بیان می‌کند و تسهیل توسعه استراتژی بازاریابی را به همراه دارد. [۱۶]

در سال‌های اخیر استفاده از ترکیبی از چند روش برای عملیات اعتبارسنجی استفاده شده از جمله این کارها در مقاله‌ای لویز و همکارانش (۲۰۱۱) الگوریتمی را برای رتبه‌بندی اعتباری مشترکان برق برای دادن تعرفه‌های خاص و مختلف ارائه کردند. آنان بیان نمودند که این الگوریتم مشکلات الگوریتم‌های رایج بکار رفته را حل کرده است. الگوریتم بکار رفته Hopfield-K-Means است. این رویکرد اتفاقی بودن راه حل‌های اولیه شده به وسیله الگوریتم K-Means را حذف می‌کند و بر اساس آن به مطلوب جهانی نزدیک می‌شود. در این الگوریتم از ویژگی‌های H-ANN^{۱۰} استفاده شده است. الگوریتم از لحاظ شاخص اعتبار با الگوریتم‌های دیگر از جمله الگوریتم‌های سلسله مراتبی (DM and DW)، تغییرات به دنبال لیدر (F)، شبکه‌های عصبی مکرر هاپفیلد (H)، K-means (K)، SOM-K-Means، همواره با استفاده از توصیف داده‌ها از هر الگوریتم و هر شاخص (CH^{۱۱}, DBI^{۱۲} and MIA^{۱۳}) مقایسه شده است. در تجزیه و تحلیل شاخص‌ها، نتایج به دست آمده توسط H-ANN-K بهتر از آن‌هایی است که به وسیله الگوریتم پایه (H و K) و KSOM به دست آمده و شبیه به آن‌هایی است که توسط DM، DW و الگوریتم‌های F به دست آمده هستند [۱۸].

۳- روش تحقیق

۳-۱- مدل RFM

مدل RFM متداول‌ترین روش تقسیم‌بندی مشتریان است. این روش شامل سه متغیر است که عبارتند از تاخر، فرکانس، پول. مدل RFM یک مدل مبتنی بر رفتار است که برای آنالیز رفتار یک مشتری و سپس پیش‌بینی کردن او بر اساس رفتار در بانک اطلاعات استفاده می‌شود [۱۴، صص ۲-۳]. در بین متغیرهای RFM شاخص تأخر اغلب مهم‌ترین متغیر می‌باشد. با این حال براساس مطالعات انجام شده گذشته متغیرهای RFM در واقع Firm-Specific هستند و بر اساس طبیعت محصولات شرکت اهمیت آنها فرق می‌کند [۱۹]. در این الگو R، فاصله زمانی



آخرین خرید مشتری تا زمان حال، F ، تعداد خریده‌ها در یک دوره زمانی مشخص و M ، مبلغ اسمی خریداری شده در دوره مورد نظر تعریف می‌شود [۲۰].

این تحقیق از مدل RFM برای تبدیل داده‌های اولیه به فرم دلخواه برای استفاده در الگوریتم خوشه‌بندی استفاده کرده است در نتیجه تغییراتی در نحوه محاسبه آن‌ها انجام شده است. شاخص تاخر: تعداد ماه‌های که مشتری مقدار منفی در فیلد مانده، ندارد تقسیم بر تعداد کل ماه‌ها؛ شاخص تناوب: تعداد ماه‌های که مشتری تراکنش بدهکاری و بستانکاری دارد؛ شاخص پولی: مجموع گردش مشتری را در طول این دوره زمانی در نظر گرفتیم که در ۱۲ بازه برچسب گذاری شد.

۳-۲- شبکه عصبی SOM

برای امتیازی اعتباری و یا تجزیه و تحلیل امتیازی رفتاری، بسیاری از مطالعات ارائه شده است که با شبکه‌های عصبی انجام می‌شود که به میزان قابل توجهی بهتر از روش‌های آماری از قبیل تحلیل تفکیک‌کننده خطی (LDA)، تحلیل تفکیک کننده چندگانه (MDA)، تجزیه و تحلیل رگرسیون لجستیک (LRA) و غیره می‌باشد [۱۶، ص ۳].

الگوریتم SOM خوشه‌بندی از نوع شبکه عصبی است که در سال ۱۹۸۱ توسط پژوهشگر فنلاندی کوهانن اختراع شد. این الگوریتم به طور معمول متشکل از دولایه نرون‌های ورودی و خروجی است [۲۱]. به طور کلی شبکه‌های عصبی از لایه‌های نرونی تشکیل شده است. این نوع نرون‌ها از طریق ورودی‌های خود با جهان واقعی در ارتباطند و گروه دیگری از نرون‌ها نیز از طریق خروجی‌های خود، جهان خارج را می‌سازند [۲۲].

برای تعیین بهترین تعداد خوشه‌ها از روشی به نام SSE^۴ استفاده می‌شود. در این روش نخست مراکز خوشه در نظر گرفته می‌شود و سپس فاصله نقطه مورد نظر از مراکز خوشه محاسبه می‌گردد. خوشه‌ای که SSE پایین‌تر دارد نشان دهنده بهترین خوشه‌بندی (تعداد خوشه‌ها) است برای این امر از فرمول زیر استفاده می‌شود:

$$SSE = \sum_{i=1}^k \sum_{P \in C_i} d(P, m_i)^2$$

۳-۳- روش امتیازدهی

ارزش هر مشتری را می‌توان بر اساس تازگی (تاخر)، تکرار و ارزش پولی به صورت زیر مشخص کرد: [۲۳، ص ۱۶۴]

$$V(C_i) = W^R \times R(C_i) + W^F \times F(C_i) + W^M \times M(C_i)$$

که $R(C_i)$ ، $F(C_i)$ و $M(C_i)$ به ترتیب امتیازات مشتری C_i با توجه به معیارهای R ، F و M است. w^R ، w^F و w^M اهمیت وزنی برای معیارهای R ، F و M را به ترتیب نشان می‌دهد. به علاوه داریم:

$$W^R + W^F + W^M = 1$$

سود آوری خوشه O^n با محاسبه میانگین ارزش همه مشتری‌های خوشه n حاصل می‌شود. از این رو می‌توان آن را از طریق معادله زیر تعریف کرد [۲۳، ص ۱۶۴]:

$$V(O^n) = W^R \times R(O^n) + W^F \times F(O^n) + W^M \times M(O^n)$$

$$R(O^n) = \frac{\sum_{C_i \in O^n} R(C_i)}{\|O^n\|} \quad F(O^n) = \frac{\sum_{C_i \in O^n} F(C_i)}{\|O^n\|} \quad M(O^n) = \frac{\sum_{C_i \in O^n} M(C_i)}{\|O^n\|}$$

که $R(O^n)$ ، $F(O^n)$ و $M(O^n)$ امتیاز n امین خوشه با توجه به معیارهای R ، F و M است.

۳-۴- مطالعه موردی

سازمان بورس به عنوان یکی از نهادهای مالی پیوند تنگاتنگی با مشتریان دارد. سازمان بورس دارای بخش‌های مختلفی می‌باشد، یکی از این بخش‌ها شرکت‌های کارگزاری هستند. وظیفه این شرکت‌ها جذب سرمایه مشتریان و انجام امور معاملاتی آنها می‌باشد. شرکت کارگزاری سرمایه گذاری ملی ایران به عنوان یکی از کارگزاری‌های شناخته شده در بازار بورس فعالیت می‌کند. همچنین این شرکت دارای نمایندگی در ۹ استان و ۳ تالار در تهران و یک تالار الکترونیکی برای انجام معاملات می‌باشد. این شرکت با داشتن بیش از ۳۰۰۰۰ مشتری فعال در سطح کشور یکی از بزرگترین جامعه‌های آماری را بین شرکت‌های کارگزاری دارا می‌باشد. در این راستا شرکت با تکیه بر نیروهای متعهد و متخصص خود و با اخذ مجوزهای لازم از سازمان بورس و اوراق بهادار، خدماتی را به مشتریان خود ارائه می‌نماید.



۴- مدلسازی و تحلیل داده‌ها

روند اجرای عملیات در این تحقیق در شکل ۱ نشان داده شده است در ادامه به شرح این روند پرداخته می‌شود.

۴-۱- آماده‌سازی و پیش پردازش داده‌ها

در این پژوهش از داده‌های تراکنشی مشتریان مربوط به تالار اصلی - که یکی از تالارها واقع در شهر تهران می‌باشد - استفاده شد. در مجموع اطلاعات مربوط به ۹۵۸۶ مشتری در فاصله زمانی دی ماه ۸۹ تا دی ماه ۹۰ جمع‌آوری شد. با توجه به محدودیت‌های که در تحویل داده‌های دموگرافیک وجود داشت، تنها داده‌های تراکنشی در اختیار قرار گرفت که در کل شامل ۹ فیلد می‌باشد.

مرحله آماده‌سازی داده‌های یکی از مهمترین و پیچیده‌ترین مراحل در داده‌کاوی می‌باشد. مراحل بکار رفته شامل فرایند پاکسازی و کاهش بعد می‌باشد. در نهایت تعداد ۱۴۰۲ مشتری برای انجام عملیات اعتبارسنجی باقی مانده است. فیلدهای که در این مقاله استفاده شده در جدول ۱ نشان داده شده است.

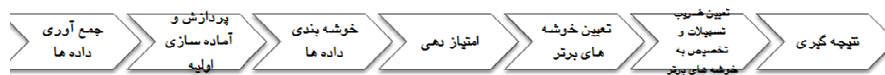
جدول ۱ فیلدهای مورد استفاده در تحقیق

مانده	گردش بستانکار	گردش بدهکار	مانده از قبل	نام مشتری	کد بورس
-------	---------------	-------------	--------------	-----------	---------

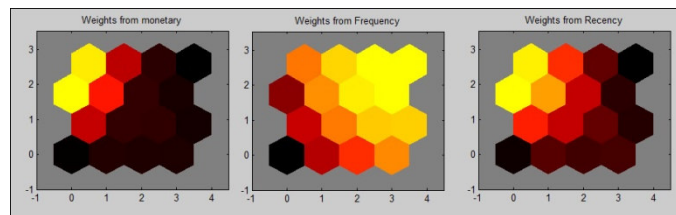
با مدل RFM که در بالا تعریف شد، داده‌ها آماده برای استفاده در عملیات اعتبارسنجی می‌شوند. در این راستا این سه شاخص تولید شده برای استفاده در الگوریتم خوشه‌بندی بکار می‌رود. همان طور که شکل ۲ نمودار صفحات وزن‌ها را برای هر یک از عناصر سه شاخص ورودی نشان می‌دهد. صفحات وزن‌ها به ازای هر عضو از بردار ورودی رسم می‌شوند. این نمودار وزن‌های اتصالی بین ورودی‌ها و نوروں را رسم می‌کند. رنگ‌های تیره‌تر نشان دهنده مقدار وزن‌های بیشتر و رنگ‌های روشن‌تر مقادیر کوچکتر وزن‌ها را نشان می‌دهد. در صورتی که الگوهای اتصالی بین دو ورودی مشابه باشند، می‌توان نتیجه گرفت دو ورودی دارای همبستگی بالایی هستند. در اینجا الگوی اتصالی بین ورودی‌ها متفاوت است.

۲-۴. خوشه بندی SOM بر روی داده های RFM

مدل امتیازدهی رفتاری در این مطالعه در نخست با به‌کارگیری الگوریتم SOM و شبکه‌ای در ابعاد 4×4 و نورون‌های شش ضلعی تولید شده است. هر یک از این سلول‌های عصبی از راه وزن‌های سیناپسی که در طول یادگیری به بردار ورودی متصل است، تنظیم می‌شوند. فاز اول SOM فاز برآورد ناهموار است که برای تولید الگوهای داده‌های ناخالص استفاده می‌شود. فاز دوم فاز تنظیم، مورد استفاده برای تنظیم نقشه شبکه به مدل ویژگی‌های خوب از داده‌ها می‌باشد [۱۵، ص ۲-۷].



شکل ۱ روند اجرای عملیات تحقیق



شکل ۲ بردارهای ورودی

انجام الگوریتم خوشه‌بندی به این علت است تا بتوان گروه‌هایی از مشتریان را برای رتبه‌بندی و اعطای تسهیلات شناسایی کرد. نتایج اجرای الگوریتم SOM در نرم‌افزار متلب در شکل‌های ۳ و ۴ نمایش داده شده است. در شکل ۳ تعداد مشتریانی را که در هر نورون تقسیم شده‌اند، نمایش می‌دهد. شکل ۴ نشان دهنده میزان فاصله بین نورون‌های همسایه از یکدیگر می‌باشد که هرچقدر میزان فاصله آن‌ها از یکدیگر بیشتر باشد، با رنگ تیره‌تر نشان داده می‌شود و هرچقدر این فاصله کمتر باشد، با رنگ‌های روشن‌تر نمایش می‌دهد. همانطور که در شکل ۴ مشاهده می‌شود، نورون‌هایی را که فاصله نزدیکی به هم دارند، می‌

توان یک خوشه در نظر گرفت. در جدول ۲ نتایج محاسبه SSE بر اساس تعداد خوشه‌های مختلف ممکن را مشاهده می‌کنید. در نتیجه ۱۰ خوشه از این شبکه می‌توان استخراج کرد. جدول ۳ شماره‌های خوشه‌ها را نشان می‌دهد.

جدول ۲ نرخ پارامتر SSE به ازای تعداد خوشه‌ها در SOM

SSE	تعداد خوشه‌ها
۱۷/۷۰۹۶۹	۸
۱۵/۷۵۸۳۱	۹
۱۴/۶۱۳۸۹	۱۰
۱۸/۶۱۰۴۹	۱۱
۱۹/۵۰۴۸۷	۱۲

جدول ۳ پراکنندگی خوشه‌ها

۹	۱۰	۱۰	۱۰
۷	۸	۱۰	۱۰
۵	۶	۴	۱۰
۱	۲	۳	۴

۳-۴. امتیاز دهی خوشه

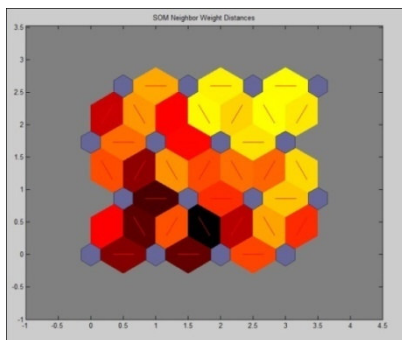
بعد از تعیین خوشه‌ها حالا می‌خواهیم ارزش هر خوشه را به دست آوریم تا از این راه رتبه‌بندی خوشه‌ها را مشخص کنیم. از این رو در ادامه عملیات مربوط به اعطای تسهیلات به خوشه‌های برتر را انجام می‌دهیم. وزن خوشه‌ها با استفاده از نظر خبرگان تعیین شد. با توجه به اهمیت شاخص ارزش پولی برای کارگزاری که نشان دهنده میزان حجم پولی معاملات افراد می‌باشد، وزن این متغیر نسبت به دو متغیر دیگر بیشتر است و مقدار ۰/۵ در نظر گرفته شد برای متغیر فرکانس مقدار ۰/۳ و برای متغیر تأخر ۰/۲ در نظر گرفته شده است. نتایج مربوط به هر خوشه در جدول ۳ نشان داده شده است. همان طور که مشاهده می‌کنید به ترتیب خوشه‌های ۵، ۱ و ۷ بالاترین امتیازها را به خود اختصاص داده‌اند. از آن جایی که در این پژوهش تمرکز بر مشتریان

با ارزش است، در نتیجه خوشه‌های که بالاترین رتبه را در مدل امتیازدهی دارد، مورد توجه قرار گرفته است. از این رو در ادامه به بررسی این سه خوشه پرداخته می‌شود.

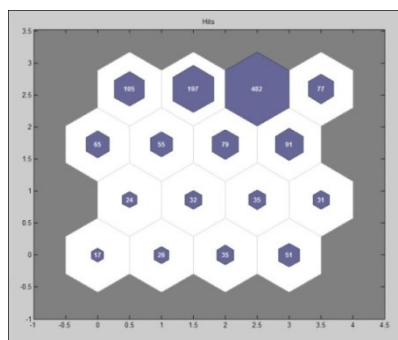
خوشه ۵. مشتریانی هستند که میزان گردش آنها بالای ۲۵ میلیون می‌باشد که معمولاً بازه به نسبت عالی دارند. تعداد ماه که تراکنش دارند، بین ۹ تا ۱۲ بار می‌باشد و تعداد مانده منفی زیر ۵ ماه دارند به این معنا است که حضور عالی در بازار دارند. حجم تراکنش در ماه آن‌ها نسبتاً خوب می‌باشد و تسویه خوبی دارند. تعداد مشتریان در این خوشه ۲۴ نفر می‌باشد.

خوشه ۱. مشتریانی هستند که متوسط میزان گردش آنها از ۵۰ میلیون به بالا است، تعداد ماه که تراکنش دارند بین ۱۰ تا ۱۲ می‌باشد که بسیار عالی است ولی تعداد مانده منفی بین ۶ تا ۱۱ ماه دارند که به این خاطر می‌باشد که چون حجم تراکنش بالایی دارند، این مقدار قابل اغماض می‌باشد. تعداد مشتریان در این خوشه ۱۷ نفر می‌باشد.

خوشه ۷. مشتریانی هستند که میزان گردش آنها بین ۲۵ میلیون تا ۷۵۰ میلیون می‌باشد که بازه به نسبت عالی دارند. تعداد ماه که تراکنش دارند بین ۸ تا ۱۲ بار می‌باشد و تعداد مانده منفی زیر ۵ ماه دارند؛ به این معنا که حضور عالی در بازار دارند. حجم تراکنش در ماه آن‌ها به نسبت خوب است و تسویه خوبی دارند. تعداد مشتریان در این خوشه ۶۵ نفر می‌باشد.



شکل ۴ میزان فاصله نرون‌ها از یکدیگر



شکل ۳ تعداد اعضای هر نرون



سه خوشه‌ای که در رتبه‌های بالاتری در امتیاز بندی قرار گرفتند ۸۰ درصد حجم مجموع گردش‌های بدهکاری و بستانکاری را به خود اختصاص داده‌اند. در نتیجه می‌توان تایید کرد که اعضای سه خوشه ۵، ۱ و ۷ بیشترین حجم معاملات و گردش‌های مالی را دارند در نتیجه باید توجه بیشتری را به آنها اختصاص داد تا از این طریق سودآوری بیشتری عاید شرکت کارگزاری شود. نتایج مربوط به درصد گردش مالی در شکل ۵ نشان داده شده است.

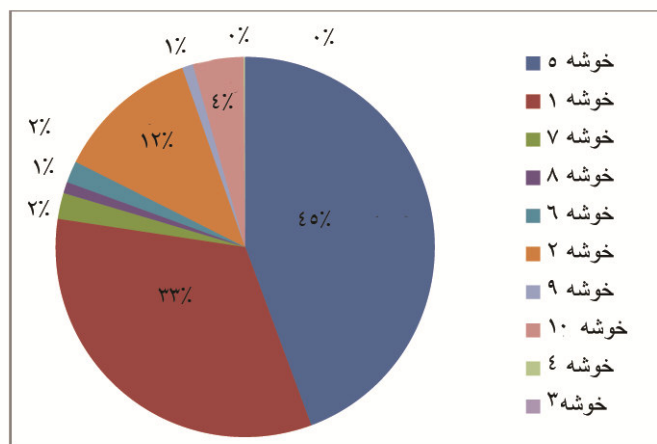
جدول ۴ ارزشیابی خوشه‌ها

تعداد اعضای خوشه	ارزش خوشه	میانگین ارزش تاخر	میانگین ارزش فرکانس	میانگین ارزش پول	خوشه
۲۴	۰/۷۷۰۶۴	۰/۷۵۳۹۴	۰/۹۴۰۹۷	۰/۶۷۳۹۳	۵
۱۷	۰/۶۹۶۵۷	۰/۳۳۸۲۴	۰/۹۶۰۷۸	۰/۶۸۱۳۷	۱
۶۵	۰/۵۸۳۹۲	۰/۸۰۲۵۶	۰/۷۸۳۳۳	۰/۳۷۶۸۱	۷
۵۵	۰/۴۸۶۸۲	۰/۹۴۳۹۴	۰/۵۶۶۶۷	۰/۲۵۶۰۶	۸
۳۲	۰/۳۷۹۷۸	۰/۷۴۴۷۹	۰/۴۵۵۷۳	۰/۱۸۸۲۱	۶
۲۶	۰/۳۷۹۵۸	۰/۴۴۲۳۱	۰/۵۳۵۲۶	۰/۲۶۱۰۷	۲
۱۰۵	۰/۳۷۳۶۷	۰/۹۵۷۱۴	۰/۳۶۳۴۹	۰/۱۴۶۳۹	۹
۹۵۷	۰/۲۴۶۹۱	۰/۹۷۷۶۲	۰/۱۰۶۱۵	۰/۰۳۹۰۹	۱۰
۸۶	۰/۱۶۹۴۲	۰/۵۰۱۹۴	۰/۱۴۹۲۳	۰/۰۴۸۵۴	۴
۳۵	۰/۰۳۶۸۲	۰/۱۰۲۳۸	۰/۰۳۵۷۱	۰/۰۱۱۲۶	۳

۴-۴. امتیازدهی اعضای خوشه‌های برتر برای ارائه تسهیلات

پس از تعیین خوشه‌های برتر در مرحله قبل، حالا نوبت آن رسیده است که به نحوی تسهیلات مورد نظر به اعضای این خوشه‌ها اختصاص داده شود. برای اینکه بتوان در مرحله بعد این اعتبارات را به مشتریان برتر واگذار کرد نیازمند آن می‌باشد که نخست ضریب تسهیلات اعطایی این سه خوشه برتر تعیین شود. برای این کار باید ارزش مشتریان این سه خوشه برتر محاسبه شود. با استفاده از فرمول محاسبه ارزش مشتری که در بخش ۳-۳ به آن اشاره شد، عملیات مربوط به امتیازدهی مشتریان موجود در این سه خوشه انجام می‌گیرد. در این مرحله نیز با توجه به اهمیت شاخص ارزش پولی ضریب این متغیر ۰/۵ در نظر گرفته شده همچنین

متغیر تناوب با ضریب ۰/۳ و متغیر تأخر با وزن ۰/۲ در فرمول اعمال می‌شوند. تعداد ۱۰۶ مشتری در مجموع در این ۳ خوشه قرار دارند که برای تمامی آنها مقدار ارزش تعیین می‌شود. سپس از تقسیم ارزش هر مشتری بر مجموع ارزش ۱۰۶ مشتری برتر، وزن هر مشتری تعیین می‌شود. مجموع وزن‌های تعیین شده برای مشتریان هر خوشه در انتها ضریب اعطای تسهیلات به اعضای آن خوشه را تعیین می‌کند. نتایج آن در جدول ۵ نشان داده شده است. با این کار مشخص می‌کنیم که به چه میزان از تسهیلات به این خوشه‌های برتر اختصاص پیدا می‌کند. با توجه به جدول ۵ به ترتیب برای خوشه‌های ۵، ۱ و ۷ از مقدار تسهیلات در نظر گرفته شده میزان ۲۷/۱، ۱۷/۳ و ۵۵/۶ درصد در نظر گرفته می‌شود. خوشه ۵ که برترین خوشه است میزان کمتری تسهیلات به نسبت خوشه ۷ که در رتبه سوم قرار دارد اختصاص پیدا می‌کند و دلیل این می‌باشد که تعداد اعضای خوشه ۷ بیشتر از خوشه ۵ است. در نتیجه این مقدار تسهیلات به نسبت کمتری به اعضای خوشه ۷ تقسیم می‌شود، از این رو تعداد اعضای خوشه‌ای که بالاترین رتبه را دارند به نسبت میزان بیشتری از تسهیلات به آنها اختصاص می‌یابد



شکل ۵ درصد گردش مالی

جدول ۵ ضریب تسهیلات برای خوشه های برتر با روش وزن دهی خطی

خوشه	ضریب تسهیلات
۵	۰/۲۷۱
۱	۰/۱۷۳
۷	۰/۵۵۶

برای اینکه محققین بتوانند مدل خود برای اعطای اعتبار را کامل کنند، از ضریب (وزن) خوشه تولید شده در مرحله قبل استفاده می کنند. روند به این گونه است که برای هر خوشه نخست مجموع ارزش های تمامی مشتریان آن ها محاسبه می شود، سپس برای هر فرد ارزش مشتری آن تقسیم بر مقدار مجموع به دست آمده می شود، با این کار ضریب اولیه اعطای تسهیلات برای هر مشتری تعیین می شود از ضرب وزن آن خوشه در وزن هر مشتری، ضریب اعطای تسهیلات مشتری تعیین می شود. نتایج مربوطه در جدول ۶ و ۷ و ۸ نشان داده شده است.

همانطور که در جدول نشان داده شد، ضرایب مربوط به اعضای خوشه های برتر به تفکیک بر اساس ارزش مشتریان هر خوشه تعیین شد. همچنین می توان ضریب بدست آمده برای هر مشتری را به عنوان ضریب جداگانه برای دریافت تسهیلات در نظر گرفت. در حال حاضر میزان اعتباری که شرکت می تواند در هر دوره به مشتریان خود اختصاص دهد ریال مبلغی بالغ بر ۱۰۰ میلیارد می باشد.

جدول ۶ ضریب و مقدار تسهیلات برای خوشه ۵ با روش وزن دهی خطی

شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری	شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری
۵۱	۰/۰۴۶۸۹۷	۰/۰۱۲۵۹۰۸	۱۲۵۹۰۸۲۵۸۹	۵۱۳	۰/۰۴۱۶۵۶۴۲۷	۰/۰۱۱۲۸۱۸۲۴	۱۱۲۸۱۸۲۳۷۳
۵۲	۰/۰۲۸۲۹۶۸	۰/۰۱۰۳۷۲۱۸	۱۰۳۷۲۱۷۸۱۶	۵۱۴	۰/۰۲۹۶۰۸۴۲۱	۰/۰۱۰۷۲۷۱۶۲	۱۰۷۲۷۱۶۷۹
۵۳	۰/۰۴۱۶۶۱۵۴	۰/۰۱۱۲۷۰۷۳	۱۱۲۷۰۷۳۰۴۹	۵۱۵	۰/۰۵۰۲۵۸۰۴۹	۰/۰۱۳۶۱۱۴۰۴	۱۳۶۱۱۴۰۳۸۵

ادامه جدول ۶

شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری	شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری
C۴	۰/۰۴۱۰۸۲۹۸۵	۰/۰۱۱۱۲۶۵۱۸	۱۱۱۲۶۵۱۸۳۹	C۱۶	۰/۰۴۳۷۸۶۳۵۲	۰/۰۱۱۸۵۸۶۷۲	۱۱۸۵۸۶۷۲۱۴
C۵	۰/۰۳۶۷۸۲۱۷۴	۰/۰۹۹۶۱۷۲۸	۹۹۶۱۷۲۸۳۲	C۱۷	۰/۰۵۱۳۶۳۹۷۱	۰/۰۱۳۹۱۰۹۲۱	۱۳۹۱۰۹۲۱۲۹
C۶	۰/۰۳۸۴۲۰۵۷۸	۰/۰۱۰۴۰۵۴۵۸	۱۰۰۵۴۵۷۸۷	C۱۸	۰/۰۳۹۹۷۷۰۶۲	۰/۰۱۰۸۳۷۰۰۱	۱۰۸۳۷۰۰۹۴
C۷	۰/۰۳۸۳۰۱۷۹	۰/۰۱۰۵۱۶۳۹	۱۰۵۱۶۳۹۰۲۶	C۱۹	۰/۰۳۶۴۱۳۵۳۳	۰/۰۰۹۸۶۱۸۸۹	۹۸۶۱۸۸۹۱۷۹
C۸	۰/۰۳۶۰۰۳۹۳۲	۰/۰۰۹۷۵۰۹۵۷	۹۷۵۰۹۵۷۹/۲	C۲۰	۰/۰۳۶۷۴۱۳۱۴	۰/۰۰۹۹۵۰۶۳۵	۹۹۵۰۶۳۵۰۸/۸
C۹	۰/۰۴۲۵۵۷۵۹	۰/۰۱۱۵۲۵۸۷۵	۱۱۵۲۵۸۷۴۹۸	C۲۱	۰/۰۴۱۳۲۸۷۴۶	۰/۰۱۱۱۹۳۰۷۸	۱۱۱۹۳۰۷۸۲
C۱۰	۰/۰۴۶۴۸۷۵۹	۰/۰۱۲۵۷۹۷۳۳	۱۲۵۷۹۳۳۶۵	C۲۲	۰/۰۴۴۵۶۴۵۹۴	۰/۰۱۲۰۶۹۴۴۴	۱۲۰۶۹۴۴۳۶۷
C۱۱	۰/۰۴۲۴۷۵۶۳۹	۰/۰۱۱۵۰۳۶۸۹	۱۱۵۰۳۶۸۵۰	C۲۳	۰/۰۴۲۹۳۶۱۹	۰/۰۱۱۶۲۵۷۱۴	۱۱۶۲۵۷۱۴۱۳
C۱۲	۰/۰۴۴۰۷۳۰۷۳	۰/۰۱۱۹۳۶۳۲۵	۱۱۹۳۶۳۲۴۸۱	C۲۴	۰/۰۳۸۲۹۷۸	۰/۰۱۰۳۷۲۱۷۸	۱۰۳۷۲۱۷۸۱۶

جدول ۷ ضریب و مقدار تسهیلات برای خوشه ۱ با روش وزن دهی خطی

شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری	شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	مقدار تسهیلات مشتری
C۲۵	۰/۰۵۰۱۸۸۷۳	۰/۰۰۹۵۴۰۱۸۵	۹۵۴۰۱۸۵۲۵/۷	C۲۴	۰/۰۶۵۱۹۰۹۶۷	۰/۰۱۱۳۰۴۰۱	۱۱۳۰۴۰۱۲۱
C۲۶	۰/۰۵۰۵۴۰۵۹۲	۰/۰۰۸۶۶۳۶۵۹	۸۶۶۳۶۵۸۵۵	C۳۵	۰/۰۵۱۶۹۲۱۵	۰/۰۰۸۹۶۳۳۳۷	۸۹۶۳۳۳۶۸۴۰۶
C۲۷	۰/۰۶۰۲۰۰۸۱۳	۰/۰۱۰۴۳۸۷۲۸	۱۰۴۳۸۷۲۵۹	C۳۶	۰/۰۶۱۰۳۲۵۶۳	۰/۰۱۰۵۸۲۹۵	۱۰۵۸۲۹۴۹۶۹
C۲۸	۰/۰۶۵۷۰۲۷۷	۰/۰۱۱۳۹۲۷۵۶	۱۱۳۹۲۷۵۶۱۲	C۳۷	۰/۰۷۶۰۰۲۸۱۵	۰/۰۱۳۱۷۸۷۶۷	۱۳۱۷۸۷۶۵۴
C۲۹	۰/۰۵۵۴۰۲۷۲۵	۰/۰۰۹۶۰۶۷۴۵	۹۶۰۶۷۴۴۶۷/۹	C۲۸	۰/۰۶۳۵۲۷۰۵	۰/۰۱۱۰۱۵۵۸۶	۱۱۰۱۵۵۸۶۰۰
C۳۰	۰/۰۶۱۱۰۵۱۴	۰/۰۱۰۶۰۵۱۳۶	۱۰۶۰۵۱۳۶۱۷	C۲۹	۰/۰۵۳۶۱۱۴۱۳	۰/۰۰۹۲۹۶۱۳۴	۹۲۹۶۱۳۴۰۰/۶
C۳۱	۰/۰۶۵۱۹۰۹۶۷	۰/۰۱۱۳۰۴۰۱	۱۱۳۰۴۰۱۰۲۱	C۴۰	۰/۰۶۴۶۷۹۱۶۳	۰/۰۱۱۲۱۵۲۶۴	۱۱۲۱۵۲۶۴۰
C۳۲	۰/۰۵۵۱۴۶۸۲۴	۰/۰۰۹۵۶۳۳۷۲	۹۵۶۳۳۷۱۷۳/۴	C۴۱	۰/۰۴۹۵۱۶۹۸۵	۰/۰۰۸۵۸۶۱۶۷	۸۵۸۶۱۶۶۷۳/۱
C۳۳	۰/۰۴۶۳۸۲۱۸۹	۰/۰۰۸۰۴۲۵۹۸	۸۰۴۲۵۹۸۰۳/۶				



جدول ۸ ضریب و مقدار تسهیلات برای خوشه ۷ با روش وزن دهی خطی

شماره مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	شماره مشتری	مقدار تسهیلات مشتری	ضریب ارزش مشتری	ضریب تسهیلات مشتری	شماره مشتری	مقدار تسهیلات مشتری
C۴۲	۰/۰۱۵۲۴۹۵۰۱	۰/۰۰۸۴۷۵۲۳۴	C۶۴	۰/۰۱۲۷۵۲۴۹۵	۰/۰۰۷۶۴۳۲۴۱	۰/۰۰۸۱۷۵۷۱۷	C۸۶	۰/۰۰۷۶۴۳۲۴۱
C۴۳	۰/۰۱۷۸۴۴۳۱۱	۰/۰۰۹۹۱۷۳۵۵	C۶۵	۰/۰۱۴۹۳۰۱۴	۰/۰۰۸۲۹۷۷۴۳	۰/۰۰۸۳۰۸۳۶	C۸۷	۰/۰۰۸۲۹۷۷۴۳
C۴۴	۰/۰۱۵۲۰۹۵۸۱	۰/۰۰۸۴۵۳۰۴۸	C۶۶	۰/۰۱۵۵۸۸۲۲	۰/۰۰۸۱۶۳۸۱۹	۰/۰۰۷۸۶۵۱۰۶	C۸۸	۰/۰۰۸۱۶۳۸۱۹
C۴۵	۰/۰۱۴۲۱۱۵۷۷	۰/۰۰۷۸۹۸۳۸۶	C۶۷	۰/۰۱۳۸۳۳۳۵	۰/۰۰۷۶۸۷۶۱۴	۰/۰۰۹۱۰۷۵۴۹	C۸۹	۰/۰۰۷۶۸۷۶۱۴
C۴۶	۰/۰۱۵۱۸۹۶۲۱	۰/۰۰۸۴۴۱۹۵۵	C۶۸	۰/۰۱۷۲۴۵۵۰۹	۰/۰۰۹۵۸۴۵۵۸	۰/۰۰۶۹۹۹۸۳۴	C۹۰	۰/۰۰۹۵۸۴۵۵۸
C۴۷	۰/۰۱۵۱۶۹۶۶۱	۰/۰۰۸۴۳۰۸۶۱	C۶۹	۰/۰۱۴۶۵۰۶۹۹	۰/۰۰۸۱۴۲۴۳۷	۰/۰۰۸۳۹۷۵۸۲	C۹۱	۰/۰۰۸۱۴۲۴۳۷
C۴۸	۰/۰۱۶۷۰۶۵۸۷	۰/۰۰۹۲۸۵۰۴۱	C۷۰	۰/۰۱۵۲۰۹۵۸۱	۰/۰۰۸۴۵۳۰۴۸	۰/۰۰۹۴۰۷۰۶۶	C۹۲	۰/۰۰۸۴۵۳۰۴۸
C۴۹	۰/۰۱۶۳۲۷۳۴۵	۰/۰۰۹۰۷۴۲۶۹	C۷۱	۰/۰۱۵۰۴۹۹	۰/۰۰۸۳۶۴۳۰۲	۰/۰۰۷۷۴۳۰۸۱	C۹۳	۰/۰۰۸۳۶۴۳۰۲
C۵۰	۰/۰۱۵۵۶۸۸۶۲	۰/۰۰۸۶۵۲۷۲۶	C۷۲	۰/۰۱۲۷۵۲۴۹۵	۰/۰۰۷۶۴۳۲۴۱	۰/۰۰۵۳۸۵۷۷	C۹۴	۰/۰۰۷۶۴۳۲۴۱
C۵۱	۰/۰۱۴۸۱۰۳۷۹	۰/۰۰۸۳۳۱۱۸۳	C۷۳	۰/۰۱۴۶۵۰۶۹۹	۰/۰۰۸۱۴۲۴۳۷	۰/۰۰۹۲۸۵۰۴۱	C۹۵	۰/۰۰۸۱۴۲۴۳۷
C۵۲	۰/۰۱۷۳۲۵۳۴۹	۰/۰۰۹۶۲۸۹۳۱	C۷۴	۰/۰۱۴۷۳۰۵۳۹	۰/۰۰۸۱۸۶۸۱	۰/۰۰۹۲۸۵۰۴۱	C۹۶	۰/۰۰۸۱۸۶۸۱
C۵۳	۰/۰۱۳۸۷۳۲۵۵	۰/۰۰۷۷۰۹۸۰۱	C۷۵	۰/۰۱۷۱۸۵۶۲۹	۰/۰۰۹۵۵۱۳۷۸	۰/۰۰۹۶۷۳۳۰۴	C۹۷	۰/۰۰۹۵۵۱۳۷۸
C۵۴	0.017145709	۰/۰۰۹۵۲۹۰۹۲	C۷۶	۰/۰۱۴۷۵۰۴۹۹	۰/۰۰۸۱۹۷۹۰۳	۰/۰۰۸۰۳۱۵۰۵	C۹۸	۰/۰۰۸۱۹۷۹۰۳
C۵۵	۰/۰۱۶۳۲۷۳۴۵	۰/۰۰۹۰۷۴۲۶۹	C۷۷	۰/۰۱۲۰۷۵۸۴۸	۰/۰۰۶۷۱۱۴۰۹	۰/۰۰۹۶۰۶۷۴۵	C۹۹	۰/۰۰۶۷۱۱۴۰۹
C۵۶	۰/۰۱۲۶۱۴۷۷	۰/۰۰۷۰۱۰۹۲۷	C۷۸	۰/۰۱۵۳۴۹۳۰۱	۰/۰۰۸۵۳۰۷۰۱	۰/۰۰۸۴۸۶۳۲۸	C۱۰۰	۰/۰۰۸۵۳۰۷۰۱
C۵۷	۰/۰۱۶۸۱۶۶۲۷	۰/۰۰۹۳۷۳۷۸۷	C۷۹	۰/۰۱۳۵۳۲۹۳۴	۰/۰۰۷۵۲۱۲۱۶	۰/۰۰۸۵۱۹۶۰۷	C۱۰۱	۰/۰۰۷۵۲۱۲۱۶
c58	۰/۰۱۵۰۴۹۹	۰/۰۰۸۳۶۴۳۰۲	C۸۰	۰/۰۱۶۶۴۶۷۰۷	۰/۰۰۹۲۵۱۷۱۱	۰/۰۰۸۲۴۲۲۷۶	C۱۰۲	۰/۰۰۹۲۵۱۷۱۱
C۵۹	۰/۰۱۴۹۵۰۱	۰/۰۰۸۳۶۴۳۰۲	C۸۱	۰/۰۱۴۰۱۱۹۷۶	۰/۰۰۷۷۸۷۴۵۴	۰/۰۰۹۱۶۳۰۱۵	C۱۰۳	۰/۰۰۷۷۸۷۴۵۴
C۶۰	۰/۰۱۶۸۴۶۳۰۷	۰/۰۰۹۳۶۳۶۹۳	C۸۲	۰/۰۱۷۱۲۵۷۴۹	۰/۰۰۹۵۱۷۹۹۹	۰/۰۰۸۶۹۷۰۹۹	C۱۰۴	۰/۰۰۹۵۱۷۹۹۹
C۶۱	۰/۰۱۵۳۰۹۳۸۱	۰/۰۰۸۵۰۸۵۱۴	C۸۳	۰/۰۱۵۴۴۹۱۰۲	۰/۰۰۸۵۸۶۱۶۷	۰/۰۰۹۴۸۴۷۱۹	C۱۰۵	۰/۰۰۸۵۸۶۱۶۷
C۶۲	۰/۰۱۳۷۷۲۴۵۵	۰/۰۰۷۶۵۴۳۳۵	C۸۴	۰/۰۱۴۶۵۰۶۹۹	۰/۰۰۸۱۴۲۴۳۷	۰/۰۰۷۹۰۹۴۷۹	C۱۰۶	۰/۰۰۸۱۴۲۴۳۷
C۶۳	۰/۰۱۷۴۲۵۱۵	۰/۰۰۹۶۸۴۳۹۷	C۸۵	۰/۰۱۵۱۲۹۷۴۱	۰/۰۰۸۴۰۸۶۷۵			

۵- نتیجه‌گیری

از آنچه گفته شده می‌توان ادعان کرد که اعتبارسنجی مشتریان موضوعی مهم برای شرکت‌های کارگزار در بازار سرمایه امروز صنعت بورس است. بنابراین از طریق خوشه‌بندی مشتریان امکان شناسایی و درک درست از مشتریان موجود و از راه روش‌های امتیازدهی امکان ارائه رتبه‌بندی رفتاری مشتریان برای بیان برترین بخش از مشتریان را خواهیم داشت. تا در نهایت بتوان مدلی برای اعتبارسنجی مشتریان ایجاد کرد تا این امکان را برای مدیران و کارگزاران فراهم سازد تا از تحلیل این نتایج به اتخاذ تصمیم‌ها، استراتژی‌های مناسب برای تعیین تسهیلات اعطایی به مشتریان بپردازد.

در این تحقیق ابتدا اقدام به آماده‌سازی و پردازش اولیه داده‌های جمع‌آوری شده از داده‌های رفتاری مشتریان کارگزاری شد. پس از تبدیل داده‌ها به شکل مدل RFM آنها با استفاده از الگوریتم شبکه عصبی SOM خوشه‌بندی شدند. خوشه‌ها برای تعیین برترین خوشه‌ها بر اساس فرمول ارایه شده رتبه‌بندی شدند. با استفاده از فرمول ارزش مشتری ضریب ارایه تسهیلات این خوشه‌های برتر تعیین شدند، سپس برای آنکه بتوان ضریب تسهیلات را مشخص کرد با استفاده از همان فرمول ارایه شده برای هر خوشه ضریب هر مشتری تعیین شد. از ضرب وزن خوشه در وزن هر مشتری آن خوشه در مقدار کل تسهیلات، ضریب تسهیلات اعطایی به هر مشتری آن خوشه بدست می‌آید.

نتایج نشان داد که سه خوشه ۵، ۱ و ۷ بالاترین امتیازها را برای دریافت تسهیلات بدست آوردند و ضریب تسهیلات دریافتی برای آنها به ترتیب مقدار ۰/۲۷۱، ۰/۱۷۳ و ۰/۵۵۶ می‌باشد. با مشخص شدن ضریب تسهیلات برای خوشه و به تبع آن برای مشتریان حاضر در این گروه‌های برتر عمل اعطای تسهیلات شفاف‌تر و هدفمندتر می‌شود و از این طریق به شرکت در افزایش سودآوری، کاهش رویگردانی مشتریان پر بازده و ارزش آفرینی برای مشتریان کمک می‌کند. این تحقیق با بیان روشی نظام‌مند برای اعطای تسهیلات راه شناخت ارزش واقعی بر-اساس توانمندی و جلوگیری از اعمال سلیقه‌ها را نشان می‌دهد.

در پژوهش‌های دیگر از جمله تحقیق «شین و سون» در سال (۲۰۰۴) و جی. لوپز و همکاران در سال ۲۰۱۱، بعد از بخش بندی مشتریان و تعیین ویژگی‌های آنها مقدار تعرفه (کمسیون) را برای هر گروه مشخص کرده‌اند. در این تحقیق به جای تعیین تعرفه یا



کارمزد شرکت برای هر گروه میزان اعطای تسهیلات مشخص شد. همچنین علاوه بر تعیین میزان تسهیلات برای گروه های برتر، مقدار این تسهیلات برای اعضای این گروه‌ها هم محاسبه شد. نوآوری دیگر پژوهش توجه به حوزه کمتر مورد توجه قرار گرفته در اعتبارسنجی یعنی صنعت بورس می باشد. از محدودیت‌های تحقیق می‌توان به پراکندگی شدید در میزان تراکنش‌های مشتریان که نتایج حاصل از الگوریتم‌ها را بی‌معنا نشان می‌داد اشاره کرد، این مشکل با استفاده از پردازش داده‌ها و تبدیل آنها به مدل RFM تا حدی زیادی حل شد. برای تحقیقات آتی استفاده از مدل‌های دیگر خوشه‌بندی از جمله روش K-means و یا سایر مدل‌های رتبه‌بندی مانند روش‌های AHP و ANP و استفاده از شاخص بیرونی و محیطی سازمان متقاضی وام، مدل دیگری را ارائه کرد.

۶- پی‌نوشت‌ها

1. Recency, Frequency, Monetary
2. Self-Organizing Map
3. Character
4. Capacity
5. Capital
6. Collateral
7. Conditions of Economic
8. Newell
9. Apriori
10. Hopfield's autonomous recurrent neural network
11. Calinski index
12. Davies Bouldin index
13. mean adequate index
14. Sum of Squared Error

۷- منابع

- [1] T. Hasangholi-poor, et al., "Market segmentation by using artificial neural network: Case study; Meat products (sausages)", JRMI331264883400, 2007.
- [2] Yi-Hui Liang, "Integration of data mining technologies to analyze customer value for the automotive maintenance industry", *Expert Systems with Applications*, 37,

pp. 7489–7496, 2010.

- [3] Peter. C. Verhoef, Bas Donkers, “Predicting customer potential value an application in the insurance industry”, *Decision Support Systems*, 32, pp. 189–199, 2001.
- [4] A. Shahlaie Mogzhadam, “Validate Credit Card Customers Using Genetic Algorithms”, *6th International Conference on Industrial Engineering*, 2008.
- [5] A. R. Hasanzadeh, et al., “Classification of mobile banking users by data mining approach: Comparison between artificial neural networks and naïve bayes techniques”, *Journal of Management Research in Iran*, Volume 16, Number 2, 2012.
- [6] M. Khanbabai, “Using clustering techniques and genetic algorithms in the construction of optimal decision tree for the classification of banks' customers”, 2009.
- [7] S. Jamshidi, “Customer authentication methods”, *Institute for Monetary and Banking*, 2009 .
- [8] S. R. Mousavi, E. Gholipour, “Rating criteria, of bank customer validation with Delphi approach”, *First International Conference on Marketing of Banking Services*, 2009.
- [9] S. Rashidian, “Bank grid customer classification based on credit risk using predictive models and multi-criteria decision”, Master's thesis, University of Engineering & Technology, Azad University of Sanandaj, 2011 .
- [10] H.H. Liu, C.S. Ong, “Variable selection in clustering for marketing segmentation using genetic algorithms”, *Expert Systems with Applications*, 34, pp. 502–510, 2008.
- [11] J. López, et. al., “Hopfield–K-Means clustering algorithm: A proposal for the segmentation of electricity customers”, *Electric Power Systems Research*, 81, pp. 716–724, 2011.
- [12] B. Beedbad, et. al., “Classification of bank customers by combining conceptual



- models, fuzzy hierarchy and clustering”, 2010.
- [13] Wu, J., & Lin, Z. “Research on Customer Segmentation Model by Clustering”. *ACM International Conference Proceeding Series*, 113. 2005.
- [14] Yeh, C., Yang, K. & Ting, T.; "Knowledge Discovery on RFM Model Using Bernoulli Sequence", *Expert Systems with Applications*, Vol. 36, pp. 5866–5871, 2009.
- [15] Jonkera, J-J., Piersmab, N. & Van den Poelc, D.; "Joint optimization of customer segmentation and marketing policy to maximize long-term profitability", *Expert Systems with Applications*, Vol. 27, pp. 159–168, 2004.
- [16] N. C. Hsieh, “An integrated data mining and behavioral scoring model for analyzing bank customers”, *Expert Systems with Applications*, 27, pp. 623–633, 2004.
- [17] B. Sohrabi, Amir Khanlari, “customer lifetime value (CLV) measurement based on RFM model” , *Journal of Iranian Accounting & Auditing Review*, Vol. 14, No. 47, pp. 7- 20, Spring 2007.
- [18] B. Foss, M. Stone, "CRM in Financial Services", ISBN 0 7494 3696 4, 2002
- [19] Lumsden SA, Beldona S, Morison AM. Customer value in an all-inclusive travel vacation club: An application of the RFM framework. *J. Hosp. Leisure Mark.*, 16(3): 2, pp. 70-285, 2008.
- [20] Wang, C.H., "Apply robust segmentation to the service industry using kernel induced fuzzy clustering techniques", *Expert Systems with Applications.*, 37: pp. 8395-8400, 2010.
- [21] S. Y. Husseini, & et al., “Segmenting and Profiling Green Consumers with Use of Self Organizing Maps”, *Journal of Management Research in Iran*, Volume 17, Number 2, 2013.
- [22] Tsai, Lu. ,”Customer churn prediction by hybrid neural networks”, *Expert Systems with Applications*, Vol. 36, pp. 12547-12553, 2009.
- [23] Ghazanfari, M. et al, “Customer segmentation export edible fruits”, *Quarterly*

Journal of Commerce, No. 55, 151 – 181, 2011.

- [24] Momeni, M. "Data Clustering (cluster analysis)", ISBN 978-964-04-7560-7, 2012.
- [25] Khademi Zare, H., T. Ali Heydari, "Assessment and clustering behavior of customer credit using the combination of AHP and Neural Network Method", 8th International Conference on Management, 2010.
- [26] B. Masnani, F Asghari, "Applying data mining to discover a scoring model and analyze the behavior of bank customers", IDMC2008, Amirkabir University, 2008.
- [27] M., Masani, "Segmenting customers using customer lifetime value," Master's thesis, University of Engineering, Tarbiat Modares University, 2009
- [28] K. Tsipstsis, A. Chorianopoulos, "Data Mining Techniques in CRM", John Wiley & Sons, Ltd, ISBN: 978-0-470-74397-3, 2009.
- [29] P. Bligh, D. Turk, "CRM Unplugged (Releasing CRM's Strategic Value)", John Wiley & Sons Inc., ISBN 0-471-48304-4, 2004.
- [30] Shiow-Yun Ch., Tsung-Yuan Y., "An artificial immune classifier for credit scoring analysis", *Applied Soft Computing*, 12; pp. 611–618, 2012.
- [31] Hussein Abdou, et. al., "Neural nets versus conventional techniques in credit scoring in Egyptian banking", *Expert Systems with Applications*, 35; pp. 1275–1292, 2008.
- [32] Tan, & Teinvach & Kumar, "Introduction to Data mining", 2004.
- [33] Jo-Ting Wei, et al, "A review of the application of RFM model", *African Journal of Business Management*, Vol. 4(19), pp. 4199-4206, 2010.
- [34] Charles R. Schmidt, Effects of Irregular Topology in Spherical Self-Organizing Maps, A Thesis Presented to the Faculty of San Diego State University, 2008.
- [35] T.S. Lee, et .al., "Mining the customer credit using classification and regression tree and multivariate adaptive regression splines", *Computational Statistics & Data Analysis*, 50, pp. 1113-1130, 2006.



- [36] Bee Wah Yap, , “Using data mining to improve assessment of credit worthiness via credit scoring models”, *Expert Systems with Applications*, 38; pp. 13274–13283, 2011.
- [37] Ngai, et .al., “Application of data mining techniques in customer relationship management: A literature review and classification”, *Expert Systems with Applications*, 36; pp. 2592–2602, 2009.
- [38] Chu Chai Henry Chan, “Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer”, *Expert Systems with Applications*, 34; 2754–2762, 2008.
- [39] G. Nie, et .al., “Credit card customer analysis based on panel data clustering”, *Procedia Computer Science*, 1; pp. 2489–2497, 2010.
- [40] Shian-Chang Huang, et. al., “A case study of applying data mining techniques in an outfitter’s customer value analysis”, *Expert Systems with Applications*, 36; pp. 5909–5915, 2009.
- [41] R.J. Kuo, et.al, “Integration of self-organizing feature maps neural network and genetic K-means algorithm for market segmentation”, *Expert Systems with Applications*, 30, pp. 313–324, 2006.
- [42] B. Sohrabi, A. Khanlari, “customer lifetime value (CLV) measurement based on rfm model”, *Iranian Accounting & Auditing Review*, Vol. 14 No. 47, pp. 7- 20, Spring 2007.
- [43] S. M. Seyed Hosseini, “Cluster analysis using data mining approach to develop CRM methodology to assess the customer”, *Expert Systems with Applications*, 37; pp. 5259–5264, 2010.