

بخش‌بندی مراجعین کتابخانه عمومی بر اساس ارزش دوره عمر آن‌ها و مدل RFM با ترکیب روش‌های تصمیم‌گیری چندمعیاره و داده‌کاوی

عظیمه مظفری^۱، ثریا ضیائی^۲، افسانه مظفری^۳

چکیده

هدف: با توجه به اهمیت مراجعین برای کتابخانه‌های عمومی، این پژوهش به بخش‌بندی و شناسایی مراجعین کتابخانه عمومی شهید دستغیب شیراز بر اساس ارزش دوره عمر آن‌ها و مدل RFM می‌پردازد.

روش پژوهش: برای پیاده‌سازی روش پیشنهادی ابتدا مقادیر شاخص‌های مدل RFM تعیین شدند؛ پس از وزن‌دهی به این سه شاخص با استفاده از روش تحلیل سلسله مراتبی، مراجعین با استفاده از شبکه عصبی خودسازمان‌ده بخش‌بندی شدند؛ سپس با رسم هرم ارزش دوره عمر خوشه‌های مراجعین کلیدی و باارزش شناسایی شدند.

یافته‌ها: از بین سه شاخص مدل RFM، شاخص تعداد مراجعه دارای بیشترین اهمیت و شاخص سابقه عضویت دارای کمترین اهمیت در تعیین ارزش مراجعین بوده و بر این اساس مراجعین به پنج خوشه تقسیم شدند. طبق هرم ارزش دوره عمر، باارزش‌ترین مراجعین در خوشه چهار قرار گرفته‌اند که تنها ۳ درصد از کل مراجعین را تشکیل می‌دهند.

نتیجه‌گیری: با شناسایی مراجعین باارزش و کلیدی، پیشنهادهایی جهت بهبود امر خدمت‌رسانی به مراجعین با در نظر گرفتن ارزشی که برای کتابخانه دارند، ارائه گردید.

کلیدواژه‌ها: ارزش دوره عمر مراجعین، تحلیل سلسله مراتبی، شبکه خودسازمان‌ده، کتابخانه عمومی شهید دستغیب شیراز.

۱. کارشناس ارشد مهندسی صنایع، دانشگاه علم و صنعت ایران (نویسنده مسئول)، azime.mozafari@yahoo.com

۲. دکترای کتابداری و اطلاع‌رسانی، عضو هیات علمی دانشگاه پیام نور مشهد

۳. کارشناس ارشد کتابداری و اطلاع‌رسانی، کتابدار کتابخانه عمومی کوثر نور شهرستان نیریز

مقدمه

کتابخانه عمومی^۱ یکی از اثربخش‌ترین ابزارها در راستای توسعه فرهنگی و اعتلای سطح آگاهی افراد جامعه با رسالت گردآوری، سازمان‌دهی و اشاعه اطلاعات است. کتابخانه عمومی به تمام اقشار جامعه خدمات می‌دهد و وظیفه اصلی آن، فراهم آوردن مواد و منابع اطلاعاتی مورد نیاز افراد و گروه‌های مختلف جامعه در اشکال مختلف چاپی و الکترونیکی است. در دنیای رقابتی امروز، اطلاعات به عنوان یکی از فاکتورهای تولیدی مهم پدیدار می‌شود و تلاش برای استخراج اطلاعات از داده‌ها توجه بسیاری از افراد دخیل در صنعت اطلاعات و حوزه‌های وابسته را به خود جلب نموده است. به کارگیری تکنولوژی اطلاعات توسط هر سازمان در عمل یک فرایند هموار نیست و این موضوع در مورد کتابخانه عمومی نیز صدق می‌کند اما، تجربه نشان می‌دهد که یک برنامه نظام‌مند می‌تواند ظهور و نگهداری تکنولوژی اطلاعات در محیط کتابخانه را تسهیل کند (پارکر^۲، ۱۹۹۰)، حتی کاربرد تکنولوژی‌های پیشرفته پردازش اطلاعات مثل سیستم‌های خبره و سیستم اطلاعات جغرافیایی در کتابخانه گزارش شده است (پاتکر^۳، ۲۰۰۵؛ باری^۴، ۲۰۰۳). امروزه با توجه به ورود و گسترش فناوری اطلاعات، نیازهای مراجعین کتابخانه‌های عمومی و الگوهای رفتاری آنان نیز به طور قابل ملاحظه‌ای تغییر کرده است (بیگدلی و فرهادیان، ۱۳۸۱)، لذا کتابخانه‌های عمومی به منظور ارضای نیازهای گوناگون مراجعین، به تطبیق و بکارگیری استراتژی‌های متنوع جذب و نگهداری آن‌ها گرایش پیدا کرده‌اند (بیگدلی و فرهادیان، ۱۳۸۱).

رویکردهای متعددی برای گردآوری، ذخیره، سازمان‌دهی و مدیریت کارآمد داده‌های موجود و رسیدن به نتایج معنی‌دار به کار گرفته شده‌اند که داده‌کاوی^۵ یکی از رویکردهای اخیر در این زمینه است. داده‌کاوی مجموعه‌ای از روش‌هایی است که به شخص امکان می‌دهد تا ورای داده‌پردازی معمولی حرکت کند و به استخراج اطلاعاتی که در انبوه داده‌ها مخفی و یا پنهان است، کمک می‌کند. داده‌کاوی همچنین می‌تواند به عنوان بخشی از فرایند بزرگتر کشف دانش در بانک‌های اطلاعاتی در محیط‌های مختلف در نظر گرفته شود (غضنفری، ۱۳۹۱). یکی از حوزه‌های مهم

^۱ Public Library

^۲ Parker

^۳ Patkar

^۴ Barry

^۵ Data Mining

داده‌کاوی، خوشه‌بندی^۱ است که یکی از پرکاربردترین روش‌های موجود در بحث پیش‌بینی رفتار است (آذر، احمدی و سبط، ۱۳۸۹). خوشه‌بندی یا بخش‌بندی شامل تقسیم کردن کلیه رکوردها به گروه‌های اصلی است که شباهت بیشتری به یکدیگر دارند (انگای، زو و چاو^۲، ۲۰۰۹).

شناسایی مراجعین مستلزم تحلیل مراجعین هدف و دسته‌بندی کردن آنها است که به یافتن گروه‌هایی از مراجعین کلیدی بر اساس ویژگی‌های آنها منجر می‌شود. به این ترتیب با استفاده از خوشه‌بندی می‌توان مراجعین با نیازهای مشابه را در گروه‌های یکسان دسته‌بندی نمود و ضمن شناسایی نیازهای مراجعین و اولویت‌بندی آنها در هر گروه، با برنامه‌ریزی صحیح از اتلاف وقت، نیروی انسانی و هزینه در کتابخانه‌ها جلوگیری نمود. همچنین توجه به نیازهای مراجعین و برآورده ساختن آنها منجر به افزایش رضایتمندی و وفاداری مراجعین و در نهایت افزایش آمار اعضای فعال کتابخانه خواهد شد.

هدف پژوهش

داده‌های موجود در بانک اطلاعاتی مراجعین کتابخانه‌های عمومی همواره در حال افزایش هستند، در حالی که در عمق این حجم داده‌ها، الگوها و روابط بسیار جالبی میان پارامترهای مختلف به صورت پنهان باقی می‌ماند که به شناسایی مراجعین، تعیین ارزش آنها برای کتابخانه و در نهایت تامین منابع موردنیاز گروه‌های مختلف مراجعین در راستای رضایتمندی و وفاداری آنها کمک می‌کند. با وجود محدودیت منابع، بودجه و زمان در کتابخانه‌های عمومی، اهمیت مراجعین فعال و باارزش برای این کتابخانه‌ها نمود بیشتری پیدا می‌کند. بنابراین ایجاد بستری برای افزایش رضایتمندی و وفاداری مراجعین از طریق ایجاد ارتباطات موثرتر، جذب و حفظ مراجعین ضروری است. این پژوهش درصدد است تا با استفاده از داده‌های موجود در بانک اطلاعاتی مراجعین که در نرم افزار کتابخانه عمومی قرار دارد، به بخش‌بندی مراجعین کتابخانه عمومی شهید دستغیب شیراز بپردازد. اما سؤال اصلی این است که چگونه می‌توان مراجعین کلیدی را شناسایی نمود؟ ارزش دوره عمر مفهومی است که می‌تواند در این راستا کمک فراوانی نماید. این مفهوم بیان‌گر ارزشی است که مراجعه‌کننده در طول دوره عمر و مراجعه خود برای کتابخانه ایجاد می‌کند.

¹ Clustering

² Ngai, Xiu and Chau

پرسش‌های پژوهش

با توضیحات ذکر شده این پژوهش درصدد پاسخ دادن به پرسش‌های زیر است:

پرسش اول: شاخص‌های RFM^۱ در بخش‌بندی مراجعین کتابخانه عمومی شهید دستغیب شیراز به چه صورت تعریف می‌شوند؟

پرسش دوم: اولویت و اهمیت شاخص‌های RFM در تعیین ارزش مراجعین کتابخانه عمومی شهید دستغیب شیراز چگونه است؟

پرسش سوم: مراجعین کتابخانه عمومی شهید دستغیب شیراز را می‌توان به چند بخش یا خوشه تقسیم نمود و مراجعین هر خوشه چه ویژگی‌هایی دارند؟

سوال چهارم: کدام مراجعین برای کتابخانه عمومی شهید دستغیب شیراز ارزش بیشتری دارند؟

پیشینه پژوهش

تحقیقات گوناگونی بر روی بخش‌بندی با استفاده از مدل RFM و محاسبه هرم ارزش دوره عمر در حوزه کسب و کار انجام شده است، برای نمونه هسیه^۲ (۲۰۰۴) در پژوهشی با ایجاد یکپارچگی بین متغیرهای رفتاری و متغیرهای امتیازدهی RFM، با کمک شبکه عصبی مصنوعی مشتریان بانک را بخش‌بندی نمود. نیآگاس، سربویهوک و کیتیسین^۳ (۲۰۰۶) با استفاده از الگوریتم شبکه خودسازمان‌ده^۴ ابتدا به تعیین تعداد بهینه خوشه‌های مشتریان و سپس با استفاده از الگوریتم کا-میانگین^۵ به خوشه‌بندی مشتریان پرداختند و در نهایت اقدام به توصیف ارتباط بین ویژگی‌های مشتریان و تولید دانش از خوشه‌های ایجاد شده نمودند. سهرابی و خانلری (۲۰۰۷) در مطالعه خود بر اساس مدل RFM به محاسبه ارزش دوره عمر مشتریان یک بانک خصوصی پرداخته و مشتریان را به هشت خوشه تفکیک نموده و ویژگی‌های آن‌ها را مورد تحلیل قرار دادند. هو و جینگ^۶ (۲۰۰۸) به بررسی قابلیت مدل RFM در بخش‌بندی مشتریان در شرکت‌های خدمات پس از فروش خودرو

^۱ Recency, Frequency, Monetary (RFM)

^۲ Hsieh

^۳ Niyagas, Srivihok and Kitisin

^۴ Self Organization Map (SOM)

^۵ K-Means

^۶ Hu & Jing

پرداختند، در این پژوهش وزن مربوط به هر مؤلفه مدل از طریق فرایند تحلیل سلسله مراتبی^۱ تعیین و مشتریان به هشت خوشه تفکیک شدند و پس از تحلیل ویژگی‌های مشتریان، ارزش دوره عمر آنها در هر خوشه تعیین گردید. وو، چانگ و لو^۲ (۲۰۰۹) با استفاده از مدل RFM به تحلیل ارزش مشتریان یک شرکت ساخت تجهیزات صنعتی پرداختند که طبق نتایج به دست آمده مشتریان بر اساس شاخص‌های RFM در شش گروه خوشه‌بندی شدند و ویژگی‌های مشتریان در قالب خوشه‌ها و با استفاده از تحلیل ارزش دوره عمر مورد تحلیل قرار گرفت. نامور، غلامیان و خاک‌آبی (۲۰۱۰) در پژوهش خود برای تحلیل از داده‌های RFM و محاسبه ارزش دوره عمر مشتریان استفاده نمودند. سیدحسینی، ملکی و غلامیان (۲۰۱۰) با استفاده از مدل توسعه یافته RFM به داده کاوی در پایگاه داده یک شرکت تأمین قطعات خودرو پرداختند. در این تحقیق، وزن‌های مربوط به مؤلفه‌های مدل از طریق روش تحلیل سلسله مراتبی تعیین گردید و سپس به خوشه بندی داده‌ها پرداخته شد. خواجه‌وند و تارخ (۲۰۱۱) در پژوهش خود برای تخمین ارزش آینده مشتری از مدل RFM، الگوریتم خوشه‌بندی کا-میانگین و محاسبه ارزش دوره عمر مشتری استفاده کردند. لی، دای و تسنگ^۳ (۲۰۱۱) نیز به تحلیل ویژگی‌های مشتریان یک کارخانه بافندگی بر مبنای مدل توسعه یافته RFM پرداختند که مشتریان به پنج خوشه تفکیک شدند. کفاش پور، توکلی و علیزاده زوارم (۱۳۹۱) نیز در پژوهش خود مشتریان شرکت بازرگانی طوس شرق را بر مبنای مدل RFM و ارزش دوره عمر مشتریان بخش‌بندی کردند که مشتریان در هشت خوشه اصلی گنجانده شدند.

کاناوو^۴ (۲۰۰۳) معتقد است که داده‌کاوی می‌تواند برای پاسخ دادن به یک سوال خاص مربوط به کتابخانه و نیز برای کشف روندهای عمومی که به تصمیم‌گیری کمک می‌کنند، استفاده شود. درک الگوی استفاده کلی مجلات الکترونیکی یا تحلیل درخواست‌های اعضا برای میکروفیلم‌ها طی ۵ سال گذشته همگی مثال‌هایی از کشف روندهای عمومی هستند. دامنه تحلیل اسنادی نیز می‌تواند با استفاده از داده‌کاوی گسترش داده شود. پاتکر (۲۰۰۵) در پژوهشی به تبیین مشخصه‌های اساسی فرایند داده‌کاوی و کشف کاربردهای ممکن آن در کتابخانه و مؤسسات دانشگاهی می‌پردازد. با

¹ Analytical Hierarchy Process

² Wu, Chang and Lo

³ Li, Dai and Tseng

⁴ Cannavo

وجود اهمیت و دقت روش‌های نام برده و کاربردهای فراوانی که در حوزه‌های مختلف دارند، تاکنون پژوهشی در راستای بخش‌بندی مراجعین کتابخانه‌ها با استفاده از روش‌های داده‌کاوی و مدل RFM انجام نشده که در این پژوهش سعی شده است تا در یک چارچوب پیشنهادی و منطقی از این روش استفاده شود.

تعاریف مفهومی و عملیاتی

خوشه‌بندی. یکی از کاربردهای داده‌کاوی است که به معنای تقسیم‌بندی یک مجموعه داده به خوشه‌ها است، به نحوی که شباهت اعضای یک خوشه به همدیگر بیشتر از شباهت آن‌ها به اعضای خوشه‌های دیگر باشد (کیم و سون^۱، ۲۰۰۴؛ گوها، راستوگی و شیم^۲، ۱۹۹۸). نتیجه کلی از تحلیل خوشه‌بندی، توصیف کلی از هر خوشه است که برای تحلیل عمیق‌تری از خواص مجموعه داده‌ها لازم است (لی و چن^۳، ۲۰۰۵). با استفاده از خوشه‌بندی می‌توان مراجعین کتابخانه‌های عمومی را به بخش‌های جداگانه‌ای تفکیک نمود و با در نظر گرفتن اینکه مراجعین هر بخش نیازها و رفتارهای مشابهی دارند، به بررسی آن‌ها پرداخت.

داده‌کاوی. رویکرد ارزشمندی است که از آن به طور گسترده‌ای جهت استخراج اطلاعات، جستجوی روابط و الگوها در حجم عظیم داده‌ها استفاده شده است. داده‌کاوی فرایند کشف رابطه‌ها، الگوها و روندهای معناداری است که به بررسی حجم عظیمی از اطلاعات ذخیره شده در پایگاه داده‌ها با فناوری تشخیص الگو می‌پردازد (کانترازیک^۴، ۲۰۰۳). روش‌های داده‌کاوی از ترکیب و تجمیع روش‌های علوم مختلفی ایجاد می‌شوند که با استفاده از قابلیت‌های مختلفی که در هر یک از این علوم وجود دارد سعی کرده تا محدودیت‌هایی که وجود دارند را تا حد ممکن برطرف سازند (هان و کمبر^۵، ۲۰۰۶).

¹ Kim & Sohn

² Guha, Rastogi and Shim

³ Lee & Chen

⁴ Kantardzic

⁵ Han & Kamber

روش تحلیل سلسله مراتبی (AHP). یکی از روش‌های رتبه‌بندی یا وزن‌دهی است که بر اساس مقایسات زوجی^۱ بنا شده است. این روش وزن هر یک از فاکتورها را در راستای گزینه‌های رقیب مشخص می‌سازد و در نهایت یک الگوریتم ریاضی به گونه‌ای ماتریس‌های حاصل از مقایسات زوجی را با همدیگر تلفیق می‌سازد که تصمیم بهینه حاصل شود (مهرگان، ۱۳۸۳). در پایان نیز به تعیین نرخ سازگاری^۲ پرداخته می‌شود. با توجه به اینکه رسیدن به سازگاری کامل در دنیای واقعی مشکل است، درصدی از ناسازگاری را به صورت تجربی قابل قبول دانسته و می‌پذیریم و اگر نرخ سازگاری محاسبه شده بیشتر از ۰/۱ باشد قضاوت‌ها را متضاد و تصادفی در نظر می‌گیریم که باید در آنها تجدیدنظر به عمل آید.

شبکه‌های عصبی خودسازمان‌ده. از جمله مهم‌ترین و قدرتمندترین شبکه‌های موجود جهت داده‌کاوی هستند (کوهونن^۳، ۲۰۰۱) که به منظور تحلیل فضاهای پیچیده داده‌ها مورد استفاده قرار می‌گیرند (سئو و ابرمایر^۴، ۲۰۰۴؛ پن^۵، ۲۰۰۵؛ کوهونن و ماکیسارا^۶، ۱۹۸۴). اساس فلسفه شبکه‌های خودسازمان‌ده، نگاشت فضاهای با تعداد ابعاد بالا به فضایی دو یا سه بعدی است و در حالت کلی یک ساختار دو لایه با یک لایه ورودی و یک لایه خروجی دارند. هر لایه از تعدادی نرون^۷ تشکیل شده است به طوری که نرون‌های لایه ورودی وظیفه انتقال داده‌ها به شبکه را بر عهده دارند و تعداد آن‌ها با بعد بردارها در فضای ورودی برابر است. نرون‌های لایه خروجی نیز با توجه به روابط همسایگی مشخصی که بین آن‌ها تعریف شده و رفتار متقابل روی همدیگر، خروجی شبکه را ایجاد می‌نمایند (چنگ و لای^۸، ۲۰۰۵). آموزش شبکه خودسازمان‌ده بر مبنای الگوریتم یادگیری بدون ناظر و فاصله اقلیدسی^۹ بنا شده است. پس از پایان مرحله آموزش، نقشه‌ای از نرون‌ها به

1 Pair wise Comparisons

2 Consistency Rate

3 Kohonen

4 Seo & Obermayer

5 Penn

6 Kohonen & Makisara

7 Neuron

8 Chang, & Lai

9 Euclidean distance

دست می‌آید که در واقع چکیده‌ای از فضای مورد تحلیل شبکه است (کریدلر^۱، ۲۰۰۸؛ سگف، توکسوز و مرهون^۲، ۲۰۰۳).

کتابخانه عمومی. کتابخانه نهادی اجتماعی است که با گردآوری، سازمان‌دهی و نگهداری دانش مدون بشر امکان رشد اندیشه، شکوفایی استعداد و باروری ذهن خلاق انسان پویا و جستجوگر را فراهم می‌کند. از بین انواع کتابخانه‌ها، کتابخانه عمومی برای عموم افراد جامعه به وجود آمده است و وظیفه دارد که در تمام رشته‌های مختلف، انواع کتاب‌ها، مجلات و نشریات سودمند و جدید را جمع‌آوری کرده و در اختیار مردم بگذارد (عمادخراسانی، ۱۳۸۲).

روش‌شناسی پژوهش

تحقیق حاضر با روش پیمایشی - تحلیلی به اجرا در آمده است که در آن با ترکیب روش‌های تصمیم‌گیری چندمعیاره^۳ و داده‌کاوی چارچوبی برای خوشه‌بندی و پیش‌بینی رفتار مراجعین کتابخانه‌های عمومی ارائه شده است تا زمینه را جهت شناسایی مراجعین کلیدی و انتخاب استراتژی‌های مناسب با توجه به ویژگی‌های مراجعین هر بخش فراهم نماید. داده‌های مورد استفاده از بانک اطلاعاتی مراجعین کتابخانه عمومی شهید دستغیب شیراز استخراج شده‌اند و شامل اطلاعاتی از قبیل سن، جنسیت، وضعیت تاهل، میزان تحصیلات، تاریخ آخرین مراجعه به کتابخانه، تعداد مراجعه و سابقه عضویت برای هر مراجعه‌کننده (رکورد) است. این پژوهش در قالب دوره زمانی مقطعی و در بازه زمانی مربوط به ۱۳۹۱/۵/۱ تا ۱۳۹۲/۴/۳۱ انجام شده است و بانک اطلاعاتی مراجعین مورد استفاده در پژوهش شامل ۴۰۰۷۳ رکورد یا مراجعه‌کننده می‌باشد.

از ویژگی‌های خاص این پژوهش در مقایسه با سایر مطالعات گذشته می‌توان به استفاده از مدل RFM اشاره نمود که سه فاکتور تازگی^۴ (R)، تعداد دفعات^۵ (F) و ارزش پولی^۶ (M) را برای هر رکورد در نظر می‌گیرد (رزمی و قنبری، ۱۳۸۸؛ باتل^۷، ۲۰۰۴؛ چنگ و چن^۸، ۲۰۰۹؛ هوگاس^۱،

¹ Kreidler

² Saggaf, Toksoz and Marhoon

³ Multi Criteria Decision Making (MCDM)

⁴ Recency (R)

⁵ Frequency (F)

⁶ Monetary (M)

⁷ Buttle

⁸ Cheng, & Chen

(۱۹۹۴). در واقع در این مدل، فرض بر این است که مراجعین دارای ارزش بالای هر یک از شاخص‌های مدل، بهترین مراجعین هستند، البته تا زمانی که در آینده همانند گذشته رفتار نمایند (کینینگهام، اکسوی و بیجو^۲، ۲۰۰۶). دلیل استفاده از مدل RFM آن است که در این مدل فقط نگرش‌های مالی مطرح نمی‌باشند و گرایش اصلی آن در تحلیل ویژگی‌ها، به سمت مسائل غیرمالی است (سیدحسینی و همکاران، ۲۰۱۰). در این پژوهش با توجه به اینکه مراجعین بار مالی برای کتابخانه ندارند، از شاخص سوم چشم پوشی شده و شاخص دیگری تحت عنوان سابقه عضویت مراجعین^۳ (RR) در نظر گرفته شده است.

نظرات مختلفی پیرامون اهمیت شاخص‌های مدل RFM وجود دارند (بین، پیجی و دان^۴، ۲۰۰۸) که در این پژوهش برای وزندهی به شاخص‌های مدل، از روش تحلیل سلسله مراتبی بهره گرفته شده و با توجه به هوش‌مندی و سرعت شبکه‌های عصبی در تحلیل داده‌ها، از شبکه عصبی خودسازمان‌ده برای خوشه‌بندی استفاده شده است. از دیگر ویژگی‌های این پژوهش ترکیب بخش‌بندی مراجعین در قالب هرم ارزش دوره عمر است. به این ترتیب گام‌های فرایند پژوهش حاضر به شرح زیر است:

گام اول. جمع‌آوری و پیش‌پردازش داده‌ها بر مبنای شاخص‌های مدل: این قسمت شامل تمام مراحل است که قبل از انجام پردازش‌های اصلی روی داده‌ها جهت آماده‌سازی آنها برای پردازش‌های بعدی، انجام می‌گیرند.

گام دوم. نرمال‌سازی شاخص‌های مدل: با توجه به اینکه آموزش شبکه خودسازمان‌ده بر مبنای فاصله اقلیدسی است، نرمال بودن داده‌های آموزش شبکه می‌تواند نقشه‌های خروجی را به شدت تحت تاثیر قرار دهد که با استفاده از فرمول‌های (۱) تا (۳) انجام می‌شود.

$$(۱) R' = \frac{R_{\max} - R}{R_{\max} - R_{\min}} \quad (۲) F' = \frac{F - F_{\min}}{F_{\max} - F_{\min}} \quad (۳) RR' = \frac{RR - RR_{\min}}{RR_{\max} - RR_{\min}}$$

^۱ Hughes

^۲ Keiningham, Aksoy and Bejou

^۳ Registration Record

^۴ Bin, Peiji and Dan

در روابط بالا R ، F و RR مقادیر اصلی شاخص‌ها و R_{\max} ، F_{\max} و RR_{\max} بیشترین مقادیر شاخص‌ها و R_{\min} ، F_{\min} و RR_{\min} کمترین مقدار شاخص‌ها و R' ، F' و RR' نیز مقادیر نرمال شده شاخص‌ها هستند.

گام سوم. وزن‌دهی به شاخص‌های مدل: وزن نسبی شاخص‌ها، با استفاده از روش تحلیل سلسله مراتبی محاسبه شده و با W_{RR} ، W_F و W_R نشان داده می‌شود. گام چهارم. تعیین ارزش شاخص‌های مدل برای هر مراجعه کننده: ارزش هر شاخص از ضرب مقدار نرمال شده آن در وزن آن طبق فرمول‌های (۴) تا (۶) تعیین گردیده و به ترتیب با R'' ، F'' و RR'' نشان داده می‌شوند.

$$(۴) R'' = R' \times W_R \quad (۵) F'' = F' \times W_F \quad (۶) RR'' = RR' \times W_{RR}$$

گام پنجم. تعیین متوسط ارزش شاخص‌های مدل: متوسط ارزش هر یک از این شاخص‌ها با تقسیم مجموع ارزش آن شاخص در همه رکوردها به تعداد کل مراجعین طبق فرمول‌های (۷) تا (۹) تعیین می‌گردد.

$$(۷) \bar{R}'' = \frac{\sum R''}{n} \quad (۸) \bar{F}'' = \frac{\sum F''}{n} \quad (۹) \overline{RR}'' = \frac{\sum RR''}{n}$$

گام ششم. خوشه بندی مراجعین: برای خوشه بندی از روش خوشه بندی شبکه خودسازمانده استفاده شده و نتایج حاصل از خروجی شبکه در قالب نقشه‌های گرافیکی تحلیل می‌گردد؛ این عملیات در نرم افزار MATLAB انجام شده است.

گام هفتم. تعیین متوسط ارزش شاخص‌های مدل در هر خوشه: متوسط ارزش هر یک از شاخص‌ها در هر خوشه از تقسیم مجموع ارزش شاخص در آن خوشه به تعداد مراجعین آن خوشه تعیین می‌گردد که به ترتیب با $M_{RR''}$ ، $M_{F''}$ و $M_{R''}$ نمایش داده می‌شود. فرمول‌های (۱۰) تا (۱۲) این عملیات را نشان می‌دهند.

$$(۱۰) M_{R''} = \frac{\sum R''}{n} \quad (۱۱) M_{F''} = \frac{\sum F''}{n} \quad (۱۲) M_{RR''} = \frac{\sum RR''}{n}$$

گام هشتم. تحلیل خوشه‌ای: این تحلیل از طریق مقایسه متوسط ارزش شاخص‌ها در هر خوشه با متوسط ارزش شاخص‌ها در کل داده‌ها و همچنین با مقایسه رتبه‌های خوشه‌ها در هر یک

شاخص‌ها صورت می‌گیرد. این مقایسه مشخص می‌نماید که متوسط ارزش هر یک از شاخص‌های مدل در هر خوشه نسبت به متوسط ارزش این شاخص‌ها در کل داده‌ها در چه وضعیتی قرار دارند. برای هر یک از شاخص‌ها، وضعیت مطلوب، وضعیتی است که متوسط ارزش شاخص در یک خوشه بهتر از متوسط ارزش آن شاخص در کل داده‌ها باشد و با علامت (↑) نشان داده می‌شود وضعیت نامطلوب یعنی وضعیتی که متوسط ارزش شاخص در یک خوشه بهتر از متوسط ارزش آن در کل داده‌ها نباشد و با علامت (↓) نشان داده می‌شود.

گام نهم. محاسبه ارزش دوره عمر مراجعین هر خوشه: ارزش دوره عمر مراجعین هر خوشه از مجموع وزین متوسط ارزش شاخص‌های RFM در آن خوشه طبق فرمول (۱۳) محاسبه می‌گردد.

$$CLV = W_R \times M_{R^*} + W_F \times M_{F^*} + W_{RR^*} \times M_{RR^*} \quad (13)$$

گام دهم. تحلیل مراجعین بر اساس ارزش دوره عمر آنها: در این گام هر کدام از خوشه‌ها بر اساس ارزش دوره عمر در قالب هرم ارزش عمر مشخص می‌شوند. هر چه از سطح پایین این هرم به سطوح بالای آن حرکت می‌کنیم، خوشه‌های با ارزش دوره عمر بیشتر قرار گرفته‌اند.

یافته‌های پژوهش

پرسش اول: شاخص‌های RFM در بخش‌بندی مراجعین کتابخانه عمومی شهید دستغیب شیراز به چه صورت تعریف می‌شوند؟

پس از نظر خواهی از شش نفر از متخصصان علم اطلاعات و دانش‌شناسی که به طور متوسط بین هفت تا ده سال در کتابخانه عمومی فعالیت نموده‌اند، شاخص‌های مدل RFM به این صورت تعریف شدند: شاخص تازگی (R) مربوط به فاصله زمانی بین آخرین مراجعه صورت گرفته توسط مراجعه کننده تا پایان دوره مورد بررسی است که کمتر بودن این فاصله نشان‌گر بالا بودن ارزش این شاخص در مدل است. شاخص تعداد مراجعه (F) بیان‌گر تعداد دفعاتی است که مراجعه‌کننده در طول دوره زمانی مورد بررسی به کتابخانه مراجعه نموده و بیشتر بودن آن نشان‌گر بالا بودن ارزش این شاخص در مدل است. همانطور که پیشتر گفته شد، شاخص ارزش پولی (M) در این پژوهش

کارایی ندارد که به جای آن شاخص سومی تحت عنوان سابقه عضویت در کتابخانه (RR) برای تعیین ارزش مراجعه‌کننده در نظر گرفته شد.

در مرحله بعد پیش‌پردازش و آماده‌سازی داده‌ها با توجه به کد اختصاص یافته به هر مراجعه‌کننده صورت گرفت. برای پیش‌پردازش داده‌ها، فعالیت‌های حذف داده‌های نامناسب و ناقص انجام شد به طوری که در پایان، اطلاعات ۳۹۸۰۰ مراجعه‌کننده مورد بررسی قرار گرفت. همچنین برخی از تبدیل‌ها بر روی داده‌های اولیه انجام شد، به طوری که سن مشتریان از روی تاریخ تولد استخراج گردید و تازگی مراجعه با در نظر گرفتن تاریخ آخرین مراجعه محاسبه شد. پس از پیش‌پردازش داده‌ها، به دلیل تفاوت زیاد بین نوع آن‌ها، با استفاده از دستورالعمل‌های ذکر شده، داده‌ها نرمال شدند.

پرسش دوم: اولویت و اهمیت شاخص‌های RFM در تعیین ارزش مراجعین کتابخانه عمومی شهید دستغیب شیراز چگونه است؟

با استفاده از نظرات شش متخصص علم اطلاعات و دانش‌شناسی، ماتریس مقایسات زوجی برای معیارها و گزینه‌ها نسبت به شاخص‌ها تشکیل شد و با کمک نرم افزار اکسپرت چویس^۱ وزن گزینه‌ها محاسبه شد. طبق خروجی این روش، وزن‌های نسبی شاخص تازگی مراجعه معادل ۰/۳۳، تعداد مراجعه معادل ۰/۴۱ و سابقه عضویت مراجعه‌کننده معادل ۰/۲۶ تعیین گردید. همچنین در پایان برای اطمینان از درستی جواب، نرخ سازگاری تمام ماتریس‌ها محاسبه شد. همانطور که گفته شد نرخ سازگاری شاخصی است که میزان صحت و درستی قضاوت‌ها را نشان می‌دهد و بدین منظور باید مقدار آن کمتر از ۰/۱ باشد و اگر این مقدار بیشتر از ۰/۱ باشد قضاوت‌ها را متضاد و تصادفی در نظر می‌گیریم که باید در آن‌ها تجدیدنظر به عمل آید. در این پژوهش برای همه ماتریس‌ها نرخ سازگاری کمتر از ۰/۱ شد. وزن‌های به دست آمده میزان اهمیت شاخص‌ها را نشان می‌دهند، به این ترتیب شاخص تعداد مراجعه (F) دارای بیشترین اهمیت و شاخص سابقه عضویت (RR) دارای کمترین اهمیت در تعیین ارزش مراجعین می‌باشد.

^۱ Expert Choice

پس از تعیین ارزش شاخص‌های مدل، مقدار متوسط ارزش شاخص‌ها طبق دستورالعمل گفته شده، محاسبه شد که برای فاکتور تازگی مراجعه معادل ۰/۲۶۵۳، تعداد مراجعه معادل ۰/۰۸۶۱ و سابقه عضویت مراجعه‌کننده معادل ۰/۰۸۴۹ به دست آمد.

پرسش سوم: مراجعین کتابخانه عمومی شهید دستغیب شیراز را می‌توان به چند بخش یا خوشه تقسیم نمود و مراجعین هر خوشه چه ویژگی‌هایی دارند؟

برای انجام خوشه‌بندی، داده‌های آموزش شبکه از ۳۹۸۰۰ بردار سه بعدی شاخص‌های مدل تشکیل شده است که هر بردار نماینده یک رکورد از رکوردهای پایگاه داده مراجعین کتابخانه است. پس از انجام مراحل در پایان، پنج بخش یا خوشه اطلاعاتی به دست می‌آید و هر یک از مراجعین در یکی از این خوشه‌ها قرار می‌گیرند. نتایج خوشه‌بندی در جدول ۱ ذکر شده است.

جدول ۱. خوشه‌بندی مراجعین کتابخانه به روش شبکه عصبی خودسازمان‌ده

خوشه‌ها	۱	۲	۳	۴	۵
فراوانی مراجعین	نفر ۱۹۵۰۰	نفر ۱۵۵۲۰	نفر ۲۷۹۰	نفر ۱۱۹۰	نفر ۸۰۰
درصد مراجعین	درصد ۴۹	درصد ۳۹	درصد ۷	درصد ۳	درصد ۲
میانگین سنی مراجعین	۱۸/۳۹	۲۵	۳۴/۱۴	۵۶	۴۴/۷۵
جنسیت	زن	درصد ۷۴	درصد ۶۵	درصد ۶۴	درصد ۲۰
	مرد	درصد ۲۶	درصد ۳۵	درصد ۳۶	درصد ۸۰
تاهل	مجرد	درصد ۹۹	درصد ۷۷	درصد ۷	درصد ۳
	متاهل	درصد ۱	درصد ۲۳	درصد ۹۳	درصد ۹۷
تحصیلات	ابتدایی	درصد ۵	درصد ۰	درصد ۰	درصد ۲
	راهنمایی	درصد ۶	درصد ۰	درصد ۰	درصد ۳
	متوسطه	درصد ۷۸	درصد ۴	درصد ۱۴	درصد ۰
	دیپلم	درصد ۹	درصد ۶	درصد ۰	درصد ۴۳
	فوق دیپلم	درصد ۱	درصد ۱۲	درصد ۲۱	درصد ۱۱
	لیسانس	درصد ۱	درصد ۷۳	درصد ۴۳	درصد ۱۸
	فوق لیسانس	درصد ۰	درصد ۴	درصد ۲۲	درصد ۱
	دکتر و بالاتر	درصد ۰	درصد ۱	درصد ۰	درصد ۲۲

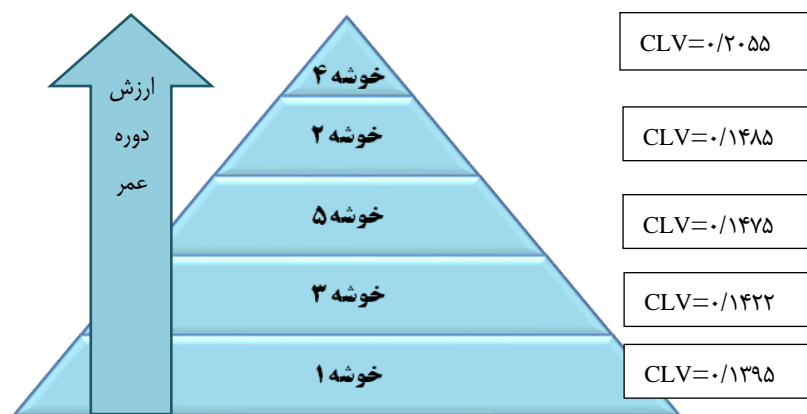
پس از تعیین خوشه‌ها، برای هر خوشه متوسط ارزش شاخص‌های مدل محاسبه و وضعیت آن با مقدار میانگین کل سنجیده شد و سپس ارزش دوره عمر برای هر خوشه محاسبه گردید که نتایج حاصل در جدول ۲ نشان داده شده است.

جدول ۲. وضعیت خوشه‌ها

خوشه‌ها		۱	۲	۳	۴	۵
تازگی مراجعه ($M_{R''}$)	متوسط ارزش	۰/۲۶۳۵	۰/۲۶۷۲	۰/۲۶۵۱	۰/۲۵۷۹	۰/۲۸۴۵
	رتبه	دوم	چهارم	سوم	اول	پنجم
تعداد دفعات مراجعه ($M_{F''}$)	متوسط ارزش	۰/۰۷۵۷	۰/۰۹۰۷	۰/۰۹۸۲	۰/۱۸۶۵	۰/۰۸۹۴
	رتبه	پنجم	سوم	دوم	اول	چهارم
سابقه عضویت ($M_{RR''}$)	متوسط ارزش	۰/۰۸۲۵	۰/۰۸۸۹	۰/۰۵۵۷	۰/۱۶۹۰	۰/۰۶۵۰
	رتبه	سوم	دوم	پنجم	اول	چهارم
وضعیت متوسط ارزش شاخص‌ها ($M_{R''}$ ، $M_{F''}$ ، $M_{RR''}$)		(↑، ↓، ↓)	(↓، ↑، ↑)	(↑، ↑، ↓)	(↑، ↑، ↑)	(↓، ↑، ↓)
ارزش دوره عمر		۰/۱۳۹۵	۰/۱۴۸۵	۰/۱۴۲۲	۰/۲۰۵۵	۰/۱۴۷۵

پرسش چهارم: کدام مراجعین برای کتابخانه عمومی شهید دستغیب شیراز، ارزش دوره عمر بیشتری دارند؟

خوشه‌ها بر اساس ارزش دوره عمر مراجعین آن‌ها، در قالب هرم ارزش عمر که در شکل (۱) نشان داده شده است، مشخص شدند. در این هرم، به ارزش دوره عمر مراجعین در ۵ خوشه اصلی اشاره شده است. هر چه از سطح پایین این هرم به سطوح بالای آن حرکت می‌کنیم، خوشه‌های با ارزش دوره عمر بیشتر قرار گرفته‌اند. به عبارتی، با حرکت به سمت رأس این هرم، با مراجعینی مواجه می‌شویم که ارزش و اهمیت بیشتری از لحاظ عمر عضویت و سابقه فعالیت برای کتابخانه دارند.



شکل ۱. هرم ارزش دوره عمر مراجعین کتابخانه عمومی

به این ترتیب طبق هرم ارزش دوره عمر، مراجعین خوشه ۴ دارای بیشترین ارزش عمر برای کتابخانه هستند. مراجعین این خوشه تنها ۳ درصد از کل مراجعین را تشکیل می‌دهند، میانگین سنی آنها ۵۶ سال است، بیشتر مرد و متأهل هستند که عمدتاً تحصیلات دیپلم دارند، این مراجعین به تازگی به کتابخانه مراجعه داشته‌اند، تعداد مراجعه آنها در این دوره یکساله بالا بوده است و طبیعی است که سابقه عضویت زیادی نیز داشته باشند.

بعد از خوشه ۴ مراجعین خوشه ۲ از اهمیت بالایی برای کتابخانه برخوردارند، مراجعین این خوشه ۳۹ درصد از کل مراجعین را تشکیل می‌دهند، میانگین سنی آنها تقریباً ۲۵ سال است، عمدتاً زن و مجرد هستند که بیشتر آنها تحصیلات لیسانس دارند، این مراجعین به تازگی به کتابخانه مراجعه نداشته‌اند اما تعداد مراجعه آنها در این دوره یکساله بالا بوده است و سابقه عضویت زیادی نیز دارند.

اولویت بعدی با خوشه ۵ است که حدود ۲ درصد از کل مراجعین را شامل می‌شود و میانگین سنی آنها تقریباً ۴۵ سال است. عمدتاً متأهل هستند و تحصیلات دیپلم به بالا دارند. این مراجعین تعداد مراجعه زیادی دارند.

خوشه ۳ در اولویت بعدی است که ۷ درصد مراجعین را شامل می‌شود. میانگین سنی مراجعین این خوشه تقریباً ۳۴ سال است و حدوداً بیشتر آن‌ها زن و متأهل هستند که اکثر آن‌ها تحصیلات لیسانس و فوق لیسانس دارند. این مراجعین به تازگی به کتابخانه مراجعه داشته‌اند و تعداد مراجعه آن‌ها در این دوره یکساله بالا بوده است ولی سابقه عضویت زیادی ندارند.

خوشه ۱ معرف مراجعینی است که کمترین ارزش را برای کتابخانه از نظر سابقه عضویت دارند. این خوشه عمده مراجعین یعنی در حدود ۴۹ درصد آن‌ها را تشکیل می‌دهد که میانگین سنی آن‌ها تقریباً ۱۸ سال است، بیشتر آن‌ها مجرد و زن هستند، همچنین به طور عمده تحصیلات متوسطه دارند. این مراجعین به تازگی به کتابخانه مراجعه داشته‌اند اما تعداد مراجعه آن‌ها در این دوره یکساله کم بوده است و سابقه عضویت زیادی ندارند.

برای مدیریت صحیح منابع، وقت و هزینه در کتابخانه‌های عمومی، ضمن در نظر گرفتن هرم ارزش مراجعین، کتابخانه‌ها باید ابتدا نیازهای مراجعینی که در سطوح بالای هرم ارزش قرار دارند را در نظر بگیرند. بر این اساس به عنوان مثال خوشه شماره ۴ با ارزش‌ترین مراجعین را شامل شده که باید در کانون توجه کتابخانه‌ها قرار بگیرند و پس از آن مشتریان خوشه ۲ و به همین ترتیب این روند ادامه دارد. کتابخانه باید در جهت حفظ مراجعین با ارزش خود که در سطوح بالای هرم قرار دارند، تلاش نماید و اهمیت ویژه‌ای برای مراجعین این خوشه‌ها قائل شود. در مورد مراجعینی که در سطوح پایین هرم قرار دارند، نیز باید با برقراری ارتباطات و تعاملات سازنده دلیل ایجاد شکاف و فاصله در روند مراجعه آن‌ها را جویا شده و در جهت رفع آن برآید.

نتیجه‌گیری

در این تحقیق، به منظور بخش‌بندی مراجعین کتابخانه عمومی شهید دستغیب شیراز بر اساس ارزش دوره عمر آن‌ها، از مدل RFM و تکنیک‌های تصمیم‌گیری چندمعیاره و داده‌کاوی در قالب فرایندی خاص، بهره گرفته شده است. ارائه تحلیل خوشه‌ای به منظور بررسی دقیق‌تر ویژگی‌های مراجعین در خوشه‌ها و همچنین، بخش‌بندی نهایی مراجعین بر اساس ارزش دوره عمر آن‌ها در

قالب هرم ارزش دوره عمر، از مواردی است که در این مطالعه از آنها بهره گرفته شده است. یکی از ویژگی‌های خاص این پژوهش استفاده از روش شبکه عصبی خودسازمان‌ده است که یکی از روش‌های دقیق در حوزه خوشه‌بندی و داده‌کاوی است.

بر اساس یافته‌های پژوهش، مراجعین کتابخانه مذکور به ۵ خوشه تفکیک شدند که با توجه به ارزش دوره عمر مراجعین، خوشه‌های ۴، ۲، ۵، ۳ و ۱ به ترتیب در رتبه‌های اول تا پنجم قرار گرفتند. طبق هرم ارزش مراجعین، کتابخانه باید در جهت حفظ مراجعین با ارزش خود (دارای بیشترین ارزش دوره عمر) در خوشه‌های ۴ و ۲ تلاش نماید و اهمیت ویژه‌ای برای مراجعین این خوشه‌ها قائل شود. از آنجا که مراجعین خوشه ۱ دارای کم‌ترین ارزش دوره عمر هستند و تعداد قابل توجهی را نیز شامل می‌شوند (حدود نیمی از کل مراجعین)، لذا بهتر است که مطالعه دقیق‌تری پیرامون علل کم بودن مقادیر هر یک از شاخص‌های مدل در این مراجعین انجام گیرد. پیشنهاد می‌گردد که کتابداران به منظور حفظ این مراجعین، ارتباطات و تعاملات بیشتری با آنها داشته باشند تا از این روش در راستای وفادار کردن این مراجعین اقدام کنند. در این زمینه کتابخانه باید با برقراری ارتباطات و تعاملات سازنده دلیل ایجاد شکاف و فاصله در روند مراجعه آنها را جویا شده و در جهت رفع آن برآید، همچنین کتابخانه می‌تواند با در نظر گرفتن امکانات و تسهیلات ویژه در راستای ارتقای سطح شاخص‌های مدل و به تبع آن، افزایش ارزش دوره عمر مراجعین این خوشه گام بردارد در صورت امکان از رویگردانی‌های بعدی مراجعین جدید خود بکاهد.

لازم به ذکر است که رفتار مراجعین ثابت نیست و بانک اطلاعاتی مراجعین نیز تغییر می‌کند. مجموعه قوانینی که از بانک اطلاعاتی برای یک زمان خاص دریافت شده فقط برای یک زمان مشخص دارای اعتبار است و سپس منسوخ می‌شود، بنابراین پیشنهاد می‌گردد که در مطالعات آتی، از مجموعه داده وسیع‌تری به لحاظ قلمروی زمانی استفاده شود که به طور قطع منتج به ایجاد نتایج قوی‌تر و گستره وسیع‌تری از دانش کاربردی پیرامون ویژگی‌های رفتاری مراجعین خواهد شد. از طرفی، برش‌های زمانی متناوب از پایگاه داده شرکت، امکان اجرای پویای داده کاوی بر مبنای مدل RFM جهت تعیین ارزش دوره عمر مراجعین را فراهم خواهد آورد که نتایج آن از این لحاظ که

روند تغییرات در رفتار مراجعین را منعکس می کند، می تواند نقش شایانی در ارتباط با مراجعین ایفا نماید.

فهرست منابع

- آذر، عادل؛ احمدی، پرویز و محمد وحید سبط. (۱۳۸۹). طراحی مدل انتخاب نیروی انسانی با رویکرد داده کاوی (مورد: استخدام داوطلبان آزمون های ورودی یک بانک تجاری در ایران). نشریه مدیریت فناوری اطلاعات، (۲) ۴: ۳-۲۲، ۹۸۶-۱۰۰۱.
- بیگدلی، زاهد و بهروز فرهادیان. (۱۳۸۱). نظر مراجعان کتابخانه های عمومی خوزستان درباره تناسب مجموعه ها با نیازهای اطلاعاتی آنان. علوم اجتماعی و انسانی دانشگاه شیراز، ۳۶: ۲۰۴-۲۱۸.
- رزمی، جعفر و آرش قنبری. (۱۳۸۸). ارائه مدلی نوین جهت محاسبه ارزش دوره عمر مشتری، نشریه مدیریت فناوری و اطلاعات، (۱) ۱: ۳۵-۵۰.
- عمادخراسانی، نسرین دخت. (۱۳۸۲). خدمات عمومی کتابخانه و روش های آن. تهران: نشر کتابدار، چاپ سوم.
- غضنفری، مهدی (۱۳۹۱). داده کاوی و کشف دانش. تهران: دانشگاه علم و صنعت ایران.
- کفایش پور، آذر؛ توکلی، احمد و علی علیزاده زوارم. (۱۳۹۱). بخش بندی مشتریان بر اساس ارزش دوره عمر آن ها با استفاده از داده کاوی بر مبنای مدل آر. اف. ام. (RFM). پژوهش های مدیریت عمومی، (۱۵) ۵: ۶۳-۸۴.
- مهرگان، محمد رضا. (۱۳۸۳). پژوهش عملیاتی پیشرفته. تهران: انتشارات کتاب دانشگاهی، چاپ اول.

Barry, M. (2003). Data Mining. 9th Indian. Delhi: Pearson University
Bin, D., Peiji, S., and Dan, Z. (2008). Data mining for needy student identify based on improved RFM model: A case study of university, International conference on information management, innovation management and industrial engineering, New York.

- Buttle, F. (2004). *Customer Relationship Management: Concepts and Tools*. Elsevier Butterworth Heinemann.
- Cannavo, S. (2003). *The Power of Logic in Everyday Life*. Mumbai: Magna Publishing Co.Ltd.
- Chang, P. C., and Lai, C. Y. (2005). A hybrid system combining self-organizing maps with case-based reasoning in wholesaler's new-release book forecasting. *Expert Systems with Applications*, 29, 183–192.
- Cheng, C. H., and Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, 36, 4176-4184.
- Guha, S., Rastogi, R., and Shim, K. (1998). Cure: an efficient clustering algorithm for large databases: Proceedings of the ACM SIGMOD conference, 73-84.
- Han, J., and Kamber, M. (2006). *Data mining: concepts and techniques*, San Morgan, Francisco Kaufman.
- Hsieh, N. C. (2004). An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert systems with applications*, 27, 623-633.
- Hu, W., and Jing, Z. (2008). Study of segmentation for auto services companies based on RFM model, [online], <[http:// www.pucsp.br/icim/ingles/downloads/pdf_procceeding_2008/66.pdf](http://www.pucsp.br/icim/ingles/downloads/pdf_procceeding_2008/66.pdf)>.
- Hughes, A. M. (1994). *Strategic database marketing*, Chicago: Probus publishing.
- Kantardzic, M (2003). *Data Mining: Concepts, Models, Methods, and algorithms*, wiley – inter science.
- Keiningham, T. L., Aksoy, A., and Bejou, D. (2006). Approches to measurement and management of customer value. *Journal of relationship marketing*. 5 (2), 37-54.
- Khajvand, M., and Tarokh, M. J. (2011). Estimating customer future value of different customers segments based on adapted RFM model in retail banking context. *Procedia Computer Science*, 3, 1327-1332.
- Kim, Y. S., and Sohn, S. Y. (2004). Managing loan customers using misclassification patterns of credit scoring model. *Expert Systems With Applications*, (26), 567-573.
- Kohonen, T. (2001). *Self-organizing Maps*, Springer Series in Information sciences, Springer, Berlin, New York, Volume 30.
- Kohonen, T., and Makisara, k. (1984). Phonotypic maps –insightful representation of phonological features for speech recognition, Los Alamitos, CA. IEEE Computer Soc. Press, 182 -185.
- Kreidler, M (2008). *Guide to auto insurance*, Washington State Office of the Insurance Commissioner, vived 2010/4/24 <www.insurance.wa.gov>
- Lee, T. S., and Chen, I. F. (2005). A two – stage hybrid credit scoring model using systems with applications, (28), 743-725.
- Li, D. C., Dai, W. L., and Tseng, E. T. (2011). A twostage clustering mehod to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. *Expert Systems With Applications*, 14, 1-6.

- Namvar, M., Gholamian, M. R., and KhakAbi, S (2010). A Two Phase Clustering Method For Intelligent Customer Segmentation. International Conference on Intelligent Systems, Modelling and Simulation, IEEE 2010.
- Ngai, E. W. T., Xiu, L., and Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems With Applications*, 36, 2592-2602.
- Niyagas, W., Srivihok, A., and Kitisin, S (2006). Clustering e-banking customer using dataminind and marketing segmentation. *Ectitransactions On Compter And Information Technology*, 2(1).
- Parker, V. N. (1990). determining Priorities for Compuerisation in Libraries. *Annals of Library Science and Documentation*, 37(3), 110-116.
- Patkar, V. N. (2005). Data mining applications in library and academic institutions. *Information Studies Journal of Science Direct*, 11(3): 145-156.
- Penn, B. S. (2005). Using self-organizing maps to visualize high-dimensional data. *Computers & Geosciences*, (31), 531-544.
- Saggaf, M. M., Toksoz, M. N., and Marhoon, M. I. (2003). Seismic facies classification and dentification by competitive neural networks: *Geophysics*, (44), 1041-1063.
- Seo, S., and Obermayer, K (2004). Self-organizing maps and clustering methods for matrix data. *Neural Networks*, 17, 1211-1229.
- Seyed Hosseini, M., Maleki, A., and Gholamian, M. R (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems With Applications*, 37, 5259-5264.
- Sohrabi, B., and Khanlari, A (2007). Customer lifetime value (CLV) measurement based on RFM model. *Iranian accounting & auditing review*, 14(47), 7-20.
- Wu, H. H. Chang, E. C., and Lo, C. F (2009). Applying RFM model and K-means method in customer value analysis of an outfitter, *International Conference on Concurrent Engineering*, New York.