# Customer Credit Clustering for Presenting Appropriate Facilities

**Amir Afsar[1], Rahmat Houshdar Mahjoub[2*], Behroz Minaie[3]**

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

Archive of SID

www.SID.ir

# 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: ۹۲/٤/۲۲

## 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 کمّی 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 می شوند. سپس با استفاده از شبکه عصبی SOM [۲] به عنوان یکی از الگوریتم های خوشه بندی ، مشتریان به ۱۰ خوشه تبدیل خواهند شد. در ادامه با استفاده از مدل پیشنهادی، خوشه‌ها رتبه‌بندی می شوند. خوشه‌های برتر شناسایی و عملیات اعطای تسهیلات برای اعضای این خوشه‌ها انجام می‌شود. در نهایت سه خوشه ۵، ۱ و ۷ به عنوان خوشه های برتر تعیین شدند که به عنوان مشتریان هدف می باشند . ضریب تسهیلات اعطایی به این سه خوشه برتر به ترتیب ۰/۲۷۱ ، ۰/۱۷۳ و ۰/۵۵٦ می باشد.

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

---

* نویسندهٔ مسؤول مقاله: E-mail: houshdar.mahjoob@gmail.com

Archive of SID

www.SID.ir

## ۱- مقدمه

در بازار رقابتی امروز در بسیاری از صنایع، شرکتها در جهت تحقق کامل استراتژی های کسب وکار 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 ٦۰-٦۲]

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 می پردازند، خدمات سفارش‌سازی شده‌ای را برای مشتریان متنوع ارائه می دهند این کار باعث افزایش سود-آوری آنها می شود چون از این راه می‌توان مشتریان پر ارزش را حفظ کرد. [۲، صص۱-۲].

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 [٤، صص۲-٤].

بدون شک ارزیابی مشتریان برای اعطای اعتبار موضوع مهمی در حوزه خرید اعتباری سهام به حساب می آید. وقتی مشتریان هدف به درستی شناسایی نمی شود هم ممکن است باعث ریزش گروه پر سوده شود هم باعث می شود 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 می دهد و از طرف دیگر امکان شکست اقتصادی را برای سازمان



Archive of SID

www.SID.ir

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. شرکت‌ها می‌توانند با مدل‌های داده‌کاوی مشتریان همگن با ارزش یکسان را در خوشه‌های کوچکتر مشابه قرار دهند تا به تقسیم‌بندی بازار بپردازند [۱][٥] و از این راه تمرکز و تلاش خود را به شناسایی فرصت‌ها و ارائه خدمات به شیوه سفارشی مبذول دارد [۲، صص۱-۳]. مزایای اعتبارسنجی شامل کاهش هزینه‌های تحلیل اعتبار، امکان اتخاذ سریعتر تصمیم‌های اعتباری، نظارت دقیق‌تر بر حساب های موجود و .... می باشد [٦، ص ٥].

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 می‌شود. بخش سوم اختصاص به روش تحقیق دارد. بخش چهارم نحوه آماده‌سازی داده‌ها و انجام روند اعتبارسنجی مشتریان شرح داده شده است و در انتها در قسمت پنجم نتیجه‌گیری گفته می‌شود .

## ۲- پیشینه تحقیق

اعتبارسنجی یعنی سنجیدن ظرفیت افراد در استفاده از منابع از جمله تسهیلات اعطایی می‌باشد. شیوه‌های اعتبارسنجی مشتریان، اساس رتبه‌بندی یا رتبه‌سنجی 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 سرمایه‌گذاران در بازارهای سرمایه، می‌باشد. با شیوه‌های اعتبارسنجی می‌توان ریسک سرمایه‌گذاری برای سرمایه‌گذاران را پیش‌بینی و برآورد کرد.

موسسات اعتباری در گذشته به علت محدودیت در منابع و دسترسی به اطلاعات مشتریان



Archive of SID

www.SID.ir

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

تحقیقاتی در این زمینه انجام شده که در این میان می‌توان به مطالعه «فیشر» (۱۹۳٦) به عنوان اولین سیستم ارزیابی تقاضای اعتبار و مطالعه « دوراند» (۱۹٤۱) – 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 ویژگی-های فردی[۳]، ظرفیت[٤]، سرمایه[٥]، وثیقه[٦] و شرایط اقتصادی[۷] می‌باشند [۸].

موسسات اعتباری تلاش می‌کنند تا از روش‌ها و ابزار مختلف برای تصمیم‌گیری خودکار بهره گیرند 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 [٤،صص۲-٥] [۹]. علاوه بر این چند روش برجسته در هوش مصنوعی همچون سیستم های 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 [٤،ص۲]. 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 شرکت های دیگر استفاده کنند [۲،صص۲-۷].

این روشن است که افراد مختلف نیازهای متفاوتی دارند. برای پاسخگویی به نیازهای مختلف، آن‌ها را به بخش های کوچکتر تقسیم می کنیم تا برنامه های متناسب آنها طرح ریزی



Archive of SID

www.SID.ir

شود [۱۰] . شناخت درست از مشتری اجازه می دهد تا شرکت‌ها به ارائه تعرفه خاص مطابق با ویژگی های 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 روش های پارامتریک و ناپارامتریک تقریبا 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، مدل های مختلف درخت تصمیم‌گیری، شبکه-های عصبی، ماشین بردار پشتیبان، سیستم‌های خبره، منطق فازی، مدل‌های ترکیبی طبقه کننده‌ها، الگوریتم پس انتشار شبکه عصبی، Bagging، Boosting ، نزدیک ترین همسایه و الگوریتم ژنتیک و ترکیب برخی موارد فوق با هم [٦، صص۲-٥].

فرنس کیس مدل‌های رتبه‌بندی اعتباری را در دو گروه ارزیابی پارامتری و درخت دسته-بندی، تقسیم می‌کنند. ایده جداسازی گروه‌ها در یک جمعیت به وسیله فیشر در سال ۱۹۳٦ مطرح شد. التمن در سال ۱۹۳۸ اولین سیستم ارزیابی تقاضانامه‌های اعتباری را با بکارگیری ٥ معیار توسعه داد. دوراند در سال ۱۹٤۱ یک پروژه تحقیقاتی برای اداره ملی تحقیقات اقتصادی آمریکا انجام داد. ایشان بنیانگذار سیستم‌های رتبه‌بندی اعتبار به روش کنونی می-باشند. در سال ۱۹٦۳ مایرزو هنرجی آنالیز ممیزی چند متغیره برای رتبه‌بندی اعتباری مشتریان ارائه نمودند. مور و کلن در سال ۱۹٦۸ 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.

دسایی 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 طبقه بندی مشتریان وام‌های بین المللی 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



Archive of SID

www.SID.ir

این پژوهش با استفاده از اطلاعات مربوط به اعتبارات اتخاذ شده در سه کشور ذکرشده، 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 ۱۹۹٤ 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[۸] (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 [۱٤، صص۳-۸]. 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 [۱٦، صص٤-٦]. برخی نیز از این مدل جهت محاسبه ارزش دوره عمر مشتری استفاده کرده‌اند [۱۷].

هسیه (۲۰۰٤) 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



Archive of SID
www.SID.ir

گروهی از مشتریان با استفاده از قوانین انجمنی اپریوری[۹] تعیین شد. این مطالعه نشان می‌دهد که شناسایی از راه یک مدل امتیازدهی رفتاری ویژگی‌های مفید مشتری را بیان می کند و تسهیل توسعه استراتژی بازاریابی را به همراه دارد. [١٦]

در سال های اخیر استفاده از ترکیبی از چند روش برای عملیات اعتبارسنجی استفاده شده از جمله این کارها در مقاله‌ای لوپز و همکارانش (۲۰۱۱) الگوریتمی را برای رتبه‌بندی اعتباری مشترکان برق برای دادن تعرفه‌های خاص و مختلف ارایه 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 ارائه شده به‌وسیله الگوریتم K-Means را حذف می‌کند و بر اساس آن به مطلوب جهانی نزدیک می‌شود. در این الگوریتم از ویژگی های H-ANN[۱۰] استفاده شده است. الگوریتم از لحاظ شاخص اعتبار با الگوریتم های دیگر از جمله الگوریتم های سلسله مراتبی(DM and DW) ، تغییرات به دنبال لیدر(F)، شبکه های عصبی مکرر هاپفیلد(H) ، K-means (K)، SOM–K-Means ، همواره با استفاده از توصیف داده‌ها از هر الگوریتم و هر شاخص ([۱۱]CH, [۱۲]DBI and [۱۳]MIA) مقایسه شده است. در تجزیه و تحلیل شاخص ها، نتایج به دست آمده توسط H-ANN-K بهتر از آن‌هایی است که به‌وسیله الگوریتم پایه ( Kو H) و KSOM به دست آمده و شبیه به آن‌هایی است که توسط DM، DW و الگوریتم های 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 رفتارش در بانک اطلاعات استفاده می‌شود [١٤، صص۲-۳]. 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



www.SID.ir

Archive of SID

آخرین خرید مشتری تا زمان حال، F، تعداد خریدها در یک دوره زمانی مشخص و M، مبلغ اسمی خریداری شده در دوره موردنظر 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 می باشد [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 ارتباط‌اند 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 [۱۴] استفاده می‌شود. در این روش نخست مراکز خوشه در نظرگرفته 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:

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



Archive of SID

www.SID.ir

### ۳-۳- روش امتیازدهی

ارزش هر مشتری را می‌توان بر اساس تازگی( تاخر)، تکرار و ارزش پولی به صورت زیر مشخص کرد:[۲۳، ص ١٦٤]

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

که $R(c_i)$ ، $F(c_i)$ و $M(c_i)$ به ترتیب امتیازات مشتری $c_i$ باتوجه به معیارهای R، F و M است. $W^R$، $W^F$ و $W^M$ اهمیت وزنی برای معیار های R ، F و M را به ترتیب نشان میدهد.
به‌علاوه داریم:

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

سود آوری خوشه $O^n$ با محاسبه میانگین ارزش همه مشتری‌های خوشه nام حاصل می‌شود. از این رو می‌توان آن را از طریق معادله زیر تعریف کرد [۲۳ ، ص ١٦٤]:

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

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

که $R(O^n)$ ، $F(O^n)$ و $M(O^n)$ امتیاز $n$امین خوشه باتوجه 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 به‌عنوان 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.



Archive of SID

www.SID.ir

## ۴- مدلسازی و تحلیل داده ها

روند اجرای عملیات در این تحقیق در شکل ۱ نشان داه 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 داده-های 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 هریک 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 نورون 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.

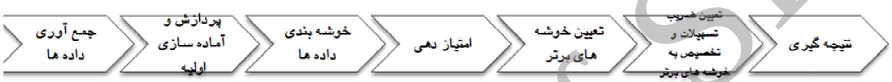


Archive of SID

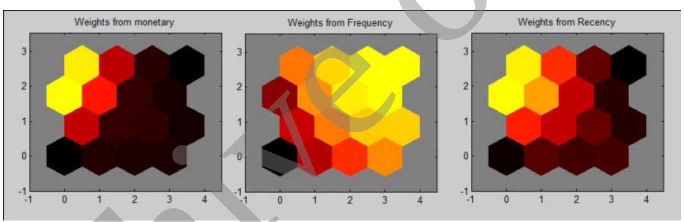
www.SID.ir

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

مدل امتیازدهی رفتاری در این مطالعه در نخست با به‌کارگیری الگوریتم SOM و شبکه‌ای در ابعاد ٤×٤ 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، ص۲-۷].




**شکل ۱** روند اجرای عملیات تحقیق




**شکل ۲** بردارهای ورودی

انجام الگوریتم خوشه‌بندی به این علت است تا بتوان گروه‌هایی از مشتریان را برای رتبه-بندی و اعطای تسهیلات شناسایی کرد. نتایج اجرای الگوریتم SOM در نرم‌افزار متلب در شکل‌های ۳ و ٤ نمایش داده شده است. در شکل ۳ تعداد مشتریانی را که در هر نورون تقسیم شده‌اند، نمایش می دهد . شکل ٤ نشان دهنده میزان فاصله بین نرون‌های همسایه از یکدیگر می‌باشد که هرچقدر میزان فاصله آن‌ها از یکدیگر بیشتر باشد، با رنگ تیره‌تر نشان داده می-شود و هر چقدر این فاصله کمتر باشد، با رنگ‌های روشن‌تر نمایش می‌دهد.

همانطور که در شکل ٤ مشاهده می‌شود، نرون‌هایی را که فاصله نزدیکی به هم دارند، می



Archive of SID

www.SID.ir

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

**جدول ۲** نرخ پارامتر SSE به ازای تعداد خوشه‌ها در SOM

| SSE | تعداد خوشه ها |
|---|---|
| ۱۷/۷۰۹۶۹ | ۸ |
| ۱۵/۷۵۸۳۱ | ۹ |
| ۱۴/۶۱۳۸۹ | ۱۰ |
| ۱۸/۶۱۰۴۹ | ۱۱ |
| ۱۹/۵۰۴۸۷ | ۱۲ |

**جدول ۳** پراکندگی خوشه‌ها

| | | | |
|---|---|---|---|
| ۹ | ۱۰ | ۱۰ | ۱۰ |
| ۷ | ۸ | ۱۰ | ۱۰ |
| ۵ | ۶ | ۴ | ۱۰ |
| ۱ | ۲ | ۳ | ۴ |

## ۴-۳. امتیاز دهی خوشه

بعد از تعیین خوشه‌ها حالا می‌خواهیم ارزش هر خوشه را به‌دست 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 مقدار۰/۳ 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



Archive of SID

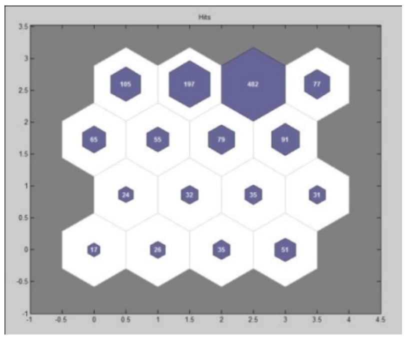
www.SID.ir

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 می شود.

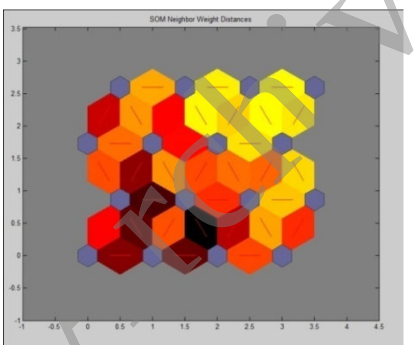
خوشه ٥. مشتریانی هستند که میزان گردش آنها بالای ۲۵ میلیون می باشد که معمولا بازه به‌نسبت عالی دارند. تعداد ماه که تراکنش دارند، بین ۹ تا ۱۲ بار می باشد و تعداد مانده منفی زیر ٥ ماه دارند به این معنا است که حضور عالی در بازار دارند. حجم تراکنش در ماه آن‌ها نسبتاً خوب می‌باشد و تسویه خوبی دارند. تعداد مشتریان در این خوشه ۲٤ نفر می‌باشد.

خوشه ۱. مشتریانی هستند که متوسط میزان گردش آنها از ٥۰ میلیون به بالا است، تعداد ماه که تراکنش دارند بین ۱۰ تا ۱۲ می باشد که بسیار عالی است ولی تعداد مانده منفی بین ٦ 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 می باشد.

خوشه ۷. مشتریانی هستند که میزان گردش آنها بین ۲٥ میلیون تا ۷٥۰ میلیون می‌باشد که بازه به‌نسبت عالی دارند. تعداد ماه که تراکنش دارند بین ۸ تا ۱۲ بار می‌باشد و تعداد مانده منفی زیر ٥ ماه دارند؛ به این معنا که حضور عالی در بازار دارند. حجم تراکنش در ماه آن‌ها به‌نسبت خوب است و تسویه خوبی دارند. تعداد مشتریان در این خوشه ٦٥ نفر می باشد.




**شکل ۳** تعداد اعضای هر نرون




**شکل ۴** میزان فاصله نرون‌ها از یکدیگر



Archive of SID
www.SID.ir

سه خوشه‌ای که در رتبه های 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

| خوشه | میانگین ارزش پول | میانگین ارزش فرکانس | میانگین ارزش تاخر | ارزش خوشه | تعداد اعضای خوشه |
|---|---|---|---|---|---|
| ٥ | ٠/٦٧٣٩٣ | ٠/٩٤٠٩٧ | ٠/٧٥٣٩٤ | ٠/٧٧٠٦٤ | ٢٤ |
| ١ | ٠/٦٨١٣٧ | ٠/٩٦٠٧٨ | ٠/٣٣٨٢٤ | ٠/٦٩٦٥٧ | ١٧ |
| ٧ | ٠/٣٧٦٨١ | ٠/٧٨٣٣٣ | ٠/٨٠٢٥٦ | ٠/٥٨٣٩٢ | ٦٥ |
| ٨ | ٠/٢٥٦٠٦ | ٠/٥٦٦٦٧ | ٠/٩٤٣٩٤ | ٠/٤٨٦٨٢ | ٥٥ |
| ٦ | ٠/١٨٨٢١ | ٠/٤٥٥٧٣ | ٠/٧٤٤٧٩ | ٠/٣٧٩٧٨ | ٣٢ |
| ٢ | ٠/٢٦١٠٧ | ٠/٥٣٥٢٦ | ٠/٤٤٢٣١ | ٠/٣٧٩٥٨ | ٢٦ |
| ٩ | ٠/١٤٦٣٩ | ٠/٣٦٣٤٩ | ٠/٩٥٧١٤ | ٠/٣٧٣٦٧ | ١٠٥ |
| ١٠ | ٠/٠٣٩٠٩ | ٠/١٠٦١٥ | ٠/٩٧٧٦٢ | ٠/٢٤٦٩١ | ٩٥٧ |
| ٤ | ٠/٠٤٨٥٤ | ٠/١٤٩٢٣ | ٠/٥٠١٩٤ | ٠/١٦٩٤٢ | ٨٦ |
| ٣ | ٠/٠١١٢٦ | ٠/٠٣٥٧١ | ٠/١٠٢٣٨ | ٠/٠٣٦٨٢ | ٣٥ |

### ۴-۴. امتیازدهی اعضای خوشه‌های برتر برای ارایه 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 می باشد 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



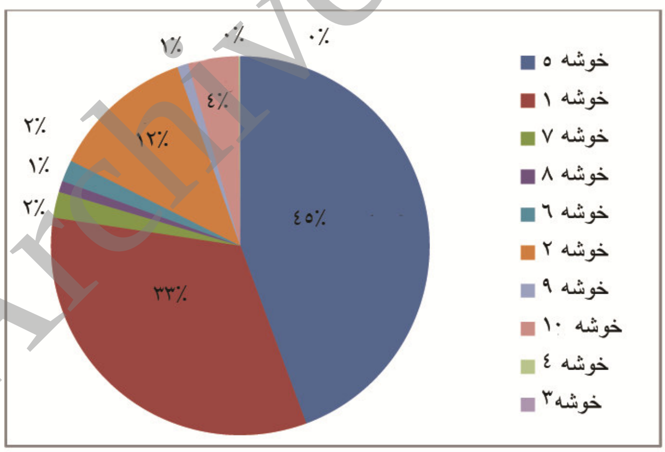
Archive of SID

www.SID.ir

متغیر تناوب با ضریب ۰/۳ و متغیر تأخر با وزن ۰/۲ در فرمول اعمال می‌شوند.

تعداد ١٠٦ 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 ١٠٦ 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 ٥ 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 ٥ به ترتیب برای خوشه‌های ٥، ١و ٧ از مقدار تسهیلات در نظر گرفته شده میزان ٢٧/١ ، ١٧/٣ و ٥٥/٦ درصد درنظر گرفته می شود. خوشه ٥ که برترین خوشه است میزان کمتری تسهیلات به نسبت خوشه ٧ که در رتبه سوم قرار دارد اختصاص پیدا می‌کند و دلیل این می باشد که تعداد اعضای خوشه ٧ بیشتر از خوشه ٥ است. در نتیجه این مقدار تسهیلات به نسبت کمتری به اعضای خوشه ٧ تقسیم می‌شود، از این رو تعداد اعضای خوشه‌ای که بالاترین رتبه را دارند به‌نسبت میزان بیشتری از تسهیلات به آنها اختصاص می‌یابد




**شکل ۵** درصد گردش مالی



Archive of SID

www.SID.ir

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

| خوشه | ضریب تسهیلات |
|---|---|
| ٥ | ۰/۲۷۱ |
| ۱ | ۰/۱۷۳ |
| ۷ | ۰/٥٥٦ |

برای اینکه محققین بتوانند مدل خود برای اعطای اعتبار را کامل کنند، از ضریب (وزن) خوشه تولید شده در مرحله قبل استفاده می‌کنند. روند به این گونه است که برای هر خوشه نخست مجموع ارزش‌های تمامی مشتریان آن‌ها محاسبه می شود، سپس برای هر فرد ارزش مشتری آن تقسیم بر مقدار مجموع به‌دست آمده می‌شود، با این کار ضریب اولیه اعطای تسهیلات برای هر مشتری تعین می‌شود از ضرب وزن آن خوشه در وزن هر مشتری، ضریب اعطای تسهیلات مشتری تعیین می‌شود. نتایج مربوطه در جدول ٦ و ۷ و ۸ نشان داده شده است.

همانطور که در جدول نشان داده شد، ضرایب مربوط به اعضای خوشه‌های 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 می‌تواند در هر دوره به مشتریان خود اختصاص دهدریال مبلغی بالغ بر ۱۰۰میلیارد می‌باشد.

**جدول ۶** ضریب و مقدار تسهیلات برای خوشه ٥ با روش وزن‌دهی خطی

| شماره مشتری | ضریب ارزش مشتری | ضریب تسهیلات مشتری | مقدار تسهیلات مشتری | شماره مشتری | ضریب ارزش مشتری | ضریب تسهیلات مشتری | مقدار تسهیلات مشتری |
|---|---|---|---|---|---|---|---|
| c۱ | ۰/۰٤٦٤۸۹۷ | ۰/۰۱۲٥۹۰۸ | ۱۲٥۹۰۸۲٥۸۹ | c۱۳ | ۰/۰٤۱٦٥٦٤۲۷ | ۰/۰۱۱۲۸۱۸۲٤ | ۱۱۲۸۱۸۲۳۷۳ |
| C۲ | ۰/۰۳۸۲۹٦۸ | ۰/۰۱۰۳۷۲۱۸ | ۱۰۳۷۲۱۷۸۱٦ | c۱٤ | ۰/۰۳۹٦۰۸٤۲۱ | ۰/۰۱۰۷۲۷۱٦۲ | ۱۰۷۲۷۱٦۷۹ |
| C۳ | ۰/۰٤۱٦۱٥٤ | ۰/۰۱۱۲۷۰۷۳ | ۱۱۲۷۰۷۳۰٤۹ | c۱٥ | ۰/۰٥۰۲٥۸۰٤۹ | ۰/۰۱۳٦۱۱٤۰٤ | ۱۳٦۱۱٤۰۳۸٥ |



Archive of SID

www.SID.ir

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 |
|---|---|---|---|---|---|---|---|
| C٤ | ٠/٠٤١٠٨٢٩٨٥ | ٠/٠١١١٢٦٥١٨ | ١١١٢٦٥١٨٣٩ | C١٦ | ٠/٠٤٣٧٨٦٣٥٢ | ٠/٠١١٨٥٨٦٧٢ | ١١٨٥٨٦٧٢١٤ |
| C٥ | ٠/٠٣٦٧٨٢١٧٤ | ٠/٠٠٩٩٦١٧٢٨ | ٩٩٦١٧٢٨٣٢ | C١٧ | ٠/٠٥١٣٦٣٩٧١ | ٠/٠١٣٩١٠٩٢١ | ١٣٩١٠٩٢١٢٩ |
| C٦ | ٠/٠٣٨٤٢٠٥٧٨ | ٠/٠١٠٤٠٥٤٥٨ | ١٠٠٥٤٥٧٨٧ | C١٨ | ٠/٠٣٩٩٧٧٠٦٢ | ٠/٠١٠٨٢٧٠٠١ | ٠١٠٨٢٧٠٠٩٤ |
| C٧ | ٠/٠٣٨٨٣٠١٧٩ | ٠/٠١٠٥١٦٣٩ | ١٠٥١٦٣٩٠٢٦ | C١٩ | ٠/٠٣٦٤١٣٥٣٣ | ٠/٠٠٩٨٦١٨٨٩ | ٩٨٦١٨٨٩١٧/٩ |
| C٨ | ٠/٠٣٦٠٠٣٩٣٢ | ٠/٠٠٩٧٥٠٩٥٧ | ٩٧٥٠٩٥٦٧٩/٢ | C٢٠ | ٠/٠٣٦٧٤١٢١٤ | ٠/٠٠٩٩٥٠٦٣٥ | ٩٥٥٠٦٣٥٠٨/٨ |
| C٩ | ٠/٠٤٢٥٥٧٥٤٩ | ٠/٠١١٥٢٥٨٧٥ | ١١٥٢٥٨٧٤٩٨ | C٢١ | ٠/٠٤١٣٢٨٧٤٦ | ٠/٠١١١٩٣٠٧٨ | ١١١٩٣٠٧٧٨٢ |
| C١٠ | ٠/٠٤٦٤٤٨٧٥٩ | ٠/٠١٢٥٧٩٧٣٣ | ١٢٥٧٩٣٢٦٥ | C٢٢ | ٠/٠٤٤٥٦٤٥٩٤ | ٠/٠١٢٠٦٩٤٤٤ | ١٢٠٦٩٤٤٣٦٧ |
| C١١ | ٠/٠٤٢٤٧٥٦٢٩ | ٠/٠١١٥٠٣٦٨٩ | ١١٥٠٣٦٨٨٥٠ | C٢٣ | ٠/٠٤٢٩٢٦١٩ | ٠/٠١١٦٢٥٧١٤ | ١١٦٢٥٧١٤١٣ |
| C١٢ | ٠/٠٤٤٠٧٣٠٧٣ | ٠/٠١١٩٣٦٣٢٥ | ١١٩٣٦٣٢٤٨١ | C٢٤ | ٠/٠٣٨٢٩٧٦٨ | ٠/٠١٠٣٧٢١٧٨ | ١٠٣٧٢١٧٨١٦ |

**جدول۷** 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 |
|---|---|---|---|---|---|---|---|
| C٢٥ | ٠/٠٥٥٠١٨٨٧٣ | ٠/٠٠٩٥٤٠١٨٥ | ٩٥٤٠١٨٥٢٥/٧ | C٣٤ | ٠/٠٦٥١٩٠٩٦٧ | ٠/٠١١٣٠٤٠١ | ١١٣٠٤٠١٠٢١ |
| C٢٦ | ٠/٠٥٠٥٤٠٥٩٢ | ٠/٠٠٨٦٦٣٦٥٩ | ٨٧٦٣٦٥٨٥٥ | C٣٥ | ٠/٠٥١٦٩٢١٥ | ٠/٠٠٨٩٦٣٣٣٧ | 896333684.6 |
| C٢٧ | ٠/٠٦٠٢٠٠٨٨٣ | ٠/٠١٠٤٣٨٧٣٨ | ١٠٤٣٨٧٣٧٥٩ | C٣٦ | ٠/٠٦١٠٣٢٥٦٣ | ٠/٠١٠٥٨٢٩٥ | ١٠٥٨٢٩٤٩٦٩ |
| C٢٨ | ٠/٠٦٥٧٠٢٧٧ | ٠/٠١١٣٩٢٧٥٦ | ١١٣٩٢٧٥٦١٢ | C٣٧ | ٠/٠٧٦٠٠٢٨١٥ | ٠/٠١٣١٧٨٧٦٧ | ١٣١٧٨٦٧٥٤ |
| C٢٩ | ٠/٠٥٥٤٠٢٧٢٥ | ٠/٠٠٩٦٠٦٧٤٥ | ٩٦٠٦٧٤٤٦٨/٩ | C٣٨ | ٠/٠٦٣٥٢٧٦٠٥ | ٠/٠١١٠١٥٥٨٦ | ١١٠١٥٥٨٦٠٠ |
| C٣٠ | ٠/٠٦١١٠٥١٤ | ٠/٠١٠٦٠٥١٣٦ | ١٠٦٠٥١٣٦١٧ | C٣٩ | ٠/٠٥٣٦١١٤١٣ | ٠/٠٠٩٢٩٦١٣٤ | ٩٢٩٦١٣٤٠٠/٦ |
| C٣١ | ٠/٠٦٥١٩٠٩٦٧ | ٠/٠١١٣٠٤٠١ | ١١٣٠٤٠١٠٢١ | C٤٠ | ٠/٠٦٤٦٧٩١٦٣ | ٠/٠١١٢١٥٢٦٤ | ١١٢١٥٢٦٤٣٠ |
| C٣٢ | ٠/٠٥٥١٤٦٨٢٤ | ٠/٠٠٩٥٦٢٣٧٢ | ٩٥٦٢٣٧١٧٣/٤ | C٤١ | ٠/٠٤٩٥١٦٩٨٥ | ٠/٠٠٨٥٨٦١٦٧ | ٨٥٨٦١٦٦٧٣/١ |
| C٣٣ | ٠/٠٤٦٣٨٢١٨٩ | ٠/٠٠٨٠٤٢٥٩٨ | ٨٠٤٢٥٩٨٠٣/٦ | | | | |

Archive of SID



www.SID.ir

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

| شماره مشتری | ضریب ارزش مشتری | ضریب تسهیلات مشتری | شماره مشتری | ضریب ارزش مشتری | مقدار تسهیلات مشتری | شماره مشتری | ضریب تسهیلات مشتری | مقدار تسهیلات مشتری |
|---|---|---|---|---|---|---|---|---|
| C٤٢ | ٠/٠١٥٢٤٩٥٠١ | ٠/٠٠٨٤٧٥٢٣٤ | C٦٤ | ٠/٠١٣٧٥٢٤٩٥ | ٠/٠٠٧٦٤٣٢٤١ | C٨٦ | ٠/٠١٤٧١٠٥٧٩ | ٠/٠٠٨١٧٥٧١٧ |
| C٤٣ | ٠/٠١٧٨٤٤٣١١ | ٠/٠٠٩٩١٧٣٥٥ | C٦٥ | ٠/٠١٤٩٣٠١٤ | ٠/٠٠٨٢٩٧٧٤٣ | C٨٧ | ٠/٠١٤٩٥٠١ | ٠/٠٠٨٣٠٨٨٣٦ |
| C٤٤ | ٠/٠١٥٢٠٩٥٨١ | ٠/٠٠٨٤٥٣٠٤٨ | C٦٦ | ٠/٠١٥٥٨٨٨٢٢ | ٠/٠٠٨٦٦٣٨١٩ | C٨٨ | ٠/٠١٤١٥١٦٩٧ | ٠/٠٠٧٨٦٥١٠٦ |
| C٤٥ | ٠/٠١٤٢١١٥٧٧ | ٠/٠٠٧٨٩٨٣٨٦ | C٦٧ | ٠/٠١٣٨٣٢٣٣٥ | ٠/٠٠٧٦٨٧٦١٤ | C٨٩ | ٠/٠١٦٣٨٧٢٢٦ | ٠/٠٠٩١٠٧٥٤٩ |
| C٤٦ | ٠/٠١٥١٨٩٦٢١ | ٠/٠٠٨٤٤١٩٥٥ | C٦٨ | ٠/٠١٧٢٤٥٥٠٩ | ٠/٠٠٩٥٨٤٥٥٨ | C٩٠ | ٠/٠١٢٥٩٤٨١ | ٠/٠٠٦٩٩٩٨٣٤ |
| C٤٧ | ٠/٠١٥١٦٩٦٦١ | ٠/٠٠٨٤٣٠٨٦١ | C٦٩ | ٠/٠١٤٦٥٠٦٩٩ | ٠/٠٠٨١٤٢٤٣٧ | C٩١ | ٠/٠١٥١٠٩٧٨ | ٠/٠٠٨٣٩٧٥٨٢ |
| C٤٨ | ٠/٠١٦٧٠٦٥٨٧ | ٠/٠٠٩٢٨٥٠٤١ | C٧٠ | ٠/٠١٥٢٠٩٥٨١ | ٠/٠٠٨٤٥٣٠٤٨ | C٩٢ | ٠/٠١٦٩٢٦١٤٨ | ٠/٠٠٩٤٠٧٠٦٦ |
| C٤٩ | ٠/٠١٦٣٢٧٣٤٥ | ٠/٠٠٩٠٧٤٢٦٩ | C٧١ | ٠/٠١٥٠٤٩٩ | ٠/٠٠٨٣٦٤٣٠٢ | C٩٣ | ٠/٠١٣٩٣٢١٣٦ | ٠/٠٠٧٧٤٣٠٨١ |
| C٥٠ | ٠/٠١٥٥٦٨٨٦٢ | ٠/٠٠٨٦٥٢٧٢٦ | C٧٢ | ٠/٠١٣٧٥٢٤٩٥ | ٠/٠٠٧٦٤٣٢٤١ | C٩٤ | ٠/٠١٨٩٦٢٠٧٦ | ٠/٠١٠٥٣٨٥٧٧ |
| C٥١ | ٠/٠١٤٨١٠٣٧٩ | ٠/٠٠٨٢٣١١٨٣ | C٧٣ | ٠/٠١٤٦٥٠٦٩٩ | ٠/٠٠٨١٤٢٤٣٧ | C٩٥ | ٠/٠١٦٧٠٦٥٨٧ | ٠/٠٠٩٢٨٥٠٤١ |
| C٥٢ | ٠/٠١٧٣٢٥٣٤٩ | ٠/٠٠٩٦٢٨٩٣١ | C٧٤ | ٠/٠١٤٧٣٠٥٣٩ | ٠/٠٠٨١٨٦٨١ | C٩٦ | ٠/٠١٤٤٧١٠٥٨ | ٠/٠٠٩٢٨٥٠٤١ |
| C٥٣ | ٠/٠١٣٨٧٢٢٥٥ | ٠/٠٠٧٧٠٩٨٠١ | C٧٥ | ٠/٠١٧١٨٥٦٢٩ | ٠/٠٠٩٥٥١٢٧٨ | C٩٧ | ٠/٠١٧٤٠٥١٩ | ٠/٠٠٩٦٧٣٣٠٤ |
| C٥٤ | 0.017145709 | ٠/٠٠٩٥٢٩٠٩٢ | C٧٦ | ٠/٠١٤٧٥٠٤٩٩ | ٠/٠٠٨١٩٧٩٠٣ | C٩٨ | ٠/٠١٤٤٥١٠٩٨ | ٠/٠٠٨٠٣١٥٠٥ |
| C٥٥ | ٠/٠١٦٣٢٧٣٤٥ | ٠/٠٠٩٠٧٤٢٦٩ | C٧٧ | ٠/٠١٢٠٧٥٨٤٨ | ٠/٠٠٦٧١١٤٠٩ | C٩٩ | ٠/٠١٧٢٨٥٤٢٩ | ٠/٠٠٩٦٠٦٧٤٥ |
| C٥٦ | ٠/٠١٢٦١٤٧٧ | ٠/٠٠٧٠١٠٩٢٧ | C٧٨ | ٠/٠١٥٣٤٩٣٠١ | ٠/٠٠٨٥٣٠٧٠١ | C١٠٠ | ٠/٠١٥٢٦٩٤٦١ | ٠/٠٠٨٤٨٦٣٢٨ |
| C٥٧ | ٠/٠١٦٨٦٦٢٦٧ | ٠/٠٠٩٣٧٣٧٨٧ | C٧٩ | ٠/٠١٣٥٣٢٩٣٤ | ٠/٠٠٧٥٢١٢١٦ | C١٠١ | ٠/٠١٥٣٢٩٣٤١ | ٠/٠٠٨٥١٩٦٠٧ |
| c58 | ٠/٠١٥٠٤٩٩ | ٠/٠٠٨٣٦٤٣٠٢ | C٨٠ | ٠/٠١٦٦٤٦٧٠٧ | ٠/٠٠٩٢٥١٧٦١ | C١٠٢ | ٠/٠١٤٨٣٠٣٣٩ | ٠/٠٠٨٢٤٢٢٧٦ |
| C٥٩ | ٠/٠١٤٩٥٠١ | ٠/٠٠٨٣٦٤٣٠٢ | C٨١ | ٠/٠١٤٠١١٩٧٦ | ٠/٠٠٧٧٨٧٤٥٤ | C١٠٣ | ٠/٠١٦٤٨٧٠٢٦ | ٠/٠٠٩١٦٣٠١٥ |
| C٦٠ | ٠/٠١٦٨٤٦٣٠٧ | ٠/٠٠٩٣٦٢٦٩٣ | C٨٢ | ٠/٠١٧١٢٥٧٤٩ | ٠/٠٠٩٥١٧٩٩٩ | C١٠٤ | ٠/٠١٥٦٤٨٧٠٣ | ٠/٠٠٨٦٩٧٠٩٩ |
| C٦١ | ٠/٠١٥٣٠٩٣٨١ | ٠/٠٠٨٥٠٨٥١٤ | C٨٣ | ٠/٠١٥٤٤٩١٠٢ | ٠/٠٠٨٥٨٦١٦٧ | C١٠٥ | ٠/٠١٧٠٦٥٨٦٨ | ٠/٠٠٩٤٨٤٧١٩ |
| C٦٢ | ٠/٠١٣٧٧٢٤٥٥ | ٠/٠٠٧٦٥٤٣٣٥ | C٨٤ | ٠/٠١٤٦٥٠٦٩٩ | ٠/٠٠٨١٤٢٤٣٧ | C١٠٦ | ٠/٠١٤٢٣١٥٣٧ | ٠/٠٠٧٩٠٩٤٧٩ |
| C٦٣ | ٠/٠١٧٤٢٥١٥ | ٠/٠٠٩٦٨٤٣٩٧ | C٨٥ | ٠/٠١٥١٢٩٧٤١ | ٠/٠٠٨٤٠٨٦٧٥ | | | |



Archive of SID

www.SID.ir

## ۵- نتیجه‌گیری

از آنچه گفته شده می‌توان اذعان 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 روش های امتیازدهی امکان ارائه رتبه‌بندی رفتاری مشتریان برای بیان برترین بخش از مشتریان را خواهیم داشت. تا درنهایت 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 ۰/۵٥٦ 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 (۲۰۰٤) 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



Archive of SID

www.SID.ir

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 تراکنش-های مشتریان که نتایج حاصل از الگوریتم‌ها را بی‌معنا نشان می داد 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 رتبه-بندی مانند روش‌های AHP و ANP و استفاده از شاخص بیرونی و محیطی سازمان متقاضی وام، مدل دیگری را ارایه 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,



Archive of SID

www.SID.ir

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



Archive of SID

www.SID.ir

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*



Archive of SID

www.SID.ir

*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. Tsiptsis, 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.



Archive of SID

www.SID.ir

[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.

Archive of SID



www.SID.ir