



## تحلیل سبد بازار و بهینه سازی مدل پیشنهادی برای بسته‌های تخفیفی ساخته شده بر اساس قواعد تلازمی مطالعه موردی در فروشگاه زنجیره‌ای

نیما شریعتی<sup>۱</sup>

حسن شوندی<sup>۲</sup>

سمیه علیزاده<sup>۳</sup>

تاریخ دریافت: ۹۰/۱۰/۲۰

تاریخ پذیرش: ۹۰/۱/۲۰

### چکیده

انتخاب سبد بازار یک فرآیند تصمیم‌گیری است که مشتری اقلامی را از میان تعدادی کالا در فرآیند خرید انتخاب می‌کند. وابستگی‌های درونی در ارتباطات تقاضا در میان اقلام، مشخصه‌ای کلیدی در انتخاب سبد بازار است. در این مقاله با بکارگیری یکی از روش‌های فرآیندهای داده‌کاوی در تحلیل سبد بازار که قواعد تلازمی نام دارد، به تعیین سبد بازار در فروشگاه‌های زنجیره‌ای با جامعه آماری بزرگ از مشتریان پرداخته شده است. نتیجه این فرآیند داده‌کاوی ارائه مدلی جهت استخراج قواعد تلازمی، ارائه مدلی برای دستیابی به مقادیر بهینه تخفیف در بسته‌های پیشنهادی تخفیفی و بدست‌آمده براساس قواعد تلازمی و نیز تبیین استراتژی‌های دیگر جهت بهبود عملکرد، افزایش رضایتمندی مشتری و سودآوری در این فروشگاه‌ها خواهد بود.

**واژه‌های کلیدی:** داده‌کاوی، سبد بازار، قواعد تلازمی، بسته تخفیفی، فروشگاه زنجیره‌ای.

### ۱- مقدمه

۱- کارشناس ارشد صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی؛  
shariatinima@yahoo.com

۲- دانشیار دانشکده صنایع، دانشگاه صنعتی شریف؛ shavandi@sharif.edu

۳- استادیار دانشکده صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی؛  
s\_alizadeh@kntu.ac.ir

در دو دهه قبل توانایی‌های فنی بشر برای تولید و جمع‌آوری داده‌ها به سرعت افزایش یافته است. عواملی نظیر استفاده گسترده از بارکد برای تولیدات تجاری، به خدمت گرفتن کامپیوتر در کسب و کار، علوم، خدمات دولتی و پیشرفت در وسائل جمع‌آوری داده در این تغییرات نقش مهمی دارند. بطور کلی استفاده همگانی از وب و اینترنت به عنوان یک سیستم اطلاع‌رسانی جهانی ما را با حجم انبوهی از داده‌ها و اطلاعات مواجه می‌کند. داده‌کاوی از پایگاه داده‌های بزرگ برای کشف اطلاعات ارزشمند استفاده می‌کند. از جمله حوزه‌هایی که علم داده‌کاوی بسیار موثر بوده است، تحلیل سبد بازار می‌باشد. پیشرفت‌های فناوری، فروشگاه‌های خرده‌فروشی را قادر ساخته است تا حجم زیادی از داده‌های خرید مشتریان را جمع‌آوری و ذخیره نمایند. در فروشگاه‌های بزرگ از جمله فروشگاه‌های زنجیره‌ای، همواره مدیریت ارتباط با مشتری یک موضوع و حوزه بسیار مهم در تصمیم‌گیری مدیران ارشد بوده است. شناسایی و تحلیل سبد بازار و یا سبد خرید، همواره به عنوان یک راه حل مناسب در دستیابی به ارتباط پایدار، موثر و رضایت بخش برای ذینفعان که در واقع مدیران فروشگاه‌های زنجیره‌ای و مشتریان می‌باشند، مطرح گردیده است. هر مشتری خرید مجزایی را در مقادیر مختلف و زمان‌های متفاوت انجام می‌دهد و داده‌های موجود در سبد بازار، نشان‌دهنده خرید مشتری در یک زمان خاص است. با تجزیه و تحلیل سبد بازار، فروشندگان می‌توانند رفتار خرید مشتریان را پیش‌بینی کنند و از این طریق سودآوری خود را افزایش دهند. در تعیین سبد بازار به کمک علم داده‌کاوی از روش‌های مختلف قواعد تلازمی استفاده می‌شود. استخراج قواعد تلازمی یا با هم آیی، نوعی عملیات داده‌کاوی است که به جستجو برای یافتن ارتباط بین ویژگی‌ها در مجموعه داده‌ها می‌پردازد. در این مقاله با توجه به ضرورت زیاد تعیین سبد بازار در فروشگاه‌های زنجیره‌ای و مزایای رقابتی حاصل از آن، از قواعد تلازمی در دستیابی به اهداف سازمان استفاده شده است. سپس رویه‌ای برای ایجاد بسته‌های تخفیفی که بر مبنای قواعد تلازمی به دست آمده‌اند پیشنهاد شده و مدلی پیشنهادی برای یافتن مقدار بهینه‌ی تخفیف به کار گرفته شده است. همچنین استراتژی‌های برای ساماندهی و چیدمان بهتر محصولات نیز ارائه شده است. اهداف انجام این فرآیند داده‌کاوی در پاسخگویی به سوالات ذیل گنجانده شده است:

(۱) کدام اقلام با هم در یک سبد خرید قرار می‌گیرند؟

۲) بسته ی تخیفی بهینه چگونه تشکیل می شود.  
۳) چیدمان مناسب اقلام که به بهترین نحوه نیاز های مشتریان را تامین کند و موجب کاهش زمان سفر خرید آنان شود چگونه است؟  
۴) استراتژی های مناسب در بهبود عملکرد فروشگاه کدامند؟ و چگونه می توان با استفاده از قواعد بدست آمده به افزایش سود و رضایتمندی ذینفعان کمک کرد؟  
در این مقاله انجام فرآیند داده کاوی و استخراج قواعد تلازمی در شعبه ای از یک فروشگاه زنجیره ای که جامعه آماری بزرگی از مشتریان را به خود اختصاص می دهد و ضرورت پاسخگویی به سوالات فوق با توجه به گستردگی مجموعه آن، حجم بالای تراکنش، نیاز مبرم برای بهبود عملکرد و ایجاد مزیت رقابتی به شدت در آنجا احساس می شود، انجام گرفته است. در ادامه مروری بر قواعد تلازمی و پیشینه موضوعی تحقیق خواهیم داشت.

## ۲- مبانی نظری و پیشینه تحقیق

### ۲-۱- قواعد تلازمی<sup>۱</sup>:

استخراج قواعد تلازمی یا با هم آیی نوعی عملیات داده کاوی است که به جستجو برای یافتن ارتباط بین ویژگی ها در مجموعه داده ها می پردازد. نام دیگر روش تحلیل تلازمی، تحلیل سبد بازار می باشد؛ به عبارت دیگر، تحلیل تلازمی، مطالعه ویژگی ها یا خصوصیات می باشد که با یکدیگر همراه بوده و به دنبال استخراج قواعد به منظور کمی کردن ارتباط میان دو یا چند ویژگی است. به طور کلی قواعد تلازمی اقلام مکرری که در پایگاه داده تراکنش ها با هم می آیند را به شکل اگر - آنگاه و به همراه دو معیار پشتیبان<sup>۲</sup> و اطمینان<sup>۳</sup> که به ترتیب مجموعه اقلام مکرر و قواعد تلازمی را کشف می کند، مورد بررسی قرار می دهد. همان طور که اشاره شد، یکی از پرکاربردترین حالت های تحلیل قواعد تلازمی، تجزیه و تحلیل سبد بازار است.

### ۲-۲- پیشینه تحقیق

براساس ویفت (۲۰۰۱) و پارواتیار و شت (۲۰۰۱)، مدیریت ارتباط با مشتری از چهار بعد تشکیل شده است: (۱) شناسایی مشتری، (۲) جذب مشتری، (۳) حفظ مشتری و (۴)

توسعه مشتری. هفت نوع مدلی که بطور معمول در حوزه ارتباط با مشتری به آنها اشاره شده است شامل: تلازمی، دسته بندی، خوشه بندی، پیش بینی، رگرسیون، کشف توالی و مصور سازی می باشد.

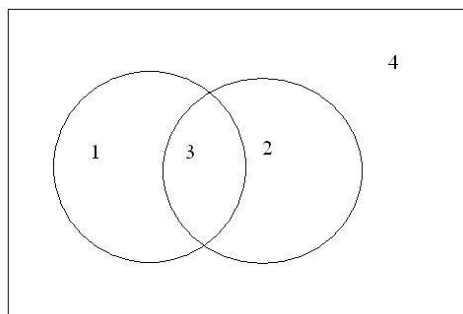
در حوزه مدیریت ارتباط با مشتری، قواعد تلازمی بیشتر روش های آماری و الگوریتم های اپریوری (اگراول، ۱۹۹۴) را شامل می شود. در این حوزه از سال ۲۰۰۰ تا ۲۰۰۶، ۱۸ مقاله در ژورنال های معتبر به انتشار رسیده است (نگی و همکاران ۲۰۰۹). ساواسر و همکاران (۱۹۹۵) الگوریتم بخش بندی را برای قواعد تلازمی گسترش داده اند. از جمله گسترش هایی که در سال های اخیر در حوزه آکادمیک در بحث قواعد تلازمی به وجود آمده است می توان به الگوریتم های بهبود (هن و همکاران، ۲۰۰۰؛ لیو و همکاران، ۲۰۰۲)، قوانین فازی (اشیبوچی و یاموتو، ۲۰۰۱؛ کواک و همکاران، ۱۹۹۸)، قوانین عمومی و چند سطحی (کلمنتینی و همکاران، ۲۰۰۰؛ هن و فو، ۱۹۹۹)، قوانین کمی (پارک و همکاران، ۱۹۹۷؛ سربینکت و اگراول، ۱۹۹۶؛ ویجسن و مرسمن، ۱۹۹۸) و قوانین فضایی (کلمنتینی و همکاران، ۲۰۰۰؛ کپرسکی و هن، ۱۹۹۵) و نیز الگوریتم های قوانین علاقه مند، قوانین درون تراکنشی، قوانین تلازمی موقت و مدل های بهینه سازی ترکیبی اشاره کرد. قوانین تلازمی در حوزه ارتباط با مشتری بیشتر در حوزه حفظ مشتری (بازار یک به یک و حوزه توسعه مشتری)، تحلیل سبد بازار و افزایش فروش کاربرد داشته است (نگی و همکاران ۲۰۰۹).

### ۳- مدل تحقیق تعیین بسته های تخفیفی

نحوه دستیابی به قواعد تلازمی و شناسایی کالاهایی که بایستی بر اساس این قواعد برای قرار گیری در بسته های تخفیفی انتخاب شوند در بخش ۴ مفصلاً شرح داده شده اند. هدف این بخش قرار دادن زوج کالای مذکور در بسته های تخفیفی و تعیین میزان بهینه ی تخفیف به نحوی است که با تأثیر گذاری بر احتمال خرید توأم زوج کالا، امید ریاضی سود سالانه حداکثر شود. بدین منظور نمودار نشان داده شده در شکل (۱) در نظر گرفته می شود. فرض می شود دو کالایی که باید در یک بسته ی تخفیفی قرار گیرند، کالاهای  $X$  و  $Y$  می باشند. از لحاظ خرید یا عدم خرید این دو کالا، مشتریان به چهار گروه مطابق شکل (۱) تقسیم بندی می شوند.

- افرادی که فقط کالای  $X$  را خریداری می کنند.

- افرادی که فقط کالای Y را خریداری می کنند.
- افرادی که هر دو کالای X و Y را خریداری می کنند.
- افرادی که هیچ یک از دو کالای X و Y را خریداری نمی کنند.



شکل (۱): نمودار رفتار مشتریان در برابر خرید یا عدم خرید کالاهای X و Y

در واقع هدف از اعمال تخفیف بر بسته‌ی خرید مشتمل بر X و Y، تغییر در احتمال وقوع ناحیه ۳ به نحوی است که امید ریاضی سود سالانه حداکثر شود. بدین منظور ابتدا باید احتمال وقوع هر یک از چهار حالت مستقل فوق را محاسبه نمود. این احتمالات به ترتیب با  $P_1, P_2, P_3, P_4$  نمایش داده می‌شوند و به صورت زیر محاسبه می‌گردند:

$$P_1 = Sup(X) - Sup(X, Y) = Sup(X)(1 - Conf(X \Rightarrow Y)) \quad (۱)$$

$$P_2 = Sup(Y) - Sup(X, Y) = Sup(Y)(1 - Conf(Y \Rightarrow X)) \quad (۲)$$

$$P_3 = Sup(X \cap Y) = Sup(X) \times Conf(X \Rightarrow Y) = Sup(Y) \times Conf(Y \Rightarrow X) \quad (۳)$$

$$P_4 = 1 - P_1 - P_2 - P_3 \quad (۴)$$

در گام بعدی و پس از محاسبه‌ی احتمالات وقوع هر یک از چهار ناحیه‌ی فوق، سود حاصل از اعمال تخفیف به ازای خرید هر واحد کالا برای هر یک از این نواحی محاسبه می‌شود. بدین منظور نماد گذاری ذیل مورد استفاده قرار می‌گیرد:

$UB_{i \rightarrow j}$  سود واحد کالای ناشی از انتقال یک مشتری از گروه i به j

$UB_i$  سود حاصل از فروش هر واحد کالای نوع i

قیمت فروش هر واحد کالای نوع  $i$   $Pr_i$

درصد تخفیف بر روی بسته‌ی کالای شامل  $i$  و  $j$   $D_{i,j}$

در این مدل فرض می‌شود که اعمال تخفیف، موجب انتقال مشتریان از گروه‌های ۱، ۲ و ۴ به گروه ۳ می‌شود ولی تأثیری بر نحوه‌ی خرید مشتریان گروه ۳ نداشته و نیز منجر به انتقال مشتریان بین گروه‌های ۱ و ۲ و ۴ نیز نمی‌شود. با این توضیح سود حاصل از اعمال تخفیف و انتقال مشتریان از گروه‌های مبدأ (۱، ۲ و ۴) به گروه مقصد (۳) به صورت زیر محاسبه می‌شود:

$$UB_{1 \rightarrow 3} = UB_Y - (Pr_X + Pr_Y)D_{X,Y} \quad (5)$$

$$UB_{2 \rightarrow 3} = UB_X - (Pr_X + Pr_Y)D_{X,Y} \quad (6)$$

$$UB_{4 \rightarrow 3} = UB_X + UB_Y - (Pr_X + Pr_Y)D_{X,Y} \quad (7)$$

$$UB_{3 \rightarrow 3} = -(Pr_X + Pr_Y)D_{X,Y} \quad (8)$$

به صورت شهودی قابل فهم است که اعمال تخفیف موجب تغییر در احتمال خرید مشتریان گروه ۳ نمی‌شود، لذا اعمال تخفیف در مورد گروه ۳ زیان ده است.

در گام بعدی تأثیر تخفیف بر احتمال انتقال مشتریان از گروه‌های مبدأ به گروه مقصد بررسی می‌شود. از آنجا که انتقال یا عدم انتقال یک مشتری، یک متغیر دو دویی است که دو مقدار صفر و یک را اتخاذ می‌کند (یک: انتقال، صفر: عدم انتقال)، لذا می‌توان از مدل رگرسیون لجستیک<sup>۴</sup> جهت تعیین احتمال انتقال از گروه مبدأ به گروه مقصد به صورت تابعی از میزان تخفیف استفاده نمود. به عبارت دیگر اگر  $P_{i \rightarrow j}$  احتمال انتقال از گروه  $i$  به گروه  $j$  باشد، این احتمال به صورت زیر محاسبه می‌شود:

$$P_{i \rightarrow j} = \frac{\exp(\beta_0 + \beta_1 D_{X,Y})}{1 + \exp(\beta_0 + \beta_1 D_{X,Y})} \quad (9)$$

که در آن  $\beta_0, \beta_1$  شیب و عرض از مبدأ پیش بینی کننده‌ی خطی می‌باشند. با در نظر گرفتن رابطه (۹)، می‌توان تابع متوسط سود سالانه (EAB<sup>۵</sup>) ناشی از اعمال تخفیف را به صورت رابطه (۱۰) محاسبه نمود:

$$EAB(D) = ASN [P_1 P_{1 \rightarrow 3} UB_{1 \rightarrow 3} + P_2 P_{2 \rightarrow 3} UB_{2 \rightarrow 3} + P_3 UB_{3 \rightarrow 3} + P_4 P_{4 \rightarrow 3} UB_{4 \rightarrow 3}] \quad (10)$$

که در آن ASN<sup>\*</sup> متوسط کل تعداد خرید در سال است.

برای به دست آوردن مقدار بهینه ی تخفیف، تابع نشان داده شده در رابطه (۱۰) باید بیشینه شود. اما علاوه بر این تعدادی محدودیت نیز در مدل وجود دارد. این محدودیت ها از این امر ناشی می شوند که مقدار تخفیف باید به گونه ای باشد که پیوستن هر یک از افراد از گروه های مبدأ به گروه ۳، سود آور باشد. لذا برای هر یک از گروه های ۱، ۲ و ۴ محدودیت های زیر باید به مدل اضافه گردند:

$$UB_Y - (Pr_X + Pr_Y) D_{X,Y} > 0 \quad (11)$$

$$UB_X - (Pr_X + Pr_Y) D_{X,Y} > 0 \quad (12)$$

$$UB_X + UB_Y - (Pr_X + Pr_Y) D_{X,Y} > 0 \quad (13)$$

$$0 < D_{X,Y} < 1 \quad (14)$$

محدودیت های (۱۱)، (۱۲) و (۱۳) را می توان به یک محدودیت به صورت محدودیت رابطه (۱۵) تقلیل داد:

$$0 < D_{X,Y} < \frac{\min\{UB_Y, UB_X\}}{Pr_X + Pr_Y} \quad (15)$$

لذا مسئله ی بهینه سازی و یافتن مقدار بهینه ی تخفیف به صورت مدل زیر در خواهد آمد:

$$Max \ EAB(D) = ASN [P_1 P_{1 \rightarrow 3} UB_{1 \rightarrow 3} + P_2 P_{2 \rightarrow 3} UB_{2 \rightarrow 3} + P_3 UB_{3 \rightarrow 3} + P_4 P_{4 \rightarrow 3} UB_{4 \rightarrow 3}]$$

$$s.t. \begin{cases} 0 < D_{X,Y} < \frac{\min\{UB_Y, UB_X\}}{Pr_X + Pr_Y} \\ 0 < D < 1 \end{cases} \quad (16)$$

در مورد مدل فوق باید به این نکته اشاره شود که احتمالات  $P_{1 \rightarrow 3}$ ,  $P_{2 \rightarrow 3}$ ,  $P_{4 \rightarrow 3}$  دارای تابعی به صورت رابطه (۹) می باشند. اما از آنجایی که مشتریان گروه های مبدأ ۱ و ۲ پیش از اعمال تخفیف، هر یک در حال خرید یک نوع کالا هستند، لذا دارای شرایط مشابه می باشند و می توان فرض کرد که برای این دو گروه، تابع احتمال انتقال از گروه مبدأ به گروه

۳ یکسان می‌باشد. اما مشتریان گروه ۴ پیش از اعمال تخفیف هیچ کالایی خریداری نمی‌کنند لذا تابع احتمال انتقال آنها از گروه مبدأ به گروه ۳ متفاوت از دو گروه ۱ و ۲ می‌باشد. می‌توان انتظار داشت که به ازای مقادیر مشخص از تخفیف، احتمال ترغیب شدن مشتریان گروه ۴ برای انتقال به گروه ۳، کمتر از مقدار مشابه برای مشتریان گروه ۱ و ۲ است. به عبارت دیگر:

$$P_{4 \rightarrow 3} < P_{1 \rightarrow 3} = P_{2 \rightarrow 3} \quad (17)$$

### تعیین مقادیر $\beta_0, \beta_1$

جهت تعیین مقادیر  $\beta_0, \beta_1$  برای توابع احتمال انتقال از دو روش می‌توان استفاده نمود:  
الف) نخستین روش، نمونه‌گیری و محاسبه‌ی مقادیر پارامترها بر اساس روش حداکثر درستنمایی است. در این روش برای هر یک از گروه‌های ۱ و ۲ و ۴ نمونه‌هایی از مشتریان انتخاب شده و نحوه‌ی پاسخ دهی آنها به مقادیر مختلف تخفیف (انتقال یا عدم انتقال به گروه ۳) ثبت می‌گردد. سپس بر اساس روش حداکثر درستنمایی پارامترها برآورد می‌شوند.

ب) در روش دوم به ازای دو مقدار مشخص برای تخفیف، احتمال خرید توأم (احتمال انتقال به گروه ۳) را مشخص می‌کنیم. دو مقدار مشخص برای تخفیف به همراه احتمالات انتقال متناظر، دو نقطه بر روی نمودار احتمال انتقال رابطه (۹) را مشخص کرده و بر اساس این دو نقطه مقادیر پارامترهای  $\beta_0, \beta_1$  تعیین می‌شوند. منطقی است که در نظر بگیریم که برای گروه‌های ۱ و ۲ در صورتیکه میزان تخفیف ۰٪ باشد، احتمال انتقال از گروه مبدأ به گروه ۳ برابر با ۰ بوده و در صورتیکه مقدار تخفیف ۱۰۰٪ باشد، احتمال انتقال برابر با ۱ خواهد بود؛ اما برای رسیدن به پارامترهای متناهی، مقادیر احتمالی مذکور را به جای ۰ و ۱ به ترتیب ۰,۰۰۱ و ۰,۹۹۹ در نظر می‌گیریم:

$$P_{1 \rightarrow 3}(0) = P_{2 \rightarrow 3}(0) = 0.001$$

$$0.001 = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)} \quad (18)$$

$$0.001001 = \exp(\beta_0)$$

$$-6.9067 = \beta_0$$

همچنین



$$P_{1 \rightarrow 3}(1) = P_{2 \rightarrow 3}(1) = 0.999$$

$$0.999 = \frac{\exp(-6.9067 + 1\beta_1)}{1 + \exp(-6.9067 + 1\beta_1)} \quad (19)$$

$$13.8135 = \beta_1$$

لذا تابع احتمال انتقال برای گروه های ۱ و ۲ برابر خواهد بود با:

$$P_{1 \rightarrow 3}(D) = P_{2 \rightarrow 3}(D) = \frac{\exp(-6.9067 + 13.8135D)}{1 + \exp(-6.9067 + 13.8135D)} \quad (20)$$

برای گروه ۴ در صورتیکه میزان تخفیف ۰٪ باشد، احتمال انتقال از گروه مبدأ به گروه ۳ برابر با ۰ بوده، ولی در صورتیکه مقدار تخفیف ۱۰٪ باشد، احتمال انتقال را برابر با ۰٫۸۵ در نظر می گیریم، چرا که این تابع به علت اینکه سابقاً افراد این گروه هیچ یک از کالا های X و Y را خریداری نمی کرده اند، از تحریک پذیری کمتری برخوردار است. برای رسیدن به پارامترهای متنهای، مقادیر احتمالی مذکور را به جای ۰، ۰٫۰۱، ۰ در نظر می گیریم:

$$P_{4 \rightarrow 3}(0) = 0.001$$

$$0.001 = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)} \quad (21)$$

$$0.001001 = \exp(\beta_0)$$

$$-6.9067 = \beta_0$$

همچنین

$$P_{4 \rightarrow 3}(1) = 0.85$$

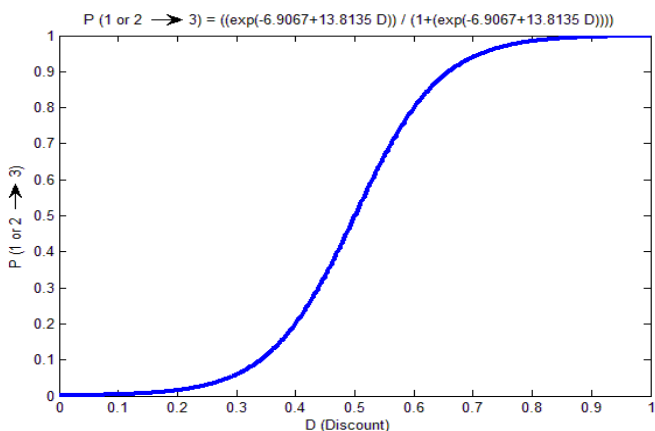
$$0.85 = \frac{\exp(-6.9067 + 1\beta_1)}{1 + \exp(-6.9067 + 1\beta_1)} \quad (22)$$

$$8.6413 = \beta_1$$

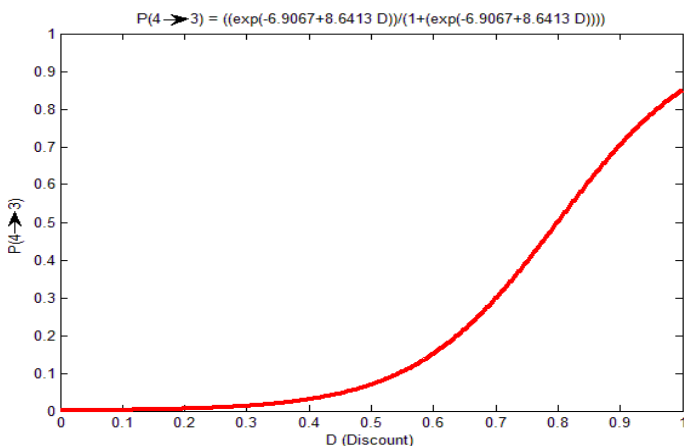
لذا تابع احتمال انتقال برای گروه ۴ برابر خواهد بود با:

$$P_{4 \rightarrow 3}(D) = \frac{\exp(-6.9067 + 8.6413D)}{1 + \exp(-6.9067 + 8.6413D)} \quad (23)$$

نمودار توابع روابط (۲۰) و (۲۳) به ترتیب در شکل‌های (۲) و (۳) نشان داده شده‌اند.



شکل (۲): نمودار احتمال انتقال به گروه ۳ برای مشتریان گروه ۱ و ۲



شکل (۳): نمودار احتمال انتقال به گروه ۳ برای مشتریان گروه ۴

با توجه به شکل‌های (۲) و (۳) مشاهده می‌شود که تمایل خرید هر دو کالا توسط مشتریان گروه‌های ۱ و ۲ به ازای مقدار یکسان تخفیف، بیشتر از مشتریان گروه ۴ بوده و یا به عبارتی دارای تحریک پذیری بیشتری نسبت به تخفیف می باشد.

#### ۴- متدلوژی انجام فرآیند

در این بخش چگونگی انجام فرآیند داده کاوی با رویکرد تعیین سبد خرید، با استفاده از الگوریتم های قواعد تلازمی در فروشگاه های زنجیره ای توضیح داده می شود. طبق متدلوژی CRISP مراحل انجام فرآیند های داده کاوی به شرح ذیل است (هن و کمبر ۲۰۰۶).

۱) گام شناخت سیستم

۲) گام شناخت داده ها

۳) گام آماده سازی داده ها

۴) گام مدل سازی

۵) گام ارزیابی

۶) گام توسعه و اجرا

در شکل (۴) جزئیات این گام ها با توجه به مورد مطالعاتی مقاله که مدل پیشنهادی استخراج قواعد تلازمی در فروشگاه های زنجیره ای است نشان داده شده است.

#### ۵- نتایج مطالعه موردی

در این بخش بر اساس متدلوژی انجام فرآیند که گام های آن در بخش قبل شرح داده شد، به استخراج قواعد تلازمی و بررسی نتایج آن در فروشگاه زنجیره ای می پردازیم.

##### ۵-۱- شناخت سیستم (درک کسب و کار)

فروشگاه های زنجیره ای با شعب متعدد و جامه ی آماری بزرگ مشتریان که با توجه به فروش و گستردگی آن، بهبود عملکرد در این حوزه اثرات فراوانی در بهره وری سازمان و رضایت و رفاه حال کلیه مشتریان آن خواهد داشت.

##### ۵-۲- شناخت داده ها

در این مرحله بعد از رایزنی هایی که با سازمان صورت گرفت به جمع آوری داده های مورد نیاز و مرتبط با هدف پروژه پرداخته شد. البته دستیابی به داده ها با مشکلات بسیار

زیادی همراه بود که از جمله آن می‌توان به حجم بالای داده‌ها که در حدود ۴ میلیون تراکنش را شامل می‌شد اشاره کرد. هر ردیف از داده‌ها که محتوی یک تراکنش است، شامل هفت ستون به شرح زیر می‌باشد: ستون اول تاریخ فاکتور به تاریخ شمسی، ستون دوم شماره صندوق، ستون سوم شماره فاکتور، ستون چهارم کد کالا، ستون پنجم نام کالا، ستون ششم تعداد کالا خریداری شده و ستون آخر قیمت کالا خریداری شده را نشان می‌دهد. در این مقاله تراکنش‌های مربوط به شعبه‌ای خاص از فروشگاه زنجیره‌ای از تاریخ ۱۳۸۸/۸/۲۰ تا ۱۳۸۹/۸/۲۳ مورد مطالعه قرار گرفته است. یک مسئله مهم در فرآیند داده‌کاوی در تعیین سبد خرید، شناسایی هر فاکتور و یا در واقع خرید‌های هر مشتری خاص است، لذا با تحلیل داده‌ها مشخص گردید که امکان یکسان بودن داده‌های هر یک از ستون‌ها به تنهایی با سایر تراکنش‌ها وجود دارد اما یکسان بودن داده‌های هر یک از ستون‌ها به تنهایی با سایر تراکنش‌ها وجود دارد اما یکسان بودن تمامی سه ستون مذکور دلالت بر خرید یک مشتری واحد است؛ لذا اگر در چند تراکنش، سه ستون اول شامل (تاریخ فاکتور، شماره صندوق و شماره فاکتور) یکسان باشند، نشانگر یک مشتری خاص بوده و لذا اقلام مربوط به آن تراکنش‌های خاص با هم خریداری شده‌اند.

### ۵-۳- آماده‌سازی داده‌ها

در این فرآیند ابتدا داده‌ها به نرم افزار صفحه گسترده (در این مقاله اکسل) وارد و به بخش‌های ۵۰۰ هزار تایی تقسیم شد. سپس با توجه به موارد مورد نیاز در الگوریتم‌های قواعد تلازمی، ستون‌های مورد نیاز برای ادامه‌ی کار انتخاب شدند. بعد از انجام این کار، راه حلی برای ترکیب این سه ستون جهت دستیابی به کد شناسه یکتا مشتریان دنبال شد؛ بدین صورت که فرض شد کد شناسه، یک عدد ۱۰ رقمی است که از ضرب ستون تاریخ خرید (می‌توان تاریخ شمسی را با تبدیل به تاریخ میلادی و اختصاص عدد یک به ۱/۱/۱۹۹۰، به شناسه‌ای یکتا از نوع عددی تبدیل کرد). در ۱۰۰ هزار به اضافه ضرب ستون دوم در ۱۰۰۰ به اضافه ستون سوم بدست می‌آید. در این صورت تمامی کالاهایی که شناسه مذکور در آنها با هم مشترک است، با یکدیگر خریداری شده‌اند. پس از ورود داده‌ها به نرم افزار و کا، عملیات مجتمع‌سازی، الحاق و یکپارچه‌سازی داده‌ها انجام شد. مقادیر کد کالا از مقادیر قابل قبول توسط نرم افزار بزرگتر بودند که نرم افزار توانایی

شناسایی آنها را نداشت، لذا به تصحیح کد کالا ها پرداخته شد. مقادیر تکراری نیز حذف شدند تا فایلی که در یک ستون آن کد کالا و در ستون دیگر نام کالا قرار دارد حاصل شود. نزدیک به ۱۰۰۰۰ نوع قلم کالا در داده ها مشاهده شد و فایل ها بصورت یک فایل یکپارچه، آماده ورود به گام مدلسازی شد.

#### ۵-۴- استخراج قواعد تلازمی (مدل سازی)

در این بخش ابتدا یک مقدار کمینه برای پشتیبان معرفی شد. این کار با کسب دیدی دقیق و تحلیلی درست نسبت به فراوانی داده ها صورت گرفت. بعد از گزارشات آماری و تحلیل هایی که نسبت به فراوانی تکرار هر قلم کالا انجام شد، مقدار پشتیبان  $0.005\%$  استفاده شد که در واقع به دلیل آنکه ۱۰۰۰۰ قلم کالا وجود داشته و متوسط تکرار هر قلم  $0.001$  یا  $0.01\%$  است، با انتخاب این مقدار پشتیبان، کالاهایی که کمتر از نصف مصرف انتظاری را دارند به نوعی در قواعد خروجی فیلتر خواهند شد و  $94.52\%$  درصد از کل تراکنش ها پوشش داده می شوند که در واقع نشان دهنده این مطلب است که  $37.76\%$  درصد از قلم کالاها  $94.52\%$  درصد از تراکنش های ما را به خود اختصاص می دهند. این مطلب در نمودار پارتو انواع کالا ها (شکل ۵) نشان داده شده است. نمودار فراوانی انواع کالاها (شکل ۶)، حاکی از آن است که بیشترین فراوانی حدود  $0.009$  است و اکثر داده ها فراوانی کمتر از  $0.001$  دارند. همچنین مقدار اطمینان  $0.25$  یا  $25\%$  انتخاب شده است. داده های موجود در این فرآیند داده کاوی دارای ویژگی های زیر هستند:

- داده ها فرمت تراکنشی دارند.
- ورودی و خروجی بودن آنها نسبت به الگوریتم های قواعد تلازمی مشخص نیست.

با توجه به مشخصات بالا یکی از الگوریتم هایی که قابلیت استخراج قواعد تلازمی را داراست، الگوریتم کارما می باشد.

#### ۵-۴-۱- الگوریتم کارما

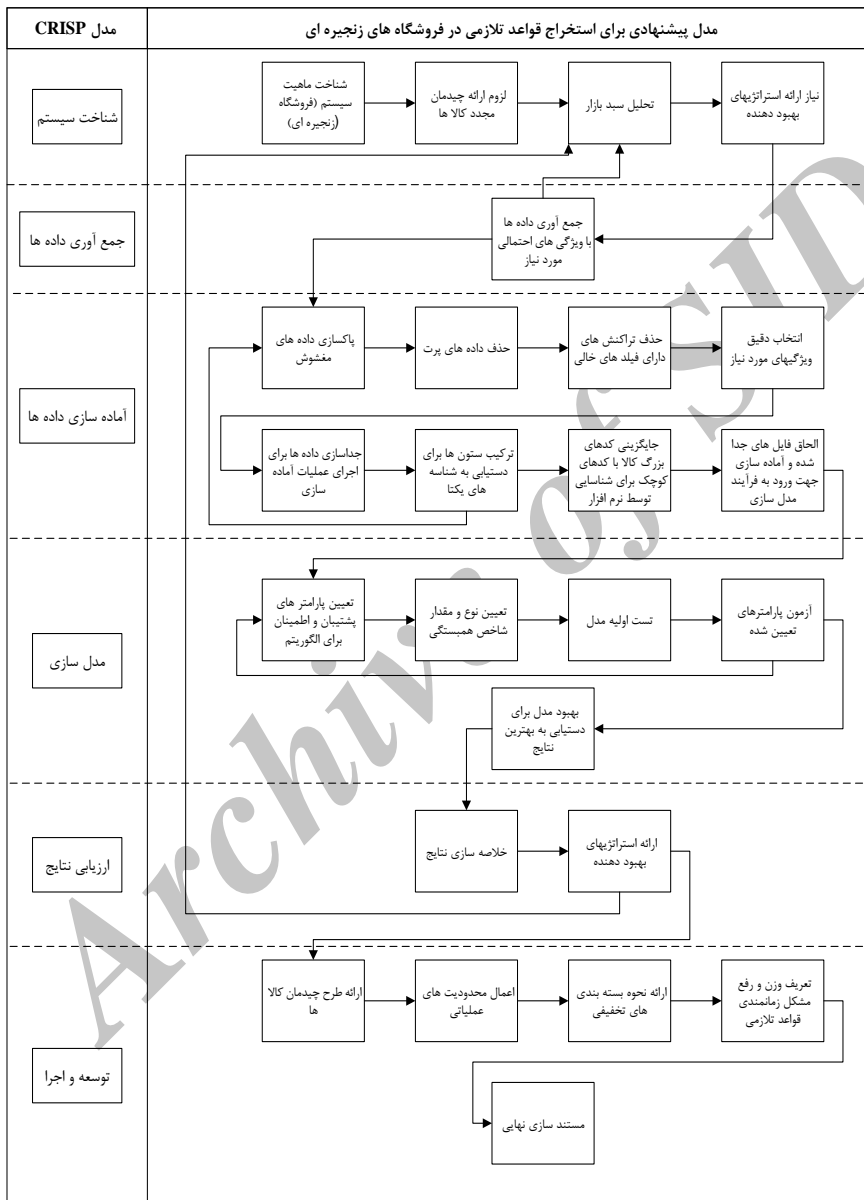
مدل کارما مجموعه ای از قواعد را از داده ها استخراج می کند، بدون اینکه نیاز باشد که ورودی (پیش بینی کننده) و خروجی (هدف) تعیین شود. این امر بدین معنی است که قواعد تولید شده می توانند برای برنامه های کاربردی متنوعی استفاده شوند. الگوریتم کارما

از نظر ساختار به الگوریتم اپریوری شبیه است اما توانایی تولید قوانین دو طرفه را نیز دارا می باشد. قوانین بدست آمده از اجرای الگوریتم کارما بر داده ها بسیار زیاد بودند که در این مقاله از آوردن آنها به علت حجم بالا خودداری شده است. این قوانین در دو دسته جای می گرفتند: قوانین با اطمینان ۱۰۰ درصد و قوانین با اطمینان بیشتر از ۲۵ درصد و کمتر از ۱۰۰ درصد. وجود اطمینان ۱۰۰٪ با توجه به تعبیر آن که لزوم خرید توام دو کالا در تمامی مواقع است، کمی دور از ذهن به نظر می رسد؛ لذا درگام بعدی به ارزیابی نتایج پرداخته می شود.

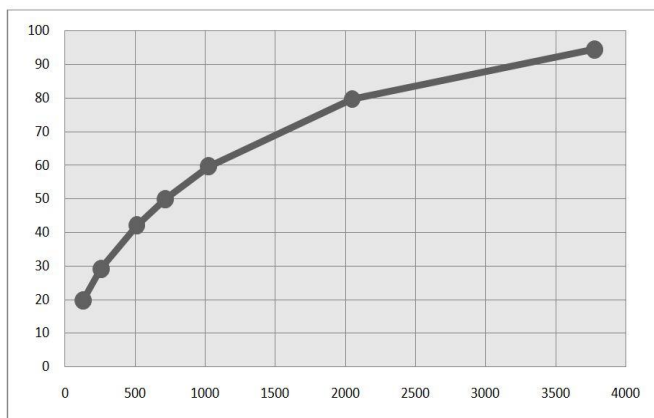
#### ۵-۵- ارزیابی نتایج

۵-۵-۱- شاخص های ارزیابی قواعد تلازمی: یکی از شاخص های متداولی که در ارزیابی قواعد تلازمی استفاده می شود و همچون شاخص لیفت<sup>۸</sup> به ارزیابی میزان وابستگی مقدم و تالی می پردازد، شاخص کساین<sup>۹</sup> (هن و کمبر ۲۰۰۶) است. این شاخص نرمال شده شاخص لیفت است و بصورت رابطه شماره (۲۴) تعریف می شود.

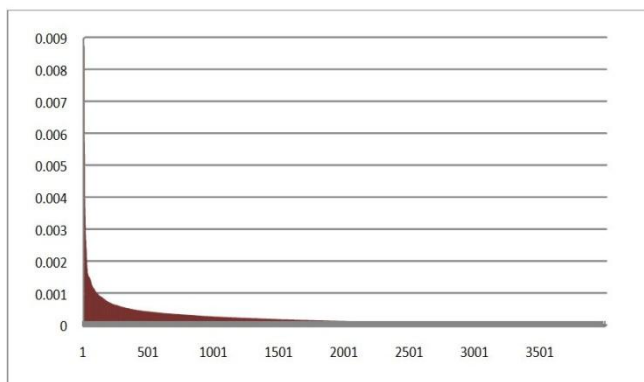
$$COSINE(A, B) = \frac{SUP(A, B)}{\sqrt{SUP(A) \times SUP(B)}} \quad (24)$$



شکل (۴): مدل پیشنهادی برای استخراج قواعد تلازمی در فروشگاه های زنجیره ای



شکل (۵): نمودار پارتو انواع کالاها



شکل (۶): نمودار فراوانی انواع کالاها

دو فرمول Lift و COSINE شبیه هم می باشند با این تفاوت که در COSINE از ریشه احتمالات استفاده می شود؛ لذا این شاخص فقط تحت تاثیر پشتیبان عناصر است و نسبت به تعداد کل تراکنش ها حساس نیست. در این مقاله ارزیابی همبستگی و آزمون غیر تصادفی بودن با محاسبه شاخص COSINE صورت پذیرفته است.

#### ۵-۲-۵- تحلیل و خلاصه سازی قواعد خروجی



در قوانین بدست آمده با اطمینان ۱۰۰ درصد، از نوع کالاها و اسامی آنها مشخص می شود که مشتری میان زوج کالاهای مقدم و تالی، عملاً تمایزی قائل نمی شوند. به عنوان مثال مشتریان بین تن ماهی کلیددار ۱۹۰ گرمی و تن ماهی کلیددار ۱۸۰ گرمی، بین رب انار ۳۶۵ گرمی یک و یک و رب انار ۲۶۵ گرمی یک و یک، بین همبرگر گوشت ۷۰٪/۵۰۰ گرمی و همبرگر گوشت ۳۰٪/۵۰۰ وکیوم گرمی و بین کیک متالایز ۹ طعم پم پم و کیک متالایز وانیل پم پم، انتخابی تصادفی انجام می دهند (جدول ۱).

جدول (۱): نمونه هایی از قواعد با اطمینان ۱۰۰٪

کد کالای مقدم	نام کالای مقدم	کد کالای تالی	نام کالای تالی	درصد اطمینان	شاخص همبستگی	درصد تکرار مقدم
6260007410361	پنیر سفید ۱۰۰ گ پگاه	6260007410367	پنیریوف ۴۰۰ گرم پگاه	100	1.000	1.699
6261111100605	رب انار ۳۶۵ گ او ۱	6261111100905	رب انار ۲۶۵ گ او ۱	100	1.000	0.173
2001415000058	قند کارتنی ۶ پانیز	2001415000059	کله کارتنی ۶ پانیز	100	0.999	0.313
6260173210086	تن ماهی ۱۸۰ گ پونل	6260173210797	تن ماهی ۱۹۰ گ پونل	100	0.999	0.251
6260187300025	کالباس ژامبون ۷۰٪/۴۰۰	6260187300125	کالباس ژامبون ۷۰٪/۵۰۰	100	0.999	0.277
6260022800642	کیک کاکائو ۱۴۰ گ درنا	6260022800532	کیک پذیرای ۱۴۰ گ مینو	100	1.005	0.027
6260731300266	ژامبون مرغ ۹۰٪/ویژه	62607313002649	ژامبون تنوری مرغ ۹۰٪	100	1.000	0.26

قواعد باقیمانده دارای اطمینان بین ۲۵٪ و ۱۰۰٪ خواهند بود که بر اساس شاخص همبستگی به ترتیب می شوند و در اینجا به دلیل حجم بالا از آوردن تمامی آنها صرف نظر شده است. در ادامه حدی برای شاخص همبستگی قرار خواهد گرفت که قواعد با مقادیر شاخص همبستگی کمتر از آن حد را عملاً با هم آبی های تصادفی و شانسی و احتمالاً غیر واقعی پنداشته و از لحاظ کردن آنها صرف نظر خواهیم کرد. در اینجا این حد را شاخص همبستگی ۲۰٪ یا ۰٫۲ لحاظ شده است. با استفاده از نظرات افراد خبره به نظر می رسد که باز هم بعضی از قواعد نیاز به بازنگری در تفسیر دارند؛ به عنوان مثال مشتری احتمالاً تمایزی بین صابون زرد و بنفش عروس قائل نمی شود، لذا این قواعد دارای اطمینان بالایی خواهند بود. در جدول (۲) (جدول قواعد مفید منتخب توسط فرد خبره با شاخص همبستگی بالای ۲۰٪)، قواعد واقعی تر و منطقی تر از دید افراد خبره گرد آوری شده اند.

به عنوان مثال افرادی که حشره کش برقی تارومار را خریداری می کنند، در ۵۵,۲٪ از مواقع قرص حشره کش تارو مار را نیز برای کشتن حشرات با استفاده از دستگاه و قرص خریداری می کنند. افرادی که ادویه فلفل را خریداری می کنند، در ۳۳,۹٪ از مواقع ادویه ماست و خیار را نیز برای تهیه غذا خریداری می کنند. افرادی که کاغذ کلاسور ۱۰۰ برگ رینگ را خریداری می کنند، در ۳۵,۴٪ از مواقع دفتر ۱۰۰ برگ رحلی را برای تکمیل لوازم التحریر خود خریداری می کنند. افرادی که ترد بزرگ مینو را خریداری می کنند، در ۳۲,۵٪ از مواقع قرص بیسگویت لیوان ساقه طلایی را نیز خریداری می کنند. افرادی که بیکنینگ پودر ۳۵ گرمی صیتی را خریداری می کنند، در ۳۲,۷٪ از مواقع وانیل شکری ۱۰ گرمی صیتی را نیز برای پخت کیک خریداری می کنند.

البته قوانین بدست آمده علاوه بر قوانین یگانه، شامل قوانین چندگانه نیز می باشند؛ مثلاً اگر مشتری پسته شوراکبری ۴۵۰گ اتکا ، بادام شور خارجی ۴۵۰گ، پسته خام اکبری ۴۵۰گ اتکا را خریداری نماید، آنگاه در ۶۳٪ مواقع تخمه جابونی شور ۴۰۰گ را نیز خریداری خواهد کرد، که البته در این مقاله از ذکر قواعد چند گانه استخراج شده به دلیل حجم بالای آنها صرف نظر شده است.

جدول (۲): قواعد مفید منتخب توسط فرد خیره با شاخص همبستگی بالای ۲۰٪

درصد تکرار مقدم	شاخص همبستگی	درصد اطمینان	نام کالای تالی	کد کالای تالی	نام کالای مقدم	کد کالای مقدم
0.023	0.329	55.285	قرص حشره کش تارومار	6260137939015	حشره کش برقی تارومار	6260137939114
0.032	0.316	36.416	سوپ قارچ ۱۱۰گ هاتی کا	6261107016668	سوپ ورمیشل ومرغ ۱۵۰گ	6261107016682
0.016	0.317	35.294	نود البت گوشت آماده	6260360050211	نودسبزیجات البت	6260360050013
0.069	0.414	34.595	پودر عصاره گوساله ۴۰گ	6261107016040	پودر عصاره مرغ ۴۰گ هات	6261107016019
0.023	0.274	34.146	دفتر ۱۰۰ برگ رحلی پ.پ	6269431216778	کاغذ کلاسور ۲۶رینگ ۱۰۰ب	6269431215610
0.053	0.292	33.916	ادویه ماست و خیار ۴۰گ	6261107016217	ادویه لیموفلفلی ۴۰گ ه	6261107016248
0.045	0.316	33.61	سوپ جو ۱۵۰گ هاتی کارا	6261107016651	سوپ قارچ ۱۱۰گ هاتی کا	6261107016668
0.032	0.275	33.526	سوپ جو ۱۵۰گ هاتی کارا	6261107016651	سوپ ورمیشل ومرغ ۱۵۰گ	6261107016682
0.576	0.201	32.955	عدس کانادایی ۹۰۰گرم	6260148310025	لوبیاقرمز جگری ۹۰۰گ ا	6261643500171
0.063	0.231	32.743	وانیل شکری ۱۰ صیتی	6260078200212	پکینگ پودر ۳۵ گ صیتی	6260078201950
2.487	0.203	32.562	بیسگویت لیوان ساقه ط	6260111310106	ترد بزرگ مینو	6260100103276
0.019	0.242	32.353	بشقاب گودپلاستیکی پر	6260268180515	بشقاب تخت پلاستیکی پ	6260268180508
0.037	0.301	31.5	کوکوسبزی ۵۰۰گرم ب آ	6260859000857	کوکوسبب زمینی ۵۰۰گ ب	6260859001069
0.069	0.276	31.351	چاشنی خورشت قیمه ۴۰گ	6261107016118	پودر عصاره مرغ ۴۰گ هات	6261107016019

50173471	آدامس سیب اربیت	42113270	آدامس هندوانه اربیت	30.344	0.267	0.152
6261107016309	ادویه کبابی ۴۰گ هاتی	6261107016118	چاشنی خورشت قیقه ۴۰گ	30.303	0.265	0.068
6261107016651	سوپ جو ۱۵۰گ هاتی کارا	6261107016668	سوپ قارچ ۱۱۰گ هاتی کا	30.224	0.316	0.05
6260268112035	بشقاب کودک ماهی ۵۰۲	6260268143046	بشقاب بزرگ ماهی ۵۰۱۴	29.607	0.279	0.062
6261107016217	ادویه ماست و خیار ۴۰گ	6261107016200	چاشنی اسپاگتی ۴۰گ هات	29.243	0.261	0.072
2000322000076	پودردستی مازادسبد ۵۰۰	2000322000083	پودرم آنزیم دارمازاد	29.189	0.212	1.18
6261107016040	پودرعصاره گوساله ۴۰گ	6261107016200	چاشنی اسپاگتی ۴۰گ هات	28.846	0.211	0.049
6260360000483	سوپ قارچ آماده لذیذ	6260360000490	سوپ سبزی آماده لذیذ	28.313	0.231	0.031
6260105031109	پودر آنزیم دار ۵۰۰ و ش	6260105011118	پودر رختشویی ۵۰۰پاک	28.263	0.344	2.892
6261107016040	پودرعصاره گوساله ۴۰گ	6261107016118	چاشنی خورشت قیقه ۴۰گ	28.077	0.208	0.049
6260268152222	قاشق آجیل ۹۳۹ پردیس	6260268152253	کوکوبرگردان ۹۴۲ پردیس	28.07	0.262	0.021
6260859000857	کوکوسبزی ۵۰۰گرم ب آ	6260859001069	کوکوسیب زمینی ۵۰۰گ ب	27.511	0.299	0.043
6260268143046	بشقاب بزرگ ماهی ۵۰۱۴	6260268112035	بشقاب کودک ماهی ۵۰۲۰	27.298	0.280	0.067
6260010521016	مایع لباس سفیدشوی ۹۰	6260010521030	مایع لباس رنگی ۹۰۰گ	26.984	0.234	0.035
6260268143022	بشقاب کوچک ماهی ۵۰۱۲	6260268112035	بشقاب کودک ماهی ۵۰۲۰	26.609	0.230	0.044
6261107016309	ادویه کبابی ۴۰گ هاتی	6261107016200	چاشنی اسپاگتی ۴۰گ هات	26.171	0.230	0.068
6260627400162	ران مرغ بی پوست ۱۸۰۰	6260627400193	سینه مرغ بی پوست ۱۸۰	26.143	0.214	0.748
6261107016668	سوپ قارچ ۱۱۰گ هاتی کا	6261107016682	سوپ ورمیشل و مرغ ۱۵۰	26.141	0.315	0.045
6261107016309	ادویه کبابی ۴۰گ هاتی	6261107016149	چاشنی قورمه سبزی ۴۰گ	25.895	0.226	0.068
8690637551338	اس.پاک کننده آشپزخان	8690637551352	اس.پاک کننده حمام ۵۰۰	25.818	0.257	0.051
6260268112417	بشقاب ملامین کودک ۵۰۲	6260268143015	کاسه کودک ۵۰۱۵	25.628	0.240	0.037
6261107003606	په ۹۰۰ گرم گلستان	6261107003644	عدس ۹۰۰ گرم گلستان	25.499	0.219	0.487
6260226920986	کرم کارامل درازه ۵۰گ	6260226920979	کرم شکلات درازه ۵۰گ	25.405	0.276	0.219
6261107016217	ادویه ماست و خیار ۴۰گ	6261107016248	ادویه لیموفلفلی ۴۰گ ه	25.326	0.290	0.072
6261107016149	چاشنی قورمه سبزی ۴۰گ	6261107016200	چاشنی اسپاگتی ۴۰گ هات	25.151	0.251	0.093

در این مقاله به استخراج اقلام پر مصرف (جدول ۳) و اقلام کم مصرف (جدول ۴) نیز پرداخته شده است. شناخت کالا های پر مصرف به خودی خود می تواند حائز اهمیت باشد. همچنین از کالا های کم مصرف واقعی می توان در استراتژیهای بهبود که در ادامه خواهند آمد بهره جست.

### جدول (۳): نمونه هایی از پر مصرف ترین کالاها

رتبه در مصرف	تعداد افراد مصرف کننده	کد کالا	نام کالا
1	34218	6260111310106	بیسکویت لیوان ساقه طلایی
2	22380	6260627401053	چرخکرده مخلوط اکیلویی مزرع
3	20931	6261101313039	صابون حمام سبز جعبه ۱۳۰گ لنار
4	15469	6260105031109	پودر آنزیم دار ۵۰۰ و ش



5	15290	6260111311028	بیسکویت کرم مانژ ساق
6	14361	6261125414347	شامپوسوپرزرد ۲۲۰ داروگ
7	13315	6260100103276	ترد بزرگ مینو
8	12318	6260002437073	نمک تصفیه ید ۷۰۰ گله‌ها
9	12092	2000322000083	پودرم آنزیم دارمازاد
10	11041	6261125415191	صابون ۵۸ نخل ۱۴۰ داروگر

#### جدول (۴): نمونه‌هایی از کم مصرف‌ترین کالاها

رتبه در مصرف	تعداد افراد مصرف کننده	کد کالا	نام کالا
6619	1	6260262100038	کیک کاکائو کلاسیک جد
6620	1	6260262112215	کیک مغزدار کیوی آشنا
6621	1	6260268157227	زیرلیوان آرکومین ۷۷۶۴
6622	1	6260268922443	مایع دست سیب ۴ لیتر
6623	1	6260276630217	خورش کشک بادمجان مای
6624	1	6260280114116	شیشه شوی تفنگی آبی ۵
6625	1	6260282014247	فیله سالمون ۷۵۰ گرمی
6626	1	6260292419346	چای تی بگ ۲۰۰ طعم لیمو
6627	1	6260418400975	نکتار پرتقال ۲۱ ع ۲۵۰
6628	1	6260428900717	مغزبان مرغ ۱۸۰۰ گروش

ارزیابی و تحلیل قوانین بدست آمده منجر به استخراج استراتژی‌های بهبودی شده است که در ادامه به آنها اشاره خواهد شد.

#### ۵-۶- توسعه و اجرا

اجرا در واقع محصول و ثمره انجام فرآیند داده‌کاوی است و در این مقاله این امر در قالب موارد زیر بیان می‌شود.

#### ۵-۶-۱- بدست آوردن مقدار بهینه تخفیف بسته‌ی تخفیفی (مثال عددی)

در این مثال عددی، هدف تعیین مقدار بهینه ی تخفیف بر روی بسته ی تخفیفی شامل حشره کش برقی تارومار با کد کالای ۶۲۶۰۱۳۷۹۳۹۱۱۴ و قرص حشره کش تارومار با کد کالای ۶۲۶۰۱۳۷۹۳۹۰۱۵ می باشد. پارامترهای مدل به شرح زیر است:

X : حشره کش برقی تارومار

Y : قرص حشره کش برقی تارومار

$$Sup(X) = 0.00023, Sup(Y) = 0.00064945, Conf(X \Rightarrow Y) = 0.55285, Conf(Y \Rightarrow X) = 0.1958$$

$$P_4 = 0.99927, Pr_x = 31553, Pr_y = 13400, UB_x = 4733, UB_y = 2680, ASN = 4000000$$

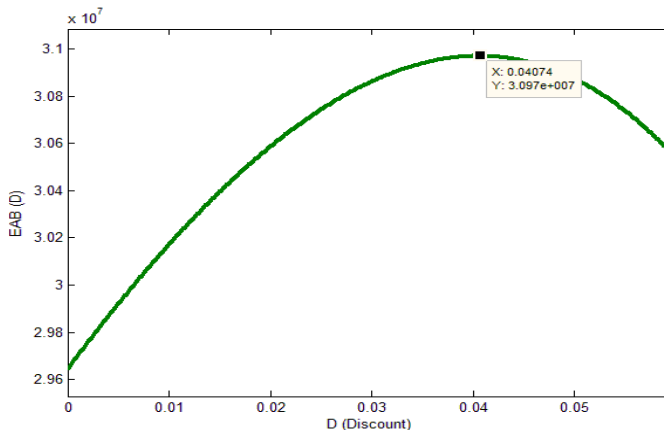
با در نظر گرفتن پارامترهای فوق، مدل بهینه سازی رابطه (۱۷) به صورت زیر در خواهد

آمد:

$$\begin{aligned} Max EAB(D) &= (4000000) \\ &\times \frac{\exp(-6.9067 + 13.8135D)}{1 + \exp(-6.9067 + 13.8135D)} [2.7476 - 28.1015D] \\ &+ \frac{\exp(-6.9067 + 8.6413D)}{1 + \exp(-6.9067 + 8.6413D)} [7407.6 - 44920.27D] \\ &- 4.6231D \end{aligned} \quad (25)$$

$$s.t. \begin{cases} 0 < D_{X,Y} < 0.0596 \end{cases}$$

بدین منظور از نرم افزار MATLAB برای بهینه سازی مدل فوق استفاده شده است. نمودار تابع هدف (EAB(D)) در شکل (۷) نشان داده شده است.



### شکل (۷): نمودار متوسط سود سالانه

نتیجه‌ی بهینه سازی نشان داد که به ازای  $4,074\%$  تخفیف بر روی بسته تخفیفی شامل زوج کالای X و Y، مقدار متوسط سود سالانه حاصل از این کار بیشینه و برابر  $30970000$  واحد پولی خواهد بود. ضمناً اگر فرایند فوق برای کالاهای پر مصرف تر اجرا شود، سودی به مراتب بالاتر حاصل خواهد شد. لازم به ذکر است که کالاهای مذکور به شکل تکی و مجزا نیز بایستی در قفسه ها وجود داشته باشند.

### ۵-۶-۲- سایر استراتژی های بهبود دهنده

علاوه بر ارائه بسته های تخفیفی بر پایه قواعد تلازمی استخراج شده و محاسبه ی مقادیر بهینه ی تخفیف آنها، می توان استراتژی های بهبود دهنده ای به شرح زیر را نیز پیشنهاد کرد:

الف) در مورد کالاهای کم مصرف واقعی که مصرف کم آنها دلایل دیگری از جمله ورود جدید به اقلام فروش و... دارد (همانند مسواک ویددافی، چای کنیا، زیر لیوان آرکومین و...)، استراتژی قرار دادن کالاهای کم مصرف در بسته های تخفیفی کالاهای با هم و یا قرار دادن آنها در بسته های تخفیفی همراه با کالاهای پر مصرف پیشنهاد می شود. ب) کالاهای مذکور در جدول (۲) از لحاظ چیدمان باید به گونه ای در قفسه ها چیده شوند که تالی در کنار مقدمش باشد، چرا که این امر منجر به افزایش اطمینان برای خرید تالی به دلیل در دید و دسترس قرار گرفتن آن و نیز سبب کاهش زمان سفر خرید و کاهش ازدحام و افزایش رضایتمندی مشتری خواهد شد.

ج) از بین زوج کالاهایی که دارای اطمینان ۱۰۰ درصد هستند، کالاهایی با سود بیشتر می تواند انتخاب می شود. مثلاً بین دو نوع کیک همسان درنا و مینو که مشتری عملاً میان آنها تمیزی قائل نمی شود، هر کدام که درصد سود بیشتری را داشته و یا ارتباط با تامین کننده ی آن مناسب تر است انتخاب می شود.

### ۶- نتیجه گیری و بحث

همانطور که بیان شد، در موضوع بهینه سازی در فروشگاه های زنجیره‌ای، انتخاب کالاهایی با سود بیشتر و یا تعامل مناسب‌تر با تامین کنندگان از میان کالاهایی با انتخاب تصادفی توسط مشتری، چیدمان کالای تالی در کنار کالای مقدم در قفسه ها به منظور افزایش اطمینان برای خرید کالای تالی و نیز کاهش زمان سفر خرید به دلیل در دید و دسترس قرار گرفتن آن بر طبق الگوی رفتاری مشتری، ارائه بسته های تخفیفی با مقادیر بهینه‌ی تخفیف بدست آمده از حل مدل پیشنهادی که امید ریاضی سود سالانه را حداکثر می کند و نیز قرار دادن کالاهای کم مصرف در بسته‌های تخفیفی کالاهای با هم آ و یا قرار دادن آنها در بسته های تخفیفی همراه با کالا های پر مصرف، به عنوان راهکارهای بهبود دهنده در مسئله سبد بازار پیشنهاد می شود. به علاوه مزیت رقابتی ناشی از ارائه این روش در فروشگاه‌های زنجیره‌ای و در سطح وسیع موجب خوش نامی هرچه بیشتر سازمان و تاثیر بر سطح رضایتمندی و فروش کل نیز خواهد شد.

به عنوان تحقیقات آتی که در اصل گام توسعه اجرایی متدولوژی می باشد، باید تمامی کالاهای کم مصرف و با هم آیی را که قرار است با یکدیگر در بسته های تخفیفی قرار داده شوند مشخص کرده و میزان تخفیف هر یک به طور دقیق و با توجه به قیمت آنها، با رعایت محدودیت‌های محاسباتی و عملیاتی و دسترسی به حداقل معینی از سود سالانه، محاسبه شوند. همچنین بسته‌های تخفیفی می‌توانند به طور یکسان در فروشگاه‌های نزدیک از لحاظ الگوی جغرافیایی و رفتاری مشتری اجرا شوند، ولی برای سایر شعب، به ویژه شعب شهرهای دیگر، بهتر است که این فرآیندها بر اساس اطلاعات بومی و منطقه‌ای مجدداً اجرا گردند.

زمینه دیگری که در تحقیقات آتی می توان به آن پرداخت، قابلیت تطبیق نتایج در طول زمان است؛ چرا که یکی از ایراداتی که به قوانین تلازمی گرفته می شود، عدم امکان استفاده مستمر از آنها می باشد و نتایج آنها فقط برای یک بازه مشخص معتبر است؛ لذا با تعریف وزن برای داده ها می توان از الگوریتم وارم که در سال ۲۰۰۸ ارائه شده است استفاده کرده و این ضعف را از قواعد تلازمی رفع کرد.

### فهرست منابع

- 1) Agrawal, R., Srikant, G., (1994). Fast algorithms for mining association rules. Proceedings of the 20<sup>th</sup> VLDB conference, Santiago, Chile, 478-499.

- 2) Clementini, P., Felice, D., Koperski, K., (2000). Mining multiple-level spatial association rules for objects with a broad boundary, *Data and Knowledge Engineering*, 34 (3), 251–270.
- 3) Han, J., Fu, Y., (1999). Mining Multiple-Level Association Rules in Large Databases, *IEEE Trans. Knowledge and Data Eng.*, 798-805.
- 4) Han, J., Kamber, M., (2006). *Data Mining Concepts and Techniques*, Second Edition, Morgan Kaufmann Publisher (imprint of Elsevier), San Francisco, CA 94111.
- 5) Han, J., Pei, J., Yin, Y., (2000). Mining frequent patterns without candidate generation. *Proceedings of the 2000 ACM-SIGMOD Int. Conf. on Management of Data*, Dallas, TX, May.
- 6) Ishibuchi, H., Nakashima, T., Yamamoto, N., (2001). Fuzzy association rules for handling continuous attributes. *Proceedings of the IEEE International Symposium on Industrial Electronics*. 118–121.
- 7) Koperski, Han, J., (1995). Discovery of spatial association rules in geographic information databases. *Proceeding of the 4th International Symposium on Large Spatial Databases (SSD95)*, Maine. 47–66.
- 8) Kuok, A., Fu, W., Wong, M.H., 1998. Mining fuzzy association rules in databases. *SIGMOD Record* 27 (1), 41–46.
- 9) Liu, Pan, Wang, Han, J., (2002). Mining frequent item sets by opportunistic projection. *Proceedings of the 2002 Int. Conf. on Knowledge Discovery in Databases*, Edmonton, Canada, July.
- 10) Ngai, E., Xiu, L, Chau, D., (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems and Applications* 36, 2592-2602.
- 11) Park, M., Chen, S., Yu. P.S., (1997). Using a hash-based method with transaction trimming for mining association rules. *IEEE Transactions on Knowledge and Data Engineering*, 9, 813–825.
- 12) Parvatiyar, A., Sheth, J. N., (2001). Customer relationship management: Emerging practice, process, and discipline. *Journal of Economic & Social Research*, 3, 1–34.
- 13) Savasere, A., Omiecinski, E., Navathe, E., (1995). An efficient algorithm for mining association rules in large database. in: *Proceedings of the 21st VLDB Conference*, 432–444.
- 14) Srikant, R., Agrawal, R., (1996). Mining quantitative association rules in large relational tables. *Proceedings of the ACM-SIGMOD, Conference on Management of Data*, Montreal, Canada, June. 1–12.
- 15) Swift, R. S., (2001). *Accelerating customer relationships: Using CRM and relationship technologies*. Upper saddle river. N.J.: Prentice Hall PTR.
- 16) Wijnsen, R., Meersman, (1998). On the complexity of mining quantitative association rules. *Data Mining and Knowledge Discovery*. 2, 263–281



## یادداشت‌ها

---

- <sup>1</sup> Association rules
- <sup>2</sup> Support
- <sup>3</sup> Confidence
- <sup>4</sup> Logistic Regression
- <sup>5</sup> Expected Value of Annual Benefit
- <sup>6</sup> Average Sales Number
- <sup>7</sup> Weka
- <sup>8</sup> Lift
- <sup>9</sup> Cosine

Archive of SID