



مجموعه مقالات اولین همایش تخصصی

"سیستهای هوشمند کامپیوتری و کاربردهای آنها"

دانشگاه پیام نور استان تهران - مجتمع فنی و مهندسی (آذرماه ۱۳۹۰)

## بسمه تعالی

# "ارائه مدلی برای ارزش گذاری مشتریان یک فروشگاه اینترنتی با استفاده از تکنیکهای داده کاوی : مطالعه موردی در سایت سیویلیکا"

نویسندگان: ۱- شری شیرین بیک مهاجر، ۲- هومن تصدیقی، ۳- طه مکفی

(۱) دانشجوی کارشناسی ارشد مهندسی فناوری اطلاعات، sherimohajer@yahoo.com

(۲) عضو هیات علمی جهاد دانشگاهی صنعتی شریف، tasdighi@gmail.com

(۳) کارشناس مهندسی صنایع، mokfi.taha@gmail.com

## چکیده

امروزه اینترنت به یکی از ابزارهای مهم در همه عرصه های زندگی بشر تبدیل شده است. وجود انواع مقالات علمی، کتب و سایر منابع الکترونیکی در اینترنت در عرصه های علمی و تحقیقاتی به طور گسترده ای توسط محققان و دانشمندان سراسر دنیا مورد پذیرش قرار گرفته است. در این راستا سایت سیویلیکا با دارا بودن بیش از 150000 مقاله ای داخلی و خارجی در زمینه های مختلف علمی، یکی از محبوب ترین مراجع به منظور سهولت دسترسی به این منابع ارزشمند است.

هدف این مقاله بکارگیری تکنیک های داده کاوی برای رده بندی و تحلیل رفتار مشتریان این سایت به منظور شناخت بهتر آنها و در نتیجه بهبود خدمات ارائه شده و تدوین استراتژی های بازاریابی می باشد. ما در این مقاله روشی را جهت بخش بندی و ارزش گذاری مشتریان بر اساس سابقه خرید آنها با استفاده از روش *CRISP-DM* و نرم افزار داده کاوی *Clementine 12.0* توصیف خواهیم نمود.

کلمات کلیدی: داده کاوی، بخش بندی مشتریان، مدیریت دانش، خوشه بندی

## ۱- مقدمه

این شده که تنوع کاربران و درخواست هایشان زیاد باشد [۱۳]. از جمله کارهایی که امروزه روی شبکه ی اینترنت رشد و توسعه یافته، خرید اینترنتی یا به طور جامع تر تجارت الکترونیکی است. با توجه به جاذبه های فراوانی که در شبکه ی جهانی اینترنت وجود دارد برخی کسب و کار های جدید به وجود آمده که هر کدام مشتریان و علاقه مندان خاص خود را دارند. با توجه به حجم عظیم داده

امروزه استفاده از فناوری های مدرن در زندگی روزمره جهت رفع نیازها امری اجتناب ناپذیر است. دنبال کردن اخبار و انجام کارها از طریق اینترنت رایج شده است سازمانها روی شبکه اینترنت بستری فراهم می کنند تا اطلاعاتشان در دسترس مردم باشد. روزانه حجم زیادی از اطلاعات در اینترنت منتشر می شود و این مسئله منجر به

داده های مشتریان و ابزارهای تکنولوژی اطلاعات مبنایی را برای هر مدیریت ارتباط با مشتری موفق تشکیل میدهند. به علاوه، سرعت رو به رشد اینترنت و تکنولوژی های وابسته باعث افزایش فرصت هایی برای بازاریابی شده است و راه های ارتباطی بین شرکت ها و مشتریانانشان را تغییر داده است [۹].

## ۲-۲- چرخه زندگی مشتری

این چرخه به مراحل ارتباط مشتری و تجارت مربوط است و آگاهی از آن سودآوری مشتری را به همراه خواهد داشت. ۴ مرحله در چرخه زندگی مشتری وجود دارد:

۱. مشتریان بالقوه: افرادی که هنوز مشتری نیستند، اما در هدف بازار قرار دارند.
۲. مشتریانی که واکنش نشان می دهند: مشتریان بالقوه یا احتمالی که به محصول یا خدمتی علاقه و واکنش نشان می دهند.
۳. مشتریان بالفعل: مشتریانی که از محصول یا خدمتی استفاده می کنند.
۴. مشتریان سابق: این دسته افراد، مشتریان مناسبی نیستند زیرا مدت طولانی، خرید خود را به سمت محصولات رقیب برده اند [۳].

## ۲-۳- ابعاد مدیریت ارتباط با مشتری

- مدیریت ارتباط با مشتری شامل ۴ بعد می باشد:
- شناسایی مشتریان: مدیریت ارتباط با مشتری با شناسایی مشتریان شروع میشود، این فاز شامل هدف قرار دادن جمعیتی است که احتمال اینکه مشتری ما شوند زیاد باشد یا برای سازمان سودآور میباشند. عناصر فاز شناسایی مشتریان، شامل تجزیه و تحلیل مشتریان هدف و بخش بندی مشتریان میباشد.
  - جذب مشتریان: این فاز بعد از فاز شناسایی صورت می گیرد. بعد از شناسایی مشتریان بالقوه،

های خرید مشتریان، داده کاوی می تواند جهت طبقه بندی و ارزش گذاری مشتریان مورد استفاده قرار گیرد. در این مقاله پس از مروری بر ادبیات تحقیق و آشنایی با مفاهیمی چون مدیریت ارتباط با مشتری، چرخه زندگی مشتری و شناخت و ارزش گذاری مشتریان، با استفاده از تکنیکهای داده کاوی به بخش بندی و ارزش گذاری مشتریان یک فروشگاه اینترنتی می پردازیم. در انتها به این نتیجه می رسیم که استفاده از تکنیکهای خوشه بندی روش مناسبتری برای طبقه بندی مشتریان می باشد. این مقاله نشان می دهد که بخش بندی و شناخت ارزش مشتریان امری مفید است و در تعیین سیاستها و تعیین نوع خدمات به بازار و به مشتریان خاص متمر ثمر خواهد بود.

## ۲-۲- شناخت و ارزش گذاری مشتریان

امروزه توجه اصلی بیشتر صنایع به سمت مشتری محوری معطوف گردیده است. شناخت نیازهای واقعی و ارائه بهترین و به موقع ترین سرویس ها مطابق با نیاز مشتریان، اصول اساسی مدیریت ارتباط با مشتری را تشکیل می دهد. توجه سریع و به موقع به نیازهای اساسی مشتریان، عاملی در صرفه جویی هزینه ها و بالا بردن صرفه اقتصادی است [۵].

داده کاوی یکی از عناصر مدیریت ارتباط با مشتری است که از تکنیک هایی برای تعامل و ارتباط با مشتریان استفاده میکند. یکی از این تکنیک های داده کاوی، وب کاوی است که در عرصه وب استفاده میشود. وب کاوی به تکنیک های خودکار بازاریابی، استخراج و ارزیابی اطلاعات از پرونده ها و سرویس های تحت وب به منظور کشف دانش اشاره دارد [۱].

## ۲-۱- مدیریت ارتباط با مشتری

مدیریت ارتباط با مشتری شامل مجموعه ای از فرآیندها و سیستم های توانا می باشند که استراتژی کسب و کار را برای ساختن روابط بلند مدت و سودآور با مشتریان حمایت می کند.



### ۱-۳- بخش بندی مشتریان

بخش بندی مشتریان یکی از روشهای شناخت مشتری بوده و فرآیندی است که بر اساس آن یک بازار به بخش های متمایزی از مشتریانی که نیازها و مشخصاتشان یکسان است، تقسیم بندی می شود. ایده اصلی این روش اینست که، از آنجا که مشتریان هر بخش نیازها و خواسته های یکسانی دارند، به یک طرح بازاریابی و یک محصول خاص به شیوه ای مشابه پاسخ خواهند داد. بخش بندی بازار کمک می کند تا با تقسیم یک بازار بزرگ به بخش های کوچکتر، شناخت گروههای مشتریان آسان تر شده، استراتژی های بازاریابی کاراتری تعریف گردد. همچنین با این روش منابع سازمان در بخش مدیریت ارتباط با مشتری به صورت بهینه تری مصرف می گردد، چراکه بخش عمده ای از منابع سازمان به مشتریان با ارزش و سود آور تخصیص خواهد یافت. بخش بندی مشتریان در فضای تجارت الکترونیک، از طریق انبوه داده های ذخیره شده در پایگاههای مختلف داده مشتریان صورت می گیرد. [۲]

بخش بندی مشتریان اساس بازاریابی و سرویس دهی اثربخشی یک سازمان است که تعداد زیاد مشتریان را در دسته هایی طبقه بندی می کند که مشتریان یک دسته دارای خصوصیات مشابهی با هم و خصوصیات متفاوتی با مشتریان سایر گروهها می باشند. در مقایسه با روش های سنتی بخش بندی مشتریان، بخش بندی مشتریان با استفاده از داده کاوی دارای مزایای زیر می باشد: [۱۱]

- نتایج بخش بندی به کمک داده کاوی بر اساس واقعیت داده ها شکل می گیرد و نقش فاعلی افرادی که داده ها را پردازش می کنند حذف می گردد که باعث می گردد نتایج نهایی بیشتر هدف نشان دادن تفاوت های میان جمعیت ها را محقق کند.
- مشخصات دسته بندی را در گروههای مختلف مشتریان به صورت جامع تری نمایش می دهد. که این موضوع منجر به شناخت کامل تر متخصصان بازاریابی از مشتریانشان می شود و از

سازمانها می توانند تلاش و منابع مستقیم جذب مشتریان هدف نمایند. یک عنصر جذب مشتری بازاریابی مستقیم میباشد.

- حفظ مشتریان : این فاز بحث اصلی مدیریت ارتباط با مشتری است. رضایت مشتری که اشاره به انتظارات مشتری با ادراک رضایت مندی میباشد، شرط اساسی برای حفظ مشتری میباشد.
- توسعه مشتریان : این فاز شامل توسعه کثرت تراکنش ها، ارزش تراکنش ها و سودآوری مشتریان میباشد.

این چهار بعد را می توان به عنوان حلقه بسته مدیریت ارتباط با مشتری در نظر گرفت [۴].

### ۳- تکنیکهای داده کاوی برای شناخت و ارزش گذاری مشتریان

ابزارهای داده کاوی در راستای تحلیل داده های مشتری در ساختار CRM تحلیلی<sup>a</sup>، بسیار مرسوم هستند. بسیاری از سازمان ها داده هایی در مورد مشتریان جاری، مشتریان بالقوه، تامین کنندگان و شرکای تجاری جمع آوری و ذخیره می کنند. عدم توانایی کشف اطلاعات ارزشمند پنهان در میان این داده ها مانع می شود که سازمانها این داده ها را به دانش مفید و با ارزش تبدیل کنند. ابزارهای داده کاوی کمک می کنند تا سازمان ها این دانش نهفته را از میان حجم عظیم داده ها استخراج کنند. با وجود داده های جامع مشتریان، تکنولوژی داده کاوی می تواند هوش تجاری با قابلیت ایجاد فرصتهای جدید فراهم آورد. کاربرد ابزارهای داده کاوی در CRM، روندی نوظهور در اقتصاد جهانی است. تحلیل و فهم رفتار و مشخصات مشتری مبنای توسعه یک استراتژی CRM رقابتی برای بدست آوردن و نگه داشتن مشتریان بالقوه و ماکزیمم کردن ارزش مشتری است. [۹]

مدلهای مختلفی برای انجام پروژه های کشف دانش و داده کاوی وجود دارد که از این میان، مدل " فرایند میان صنعتی استاندارد برای داده کاوی (CRISP-DM)"، یکی از روشهای معمول و موفق است که در بسیاری از پروژه های صنعتی مورد استفاده قرار گرفته است. مراحل اصلی این مدل شامل شناخت کسب و کار، شناخت داده، پیش پردازش داده، مدلسازی (داده کاوی)، ارزیابی و توسعه می باشد. [۸]

### ۵- ساختار اجرایی پژوهش

ماهیت این بخش پژوهش داده محور<sup>d</sup> بوده و پایه اصلی تحقیق حاضر بر کشف دانش از پایگاه داده های فروشگاه تجارت الکترونیکی مذکور نهاده شده است. از اینرو استاندارد جهانی CRISP-DM جهت انجام فرایند تحقیق مورد استفاده قرار گرفته است که در این بخش ساختار اجرایی تحقیق بر اساس مراحل این استاندارد تشریح شده است.

### ۶- بررسی موردی اجرای الگوریتم CRISP-DM

#### جهت ارزش گذاری مشتریان در سایت

#### سیویلیکا

در این بخش با اقتباس از مدل CRISP-DM، چارچوبی برای ارزش گذاری مشتریان این سایت بر اساس سابقه خرید آنها و به روش بخش بندی، با استفاده از دو روش خوشه بندی و گسسته سازی دستی ارائه می شود. در این چارچوب از روش تحلیل RFM برای تعیین ارزش مشتری استفاده شده است. در نهایت با استفاده از شبکه عصبی به بررسی و مقایسه عملکرد هر یک از روشهای استفاده شده در بخش بندی و ارزش گذاری مشتریان می پردازیم.

#### ۶-۱- درک فضای کسب و کار در ارتباط با بخش بندی

#### مشتریان

این طریق برنامه های بازاریابی هدفمند و اختصاصی قابل اجرا می باشند.

- تغییرات رفتاری مشتریان می تواند به سادگی با کنار هم گذاشتن مدل های آنالیز خوشه بندی و به روز کردن دسته مشتریان به صورت منظم، ردگیری شود.

مدل های سنتی بخش بندی مشتریان معمولاً، بر اساس داده های ویژگی های مشتری انجام می شوند. این نوع بخش بندی نتایج بسیار ساده با دقت پائینی برای محیط های کسب و کار پیچیده امروزی ارائه می کنند. اخیراً، بخش بندی مشتریان بر اساس داده های رفتاری و مبادلاتی مشتریان ( مثل نوع، حجم و سوابق خرید، داده های فعالیت در وب و غیره ) یا در واقع داده های ارزش مشتریان، که توسط سیستمهای اطلاعاتی مختلف جمع آوری می گردد، انجام می شود. [۷]

مدل تحلیل تازگی، تکرار و مبلغ<sup>b</sup> (RFM)، از روش های تحلیل ارزش مشتری می باشد که به عنوان یک ویژگی بسیار مهم در بخش بندی مشتریان نیز استفاده می شود. مدل تحلیلی RFM که توسط هوگس<sup>c</sup>، در سال ۱۹۹۴ معرفی شده، مدلی است که مشتریان را از طریق انبوه داده های تبادلات مشتریان به وسیله ۳ متغیر (ویژگی) R و F و M متمایز می سازد. جزئیات مدل به صورت زیر است:

۱. تازگی آخرین خرید (R): فاصله بین زمان آخرین خرید تا اکنون.
۲. تکرار خرید (F): تعداد مبادلات در یک پریود زمانی مشخص مثل دو بار در سال.
۳. کل مبالغ خرید (M): میزان پولی که در یک دوره زمانی معین صرف تمامی خریدهای مشتری شده است. [۱۲]

#### ۴- فرایند میان صنعتی استاندارد برای داده

#### کاوی (CRISP-DM)

### ۶-۱-۱- تعیین اهداف و نیازمندیهای تجاری

در این مرحله با توجه به مطالعه مستندات، وب سایت شرکت و مصاحبه با مدیریت و کارشناسان شرکت، مسئله موجود در زمینه شناخت و ارزش گذاری مشتریان سیستم فروش اینترنتی مورد بررسی قرار گرفت.

برای طراحی طرحهای تبلیغاتی تخصصی که حداکثر دستیابی به اهداف را در مقابل حداقل مصرف منابع ایجاد نمایند، شناخت مشتریانی که بیشترین احتمال جذب یا افزایش خرید را دارا هستند لازم است. شناخت مشتریان یکی از مهم ترین بخشهای چرخه فرایند مدیریت ارتباط با مشتری بوده و بخش بندی مشتریان از مهم ترین و متداولترین روش های آن محسوب می گردد.

### ۶-۱-۲- تعیین اهداف پروژه داده کاوی

بعد از مشخص کردن اهداف تجاری، زمان ترجمه و تبدیل آنها به شکل قابل قبول داده کاوی می باشد. اهداف تعیین شده به فرم داده کاوی در این بخش عبارتند از:

- شناسایی مشتریان ارزشمند شرکت بر اساس سابقه خرید آنها
- ساخت مدلی با استفاده از داده های مشتریان به منظور پیش بینی احتمال تغییرات رفتار مشتری
- تعیین رتبه هر یک از مشتریان بر اساس میل ارزش مشتری

جهت شناخت مشتریان و ارزش گذاری مشتریان از مدل تحلیل تازگی، تکرار و مبلغ (RFM) و جهت بخش بندی مشتریان از تکنیکهای خوشه بندی استفاده شده است.

### ۶-۲- شناخت داده های مشتریان

#### ۶-۲-۱- جمع آوری داده های اولیه

در این مرحله پایگاههای داده مشتریان مورد بررسی قرار گرفت. تنها بخشی که دارای داده های با ارزش و قابل استفاده به منظور ارزش گذاری مشتریان می باشد، داده

های مربوط به سوابق خرید مشتریان است که امکان تولید "ارزش مشتری" را فراهم می کند.

### ۶-۲-۲- تشریح و توصیف داده ها

جدول خرید مقالات، مربوط به ثبت اطلاعات درخواستهای خرید مقالات توسط کاربران سایت می باشد و اطلاعات دقیقی را در مورد تاریخ خرید، زمان خرید، ایمیل کاربر، کد مقاله خریداری شده و موضوع مقاله نشان می دهد. تعداد رکوردهای این جدول ۱۸۱۱۷۱ رکورد است که سوابق خرید ۳۷۳۲۰ مشتری در سال ۱۳۸۹ می باشد. نمونه ای از این جدول در زیر آمده است (جدول شماره ۱). این داده ها در قالب فایلی با فرمت mdb که برای نرم افزار کلمنتاین شناخته شده است با نام tblrequest وارد نرم افزار شده اند.

جدول ۱- نمونه ای از جدول خرید مقالات

موضوع	ایمیل کاربر	کد مقاله	زمان خرید	تاریخ خرید	کد پیگیری
مهندسی عمران	alisaber@gmail.com	NC_25.pdf	۰۷:۴۲:۴۹	۱/۱/۱۳۸۹	۲۶۰۹۳۶
مسکن و معماری	arabi@ut.ac.ir	ICI_30.pdf	۰۸:۵۹:۵۳	۷/۱/۱۳۸۹	۲۶۱۵۵۴
برق و کامپیوتر	eslami@yahoo.com	PC_05.pdf	۰۲:۲۱:۰۸	۷/۱/۱۳۸۹	۲۶۳۴۲۹

### ۶-۲-۳- کاوش اولیه و اعتبارسنجی کیفیت داده ها

در این مرحله داده های موجود از نظر وجود داده های پرت، انتهایی و مقادیر مفقود بررسی شدند. همانطور که در جدول زیر آمده است نتایج بدست آمده نشان می دهد که تمامی رکورد های این مرحله کامل بوده و فاقد داده های پرت و مقادیر نامعتبر می باشند.

جدول ۲- نتایج بررسی اولیه داده ها برای ارزش گذاری

مشتریان

Field Title	Type	Outlier	Extreme	Complete Records
کد پیگیری	Range	0	0	100%
تاریخ خرید	Discrete	--	--	100%
زمان خرید	Range	0	0	100%
کد مقاله	Set	--	--	100%
ایمیل کاربر	Set	--	--	100%

کل آیتم های موجود می باشد، و برای هر مشتری ممکن است سبدهای خرید بسیار کمی موجود باشد، شاید فقط یکی.

ج- ناهماهنگی یا غیر یکنواختی، تنوع قابل توجهی در رفتار خرید افراد مختلف وجود دارد.

پیش پردازش و آماده سازی داده ها مهم ترین و زمان بر ترین مرحله در پروژه های داده کاوی است. تقریباً ۶۰ تا ۹۰ درصد زمان انجام یک پروژه داده کاوی صرف این مرحله می شود و ۷۵ تا ۹۰ درصد موفقیت پروژه های داده کاوی به آن بستگی دارد. [۶]

۶-۳-۱- انتخاب و ترکیب ویژگی ها

از مجموعه فیلدهای موجود در جدول سوابق خرید مشتریان فیلدهای کد پیگیری، کد مقاله و ردیف مقاله در بخش بندی و ارزش گذاری مشتریان بی تاثیر تشخیص داده شده و پس از انجام تنظیمات لازم و به منظور رفع مشکل افزونگی داده ها، از قسمت Filter منبع نسبت به حذف این ۳ فیلد اقدام شد.

۶-۳-۲- ساخت داده

به منظور بررسی حجم مقالات خریداری شده از سایت در ساعات مختلف شبانه روز، مقادیر تغییرات در زمان خرید مقالات گسسته سازی می شوند. در جدول ۳ مقادیر گسسته سازی شده نشان داده شده است.

جدول ۳- گسسته سازی تغییرات زمان خرید مقالات برای فایل سوابق خرید مشتریان

نماینده دسته	ابتدای بازه	انتهای بازه
۱	۰	۶
۲	۶	۱۲
۳	۱۲	۱۸
۴	۱۸	۲۴

۶-۲-۴- انتخاب زیر مجموعه های مطلوب داده

با توجه به بررسی داده ها در مرحله قبل همه رکورد های سوابق خرید مشتریان را برای ساخت مدل ارزش گذاری مشتریان این سایت انتخاب کرده و وارد مدل می کنیم.

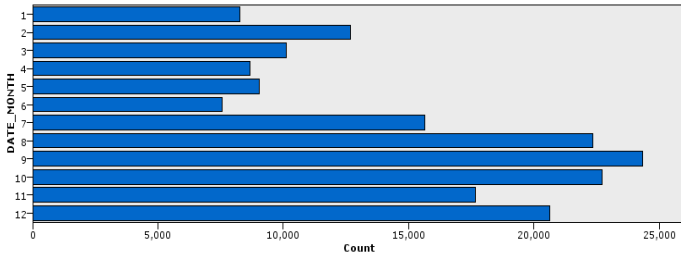
۶-۳-۳- آماده سازی داده های مشتریان

اگر چه داده های موجود در یک پایگاه داده ممکن است شامل اطلاعات کافی برای ساخت انواع مدل های داده کاوی باشد ولی همیشه در فرمی نگهداری نمی شوند که بتوانند مستقیماً به عنوان ورودی مدل ها مورد استفاده قرار گیرند. در چنین مواردی داده ها باید به منظور بدست آوردن مدل های دقیق تبدیل شوند. داده های ذخیره شده از معاملات مشتری از دیدگاه داده کاوی دارای چالش هایی می باشند که عبارتند از:

الف- حجم عظیم داده های ذخیره شده، خرده فروشی های بزرگ زنجیره ای می توانند میلیون ها تراکنش در هر روز تولید کنند.

ب- پراکندگی در داده ها، در یک فروشگاه با تعداد اقلام زیاد، یک سبد خرید معمولی تنها شامل بخش کوچکی از





شکل ۲ - هیستوگرام توزیع فروش مقالات در ماههای مختلف سال

### ۶-۳-۳- تنظیم فرمت داده ها

در فایل داده های تراکنشی، محتویات فیلد تاریخ با فرمت عددی ذخیره می شود. به منظور فعال سازی پردازش ارتباطات تاریخ در کلمنتاین باید فرمت عددی را به فرمت تاریخ تبدیل کنیم.

برای این منظور با استفاده از رابطه زیر محتویات فیلد date1 را با فرمت تاریخ جایگزین می کنیم.

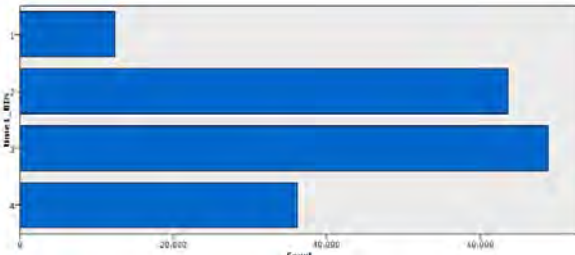
```
to_date(to_string(Date1))
```

پس از جاگذاری و تبدیل فیلد date1 به فرمت تاریخی برای

بررسی جاگذاری درست همه مقادیر از `select` و رابطه `@BLANK(date1) or @NULL(date1)`

استفاده کردیم که مشخص شد تعداد ۱۴۶۷ رکورد با مقدار null جایگزین شده اند. پس از بررسی سابقه خرید افرادی که فیلد تاریخ آنها با مقدار null جایگزین شده بود مشخص شد که همه این خریده ها در روزهای ۲۹ و ۳۰ و ۳۱ انجام شده اند و دلیل آن می تواند عدم تعریف روز ۳۱ در ماههای میلادی باشد که برای نرم افزار کلمنتاین ناشناخته است. جهت رفع این مشکل بازگشت به مرحله پاکسازی داده ضروری است. با توجه به اینکه همه داده های مذکور مربوط به روزهای ۲۹ و ۳۰ و ۳۱ این مشکل را ایجاد نمی کردند پس از بررسی های لازم تصمیم گرفته شد با توجه به حجم کم این رکوردها (حدود ۵۵ هزارم درصد از کل داده ها) که مربوط به سوابق خرید ۱۹۳ مشتری از مجموع ۴۲۷۰۹ مشتری می باشند این مجموعه داده ها با تاریخ ۰۱-۰۷-۰۷ (میانگین تاریخ های موجود) جایگزین گردند.

بررسی نمودار توزیع این فیلد (شکل شماره ۱) نشان می دهد بیشترین حجم خریده ها در ساعات ۱۲ ظهر الی ۶ بعد از ظهر بوده است.



شکل ۱- هیستوگرام توزیع فروش مقالات در ساعات مختلف شبانه روز

- نسبت به ساخت فیلد هزینه خرید با نام price اقدام شد. هدف از ایجاد این فیلد امکان بررسی متغیر هزینه در سوابق خرید مشتریان می باشد. با توجه به اینکه کاربران عضو در این سایت برای دریافت مقالات می توانند بصورت اینترنتی یا کارتهای سیویلیکا میزان اعتبار دلخواه خود را به مبلغ هر اعتبار 10000 ریال خریداری کنند و قیمت همه مقالات در این سایت برابر 10000 ریال است، مقدار این فیلد در همه رکوردها با مقدار 10000 ریال به ازای هر خرید جایگزین گردید.

- نسبت به ساخت فیلد ماه خرید مقاله با نام DATE\_MONTH با استفاده از گره Derive اقدام شد. هدف از ایجاد این فیلد بررسی توزیع حجم مقالات خریداری شده هر موضوع در هر ماه می باشد.

این فیلد با استفاده از رابطه زیر از روی مقدار فیلد تاریخ خرید در فایل سوابق خرید مشتریان ایجاد شد.

```
datetime_month(date1)
```

بررسی نمودار توزیع این فیلد (شکل شماره ۲) نشان می دهد بیشترین حجم خرید مقالات در ماههای آبان، آذر و دی بوده است. که دلیل آن می تواند شروع سال تحصیلی جدید در دانشگاهها باشد.



می باشد، در ابتدا جهت شناخت و ارزش گذاری مشتریان، داده های آماده شده در مراحل قبل به مدل تحلیل تازگی، تکرار و مبلغ (RFM) وارد شدند، سپس نتایج حاصل با استفاده از دو روش بخش بندی دستی و خوشه بندی به رده های جداگانه تقسیم شدند و در نهایت با استفاده از شبکه عصبی به ارزیابی و مقایسه کلاسهای بدست آمده در هر دو روش می پردازیم.

#### ۶-۴-۲- بخش بندی و ارزش گذاری مشتریان به روش دستی

در این بخش پس از ورود داده ها به مدل تحلیلی RFM و محاسبه ارزش هر مشتری بر اساس سوابق خرید او، به منظور بخش بندی مشتریان بر اساس ارزش آنها، پس از بررسی مقادیر بدست آمده از خروجی مدل RFM، به گسسته سازی مقادیر Recency، Frequency و Monetary بدست آمده با استفاده از تکنیک binning اقدام شد.

سپس متغیر دیگری به نام  $R^*F^*M$  تعریف کرده و مقدار آن را با استفاده از رابطه زیر برای همه مشتریان بدست می آوریم:

$$\text{Frequency\_BIN} * \text{Monetary\_BIN} * \text{Recency\_BIN}$$

نمونه ای از مقادیر بدست آمده تا این مرحله در جدول زیر (جدول شماره ۴) نشان داده شده است.

جدول ۴- نمونه مقادیر  $R^*F^*M$  بدست آمده برای مشتریان

Email	Recency	Frequency	Monetary	$R^*F^*M$
forghani@yahoo.com	76	1	10000	5
ardaghli@yahoo.com	84	63	630000	125
arash@gmail.com	70	9	90000	20

#### ۶-۳-۴- تایید کیفیت داده و ساخت فایل نهایی داده ها

با بررسی وضعیت داده ها در مراحل قبلی و استفاده از ابزار داده کاوی نسبت به رفع خطا و یا جایگزینی داده های مذکور، ملاحظه گردید که داده های اولیه کامل بوده و فاقد مقادیر تهی و نامعتبری باشد. لذا جدول نهایی بدست آمده از ساخت و آماده سازی داده های مذکور در فرمت mdb و تحت عنوان فایل TblRequestFinal استخراج گردید تا در مرحله بعدی، جهت ورود به مدل ارزش گذاری مشتریان از آن استفاده شود.

#### ۶-۴-۶- مدلسازی و اجرای مدل

در این مرحله تکنیکهای مختلف مدلسازی انتخاب و اجرا می شوند و پارامترهای آنها برای بدست آوردن مقادیر بهینه اندازه گیری می شود. عموماً تکنیکهای مختلفی برای مسائل مشابه داده کاوی وجود دارد. بعضی از تکنیکها نیازهای مشخصی در خصوص داده ها دارند بنابراین گاهی اوقات برگشت به مرحله آماده سازی داده ها ضروری است. اقداماتی که در فاز مدلسازی انجام می شود شامل مراحل زیر است:

- انتخاب و اجرای تکنیکهای مناسب مدلسازی.
- انجام تنظیمات مدل برای بهینه سازی نتایج.
- بکارگیری چندین روش مختلف داده کاوی برای حل یک مسئله داده کاوی.
- در صورت لزوم برگشت به مرحله آماده سازی داده ها جهت انجام عملیات مورد نیاز روی داده ها متناسب با نیازهای مشخص تکنیکهای داده کاوی مورد استفاده. [۱۰]

#### ۶-۴-۱- انتخاب تکنیک مدلسازی

با توجه به اینکه هدف اصلی این پروژه داده کاوی، رده بندی و ارزش گذاری مشتریان بر اساس سابقه خرید آنها



در این روش با توجه به امتیاز بدست آمده برای هر مشتری از روی حاصلضرب مقادیر گسسته Recency ، Frequency و Monetary ، و سپس گسسته سازی آن، مشتریان به ۶ گروه مشتریان خاص، طلایی، وفادار، عادی، مشکوک و رویگردان تقسیم شده و برچسب گذاری شدند.

#### ۶-۴-۳- بخش بندی و ارزش گذاری مشتریان با استفاده از تکنیکهای خوشه بندی

در این بخش پس از ورود داده ها به مدل تحلیلی RFM و محاسبه ارزش هر مشتری بر اساس سوابق خرید او، به منظور بخش بندی مشتریان، از تکنیک خوشه بندی فاصله ای K-Means استفاده نموده ایم. مشخصه های استخراج شده که به عنوان ورودی تکنیک های خوشه بندی تعریف شده اند عبارتند از:

۱. تازگی آخرین خرید (Recency): فاصله بین زمان آخرین خرید تا اکنون.
۲. تکرار خرید (Frequency): تعداد مبادلات در یک پریود زمانی مشخص

با توجه به اینکه در همه رکوردها مشخصه کل مبالغ خرید (Monetary)، همواره ضریب ثابتی از مشخصه تکرار خرید می باشد، این مشخصه بی تاثیر تشخیص داده شد و به منظور جلوگیری از افزونگی داده ها از ورودی تکنیک های خوشه بندی حذف گردید.

تکنیک خوشه بندی فاصله ای K-Means تعداد خوشه بهینه را از داده کاو دریافت می کند، در اینجا با توجه به کلاسه های از پیش در نظر گرفته برای ارزش گذاری مشتریان تعداد خوشه ها برابر ۶ در نظر گرفته شده است. نتایج خروجی خوشه بندی با این الگوریتم در جدول شماره ۷ آمده است.

جدول شماره ۷ - میانگین مشخصه های هر خوشه تکنیک K-

Means

خوشه	خوشه	خوشه	خوشه	خوشه	خوشه	مشخصه
------	------	------	------	------	------	-------

#### ۶-۴-۲-۱- برچسب گذاری مشتریان در روش اول

برای بخش بندی و برچسب گذاری مشتریان بر مبنای سوابق خرید آنها در این روش، از گسسته سازی متغیر R\*F\*M با استفاده از روابط جدول شماره ۵ استفاده می کنیم. در نتیجه فیلد جدیدی به نام CUSTOMER VALUE جهت ارزش گذاری مشتریان به جدول اضافه شد.

#### جدول ۵- گسسته سازی تغییرات R\*F\*M جهت بخش بندی و برچسب گذاری مشتریان

نماینده دسته	تعداد	درصد	ابتدای بازه	انتهای بازه	CUSTOMER VALUE
۱	62	0.15%	۱۰۰	۱۲۵	مشتریان طلایی
۲	۳۵	0.08%	۱۲۵	∞	مشتریان خاص
۳	39094	91.54%	۱	۲۵	مشتریان رویگردان
۴	196	0.46%	۵۰	۷۵	مشتریان عادی
۵	179	0.42%	۷۵	۱۰۰	مشتریان وفادار
۶	3143	7.36%	۲۵	۵۰	مشتریان مشکوک

نمونه ای از برچسب های بدست آمده برای مشتریان در این روش، در جدول زیر نشان داده شده است.

#### جدول ۶- نمونه برچسب های بدست آمده جهت ارزش گذاری مشتریان در روش اول

Email	Recency	Frequency	Monetary	R*F*M	CUSTOMER VALUE
forghni@yahoo.com	76	1	10000	5	رویگردان
hardali@yahoo.com	84	63	630000	125	طلایی
arash@gmail.com	70	9	90000	20	رویگردان

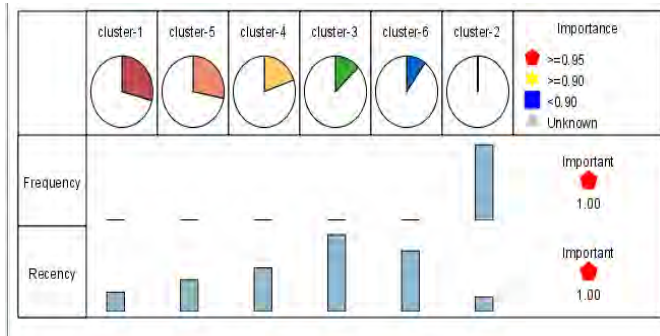
آخرین خرید آنها در مقایسه با سایر خوشه ها (به جز خوشه ۲ که فقط دارای یک مشتری می باشد) از همه کوتاهتر است. با توجه به اینکه افرادی که در این خوشه هستند فرکانس خرید معقول داشته و زمان سپری شده از آخرین خرید آنها کم است بالاترین اهمیت را برای سیستم فروش اینترنتی خواهند داشت. در نتیجه به این گروه از مشتریان برچسب مشتریان طلایی را نسبت می دهیم.

خوشه ۲: خوشه ۲ که شامل ۱ مشتری، یعنی ۰.۰۰۲٪ از کل مشتریان را دارا می باشد. با توجه به مختصات نقاط مرکزی این خوشه، اعضای این خوشه دارای فرکانس خرید بسیار بالا و غیرعادی می باشند. به این ترتیب این افراد کسانی هستند که با فرکانس خرید بسیار بالا و غیرعادی یک مشتری معمول سیستم نبوده و جزو کاربران عادی وب سایت نمی باشند. این افراد مسئولان کتابخانه ها یا کافی نت ها می باشند که با ایجاد یک حساب کاربری برای اعضای کتابخانه و یا مشتریان خود به خرید مقالات از سایت اقدام می کردند. در نتیجه به این گروه از مشتریان برچسب مشتریان خاص را نسبت می دهیم.

خوشه ۳: خوشه ۳ شامل ۵۶۰۵ مشتری بوده و ۱۳٪ از کل مشتریان را دارا می باشد. با توجه به مختصات نقاط مرکزی این خوشه، اعضای این خوشه دارای پایین ترین فرکانس خرید بوده و زمان زیادی از آخرین خرید آنها گذشته است. این افراد کسانی هستند که در دوره ای مشتری سیستم بوده اند اما بعد از تعداد محدودی خرید اینترنتی از ادامه خرید منصرف شده و سیستم را ترک کرده اند. در نتیجه به این گروه از مشتریان برچسب مشتریان رویگردان را نسبت می دهیم.

	۱	۲	۳	۴	۵	۶
تعداد	۱۲۴۱	۱	۵۶۰۵	۸۴۱۹	۱۱۹۶	۴۳۰۷
درصد	۲۹٪	۰.۰۰۲٪	۱۳٪	۱۹.۷٪	۲۸٪	۱۰٪
Recency	۹۷.۵۰	۷۴	۳۸۷.۵	۲۲۰	۱۶۰	۳۰.۵
Frequenc y	۶.۱۶	۵۷۱۶	۲.۲	۳.۱۵	۳.۸	۳.۲

شکل ۳ خوشه های تولید شده توسط تکنیک K-Means را نشان می دهد.



شکل ۳- خوشه های تولید شده توسط تکنیک K-Means

### ۶-۴-۱- نتایج بخش بندی مشتریان با تکنیک K-Means

با فرض اینکه نقطه مرکزی هر خوشه نماینده کل نقاط هر خوشه است، ویژگیهای مشتریان خوشه ها تحلیل شد. در ادامه هر یک از خوشه های تولید شده با تکنیک K-Means بر اساس داده های بدست آمده از جدول .. تحلیل می گردد:

۱. خوشه ۱: خوشه ۱ شامل ۱۲۴۱۴ مشتری بوده و ۲۹٪ از کل مشتریان را دارا می باشد. این خوشه پرجمعیت ترین خوشه نسبت به سایر خوشه های موجود می باشد. با توجه به مختصات نقاط مرکزی خوشه ها، خوشه ۱ در مقایسه با سایر خوشه ها از فرکانس خرید بالا و قابل قبول برخوردار می باشد و مدت زمان سپری شده از

۴. خوشه ۴: خوشه ۴ شامل ۸۴۱۹ مشتری بوده و ۱۹.۷٪ از کل مشتریان را دارا می باشد. با توجه به مختصات نقاط مرکزی این خوشه، این مشتریان دارای فرکانس خرید معمولی بوده و زمان سپری شده از آخرین خرید آنها نیز نرمال می باشد. در نتیجه به این گروه از مشتریان برچسب مشتریان عادی را نسبت می دهیم.

جدول ۸- برچسب گذاری مشتریان هر خوشه جهت بخش بندی و ارزش گذاری مشتریان

شماره خوشه	برچسب
۱	مشتریان طلایی
۲	مشتریان خاص
۳	مشتریان رویگردان
۴	مشتریان عادی
۵	مشتریان وفادار
۶	مشتریان مشکوک

۵. خوشه ۵: خوشه ۵ شامل ۱۱۹۶۳ مشتری بوده و ۲۸٪ از کل مشتریان را دارا می باشد. با توجه به مختصات نقاط مرکزی این خوشه، اعضای این خوشه دارای فرکانس خرید قابل قبولی بوده و زمان سپری شده از آخرین خرید آنها نیز نسبتاً کم می باشد. در نتیجه به این گروه از مشتریان برچسب مشتریان وفادار را نسبت می دهیم.

نمونه ای از برچسب های بدست آمده برای مشتریان در این روش، در جدول شماره ۹ نشان داده شده است.

جدول ۹- نمونه برچسب های بدست آمده جهت ارزش گذاری مشتریان در روش دوم

Email	Recency	Frequency	Monetary	customer value
forghani@yahoo.com	76	1	10000	طلایی
mohammad@yahoo.com	۲۲۰	۴	۴۰۰۰۰	عادی
ali@gmail.com	۱۳۱	۱۳	130000	وفادار

۶. خوشه ۶: خوشه ۶ شامل ۴۳۰۷ مشتری بوده و ۱۰٪ از کل مشتریان را دارا می باشد. با توجه به مختصات نقاط مرکزی این خوشه، اعضای این خوشه دارای فرکانس خرید نسبتاً کم بوده و زمان سپری شده از آخرین خرید آنها نیز نسبتاً زیاد می باشد. بنابراین نمی توان بطور قطعی نوع آنها را تشخیص داد و باید با گذشت زمان و بر بیشتر رفتار خرید آنان شناخت بیشتری نسبت به این افراد بدست آورده و نوع آنها را تمایز داد. در نتیجه به این گروه از مشتریان برچسب مشتریان مشکوک را نسبت می دهیم.

۶-۴-۴ - مقایسه و بررسی دقت مدل های ساخته شده برای برچسب گذاری مشتریان با استفاده از شبکه عصبی

به منظور مقایسه دقت مدل های ساخته شده برای برچسب گذاری مشتریان در دو روش دستی و خوشه بندی، با استفاده از تکنیک شبکه های عصبی به پیش بینی ارزش بدست آمده برای هر مشتری در دو روش فوق پرداختیم و

۶-۴-۳-۲- برچسب گذاری مشتریان در روش دوم  
با توجه به مشخصات خوشه ها و تحلیل های انجام شده در مرحله قبل، برچسب های زیر برای ارزش مشتریان هر خوشه در نظر گرفته شد و بر اساس روابط جدول شماره ۸ فیلد جدیدی به نام customer value جهت برچسب گذاری مشتریان اضافه شد.

می باشد، به همین دلیل نتیجه می گیریم که استفاده از تکنیکهای خوشه بندی به منظور بخش بندی مشتریان روش مناسب تری برای جداسازی مشتریان می باشد.

#### ۵-۶- ارزیابی مدل ها

در فاز ارزیابی نیز، مدل یا مدلی که بهترین کیفیت را در تحلیل داده دارند ساخته می شوند. قبل از اجرای مدل نهایی فاز ارزیابی بسیار دارای اهمیت است تا اینکه به اهداف کسب و کار دست یابیم. در پایان این مرحله تصمیماتی در خصوص استفاده از نتایج داده کاوی گرفته می شود. اقدامات این فاز شامل مراحل زیر است:

- ارزیابی مدل های تولید شده قبل از استقرار نهایی آنها.
  - بررسی دستیابی مدل به مجموعه اهداف تعیین شده در فاز اول CRISP-DM.
  - بررسی اینکه برخی از بخشهای مهم کسب و کار یا مسئله تحقیق به اندازه کافی بررسی و محاسبه نشده است.
  - تصمیم گیری جهت استفاده از نتایج داده کاوی.
- [۱۰]

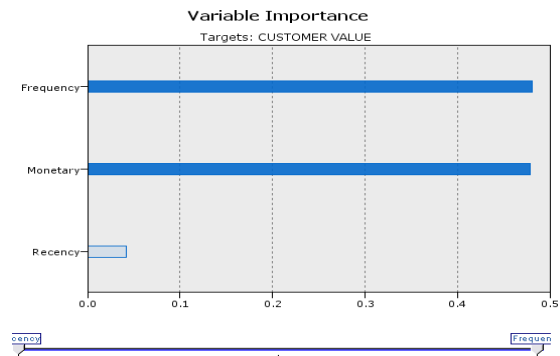
به منظور ارزیابی مدل های ساخته شده در این پژوهش، از تقسیم داده ها به دو مجموعه آموزشی و تست استفاده شد. ۷۰ درصد داده ها بصورت آموزشی و ۳۰ درصد بصورت داده های آزمایشی وارد مدل های ساخته شده در هر دو روش شدند. جداول ۱۱ و ۱۲ نتایج بدست آمده از ارزیابی مدل ها در هر دو روش را نشان می دهند.

جدول ۱۱- نتایج بدست آمده از ارزیابی مدل پیش بینی ارزش

مشتریان در روش اول

داده های آموزشی	درصد داده های آموزشی	داده های آزمایشی	درصد داده های آزمایشی
۲۷۷۹۲	۹۳.۲۲٪	۱۱۹۸۳	۹۲.۹۱٪
۲۰۲۰	۶.۷۸٪	۹۱۴	۷.۰۹٪

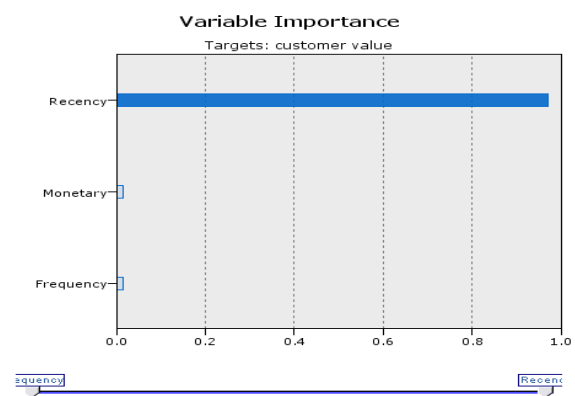
دقت بدست آمده برای شبکه عصبی در هر دو روش را با هم مقایسه کردیم. دقت بدست آمده در هر دو روش در جدول ۱۰ و نمودار اهمیت متغیرها در اشکال ۴ و ۵ نشان داده شده است.



شکل ۴- نمودار اهمیت متغیرها در روش اول

جدول ۱۰ - مقایسه دقت بدست آمده از اجرای شبکه عصبی در روش های اول و دوم

مشخصه	روش اول (بخش بندی دستی مشتریان با گسسته سازی R*F*M)	روش دوم (بخش بندی مشتریان با تکنیکهای خوشه بندی)
دقت مدل	۹۷.۷۴	۹۹.۲۷
تعداد لایه ورودی	۳	۳
تعداد لایه خروجی	۷	۷
تعداد لایه پنهان	۱	۱



شکل ۵ - نمودار اهمیت متغیرها در روش دوم

همانطور که مشاهده می شود اجرای شبکه عصبی برای پیش بینی ارزش مشتری در روش دوم دارای دقت بالاتری



دانشگاه پیام نور استان تهران

## مجموعه مقالات اولین همایش تخصصی

### "سیستهای هوشمند کامپیوتری و کاربردهای آنها"

دانشگاه پیام نور استان تهران - مجتمع فنی و مهندسی (آذرماه ۱۳۹۰)

در مورد استفاده از نتایج و تعیین استراتژی های بازاریابی در مورد گروههای مختلف مشتریان موجود تصمیم گیری نماید.

جمع کل	۲۹۸۱۲	۱۲۸۹۷	
--------	-------	-------	--

### ۷- نتیجه گیری

بخش بندی مشتریان به وسیله تکنیکهای خوشه بندی یکی از کاراترین روش ها برای شناخت ویژگی های رفتاری مشتریان به منظور مدیریت ارتباط با مشتریان می باشد. در این مقاله به منظور بخش بندی مشتریان از مدل تحلیل تازگی، تکرار و مبلغ استفاده شد، سپس مشتریان بر اساس ارزش بدست آمده از سابقه خریدشان با استفاده از دو روش بخش بندی دستی و همچنین با استفاده از تکنیکهای خوشه بندی بخش بندی شدند. بررسی و مقایسه دقت اجرای شبکه عصبی به منظور پیش بینی ارزش بدست آمده برای مشتری در هر دو روش، نشان داد که خوشه بندی روش مناسب تری برای بخش بندی مشتریان به گروههای دلخواه می باشد.

در این تحقیق مشتریان فعلی سیستم فروش اینترنتی مقالات، بر اساس سوابق خرید موجود در پایگاه داده فروش مقالات بخش بندی و ارزش گذاری شده و ۶ خوشه کشف گردید. برچسب گذاری مشتریان در خوشه های مختلف بر اساس ارزش آنها، امکان تعریف استراتژی های مناسب بر اساس ویژگی های هر خوشه را فراهم کرده است. این تحقیق یک ایده کلی در رابطه با استفاده از تکنیکهای داده کاوی برای بخش بندی و ارزش گذاری مشتریان در اختیار می گذارد.

با استفاده از نتایج این تحقیق مدیران در تدوین استراتژیهای بازاریابی دچار سردرگمی نمی شوند، از طرف دیگر دیدی کلی از تخصیص مناسب بودجه های بازاریابی و مدیریت ارتباط با مشتری پیدا می کنند. در کارهای آینده می توان با استفاده از روش های یادگیری ماشین به بررسی تغییرات دراز مدت ارزش مشتری پرداخت. مهمترین هدف

جدول ۱۲- نتایج بدست آمده از ارزیابی مدل پیش بینی ارزش مشتریان در روش دوم

داده های آموزشی	درصد داده های آموزشی	داده های آزمایشی	درصد داده های آموزشی
۲۹۵۸۹	۹۹.۲۵ %	۱۲۷۹۳	۹۹.۱۹ %
۲۲۳	۰.۷۵ %	۱۰۴	۰.۸۱ %
جمع کل	۲۹۸۱۲	۱۲۸۹۷	

همانطور که جداول فوق نشان می دهند مدل ساخته شده برای پیش بینی ارزش مشتریان در روش دوم که از تکنیکهای خوشه بندی برای بخش بندی مشتریان استفاده کرده بود دارای دقت بالاتری می باشد.

### ۶-۶- تحلیل نتایج و ارائه دانش بدست آمده

ساخت مدل معمولا پایان کار داده کاوی نیست حتی اگر اهداف مدل افزایش دانش راجع به داده ها باشد، دانش کشف شده باید سازماندهی شده و به روش قابل استفاده برای کاربران ارائه شود. مواردی که باید به آن توجه داشت عبارتند از:

- استفاده از مدل تولید شده، توجه داشته باشید که ایجاد مدل به معنای اتمام پروژه نمی باشد.
- تهیه گزارشات لازم.
- اجرا و استقرار فرایند داده کاوی در یک بخش دیگر بصورت موازی. [۱۰]

در این پژوهش نتایج حاصل از مدل ارزش گذاری مشتریان ، بصورت یک گزارش از کل کار در اختیار شرکت مورد مطالعه قرار می گیرد تا شرکت بر اساس استراتژی های خود



LUKASZ A.KURGAN and PETR MUSILEK, "A survey of Knowledge Discovery and Data Mining process models", The Knowledge Engineering Review, Vol 21:1, 1-24, 2006, Cambridge University Press

[9]

Ngai, E.W.T, Xiu, Li, Chau, D.C.K. (2009), "Application of data mining techniques in customer relationship management: A literature review and classification", Expert Systems with Applications 36 (2009) 2592-2602

[10]

T.Larose, Daniel, 2005, "DISCOVERING KNOWLEDGE IN DATA: an introduction to data mining", JOHN WILEY publishers

[11]

Xin-an Lai, "Segmentation Study on Enterprise Customers Based on Data Mining Technology", IEEE/2009

[12]

Yen-Liang Chen, Mi-Hao Kuo, Shin-Yi Wub, Kwei Tang, "Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data", Electronic Commerce Research and Applications (2009)

[13]

Y. Qu, X. Yang, & T. Huang (5 May 2000). Survey of Recommendation System and Algorithms. EE 380L: DataMining.

در مدیریت ارتباط با مشتری، افزایش سود دراز مدت مشتری می باشد.

## سپاسگزاری

در پایان از جناب آقای مهندس وجودی به جهت همکاری در انجام این تحقیق کمال تشکر را دارم.

## مراجع

[1]

امتیاز، س.، کیوان پور، م.، و پناه، ا.، "دسته بندی تطبیقی روش های وب کاوی و حوزه مدیریت ارتباط با مشتری"، دومین کنفرانس داده کاوی، تهران، تابستان ۸۷

[2]

رستگار، نگار، پایان نامه کارشناسی ارشد: "بخش بندی مشتریان مبتنی بر مدل ترکیبی RFM و LTV با استفاده از روش های هیوربستیک - مورد مطالعاتی: شرکت ارتباطات سیار(همراه اول)"، دانشگاه علم و صنعت ایران - اسفند ۱۳۸۶

[3]

شاهسمندی، پ.، "داده کاوی در مدیریت ارتباط با مشتری"، مجله تدبیر شماره ۱۵۶، ۱۳۸۴.

[4]

تقوا، م.، الفت، ل.، و حسینی، م.، "بکارگیری تکنیک های داده کاوی جهت بهبود مدیریت ارتباط با مشتری در صنعت بانکداری"، دومین کنفرانس داده کاوی، تهران، تابستان ۸۷

[5]

مینایی بیدگلی، ب.، قهرمانی، ف.، و داوری، م.ر.، "داده کاوی برای مدیریت ارتباط با مشتری در یک شرکت هواپیمایی"، دومین کنفرانس داده کاوی، تهران، تابستان ۸۷

[6]

Han, J. and Kamber, M., "Data Mining: Concepts and Techniques", Second Edition, Morgan Kaufman Publisher, 2006.

[7]

Hyunseok Hwang, Taesoo Jung, Euiho Suh, "An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry", Expert Systems with Applications 26 (2004) 181-188

[8]





مجموعه مقالات اولین همایش تخصصی  
"سیستهای هوشمند کامپیوتری و کاربردهای آنها"  
دانشگاه پیام نور استان تهران - مجتمع فنی و مهندسی (آذرماه ۱۳۹۰)

<sup>a</sup>Analytic CRM  
<sup>b</sup> Recency, Frequency, Monetary  
<sup>c</sup> Hughes  
<sup>d</sup> Data Oriented