مقایسه الگوریتم کامینز و دومرحله‌ای برای انتخاب و خوشه‌بندی تأمین‌کنندگان

محمدعلی افشار کاظمی [۱]
موسی پورمدرس سرچشمه* [۲]

چکیده

امروزه شرکت‌های خودروساز به دنبال تأمین‌کنندگانی هستند که از لحاظ قیمت، کیفیت، تحویل به‌موقع و بسیاری از عوامل دیگر در میان گروه تأمین‌کنندگان در رده بالاتر قرار داشته باشند. تأمین‌کنندگان خودروساز عناصر کلیدی و مهم در کاهش هزینه‌ها و افزایش کیفیت محصولات تولیدی خودروسازان هستند، بر این اساس ارزیابی و انتخاب تأمین‌کنندگان با توجه به معیارهای مطلوب، به یکی از ارکان اساسی تصمیم‌گیری در مدیریت زنجیره تأمین تبدیل شده است. ازاین‌رو در این پژوهش با استفاده از نرم‌افزار داده‌کاوی کلمنتاین به انتخاب متغیرهای مناسب و استفاده از تکنیک‌های داده‌کاوی مبادرت شده است. مقاله حاضر بر اساس هدف آن، پژوهشی کاربردی محسوب می‌شود و از نظر ماهیت و روش از نوع توصیفی مدل‌سازی است. متغیرهای مناسب از میان متغیرهای مندرج در مقالات علمی استخراج‌شده و سپس با استفاده از الگوریتم کامینز وی دومرحله‌ای ارزش‌گذاری شده‌اند. متغیرها با استفاده از نرم‌افزار ابتدا ارزش‌گذاری و سپس درجه اعتماد بودن آن‌ها سنجیده می‌شود و درنهایت با مقایسه این دو الگوریتم باهم خوشه‌بندی تأمین‌کنندگان خودروساز صورت می‌گیرد. نتایج به‌دست‌آمده حاکی از صحت و دقت استفاده از روش‌های داده‌کاوی در خوشه‌بندی تأمین‌کنندگان خودروساز است. متغیرهای به‌دست‌آمده از الگوریتم کامینز در بررسی متغیرها نشان‌دهنده توانایی بیشتر این الگوریتم از بقیه الگوریتم‌های خوشه‌بندی است.

واژگان کلیدی: تأمین‌کنندگان، داده‌کاوی، خوشه‌بندی، کامینز، دومرحله‌ای.

تاریخ دریافت مقاله: ۹۱/۱۲/۲۴، تاریخ پذیرش مقاله: ۹۲/۲/۱۷.

۱. دانشیار و عضو هیأت علمی دانشگاه آزاد تهران مرکز دانشکده مدیریت.
۲. دانشجوی کارشناسی ارشد دانشگاه آزاد تهران مرکز دانشکده مدیریت گرایش تولید.
E-mail: pourmodarres@yahoo.com

۱. مقدمه

دنیای رقابت کنونی ایجاب می‌کند که هر شرکت فعال توانایی واکنش سریع به محرک‌هایی را داشته باشد که حیات شرکت را تهدید می‌کنند. در این میان هیچ شرکت مستقلی نمی‌تواند در بازار رقابتی فعالیت نماید. بنابراین، فشارهای محیطی در سال‌های اخیر باعث شده است تا زنجیره تأمین و مدیریت مناسب آن یک عامل مهم جهت حضور موفق در بازارهای رقابتی مطرح شود. این عامل یک مزیت رقابتی برای شرکت‌ها به شمار می‌رود (Choi & at el, 2007؛ Liu & at el, 2000؛ Pan, 1989).

در همین راستا موضوع ارزیابی و انتخاب تأمین‌کنندگان یک بحث مهم و جدی در این حوزه است. ارزیابی و انتخاب تأمین‌کننده فرآیند ارزیابی، مقایسه و یافتن تأمین‌کننده مناسب است که این تأمین‌کننده قادر به تأمین نیازهای خریدار با بهترین کیفیت مورد انتظار، در مکان مناسب، در حجم مناسب و در زمان مناسب باشد (Mandal & Deshmukh,1994; Mohapatra, 2006).

بدیهی است که شرکت‌های خودروساز دارای تأمین‌کننده‌های فراوانی هستند که هرکدام از این تأمین‌کننده‌ها شاخص‌های آماری زیادی دارند، داده‌های این شاخص‌ها برای بررسی تأمین‌کننده از لحاظ آماری بسیار مشکل و پیچیده است برای استفاده بهینه از این اطلاعات امروزه از روش داده‌کاوی استفاده می‌شود.

در دو دهه قبل توانایی‌های فنی بشر برای تولید و جمع‌آوری داده‌ها به‌سرعت افزایش یافته است. عواملی نظیر استفاده گسترده از بارکد برای تولیدات تجاری، به خدمت گرفتن کامپیوتر در کسب‌وکار، خدمات دولتی و پیشرفت در وسایل جمع‌آوری داده و استخراج الگوهایی که دانش را بازنمایی می‌کنند که این دانش به‌صورت ضمنی در پایگاه داده‌های عظیم، انباره داده و دیگر مخازن بزرگ اطلاعات، ذخیره شده است. داده‌کاوی به‌طور همزمان از چندین رشته علمی بهره می‌برد نظیر تکنولوژی پایگاه داده، هوش مصنوعی، یادگیری ماشین، شبکه‌های عصبی، آمار، شناسایی الگو، سیستم‌های مبتنی بر دانش، حصول دانش، بازیابی اطلاعات، محاسبات سرعت‌بالا و بازنمایی بصری داده. داده‌کاوی در اواخر دهه ۱۹۸۰ پدیدار گشته، در دهه ۱۹۹۰ گام‌های بلندی در این شاخه از علم برداشته شده است و انتظار می‌رود در این قرن به رشد و پیشرفت خود ادامه دهد (David Hand & et al, 2001).

داده‌کاوی یک رشته نسبتاً جدید علمی است که از انجام تحقیقات حداقل در رشته‌های آمار، یادگیری ماشین، علوم کامپیوتر خصوصاً مدیریت پایگاه داده‌ها شکل گرفته است. البته مرزهای این رشته‌ها در داده‌کاوی در حال حاضر مبهم است. داده‌کاوی تعاریف متنوعی دارد. این تعاریف با

پیش‌زمینه‌ها و نقطه‌نظرهای افراد بستگی بالایی دارند. در اینجا بـه ارائـه بعضـی از ایـن تعـاریف پرداخته می‌شود:

۱) اصطلاح داده‌کاوی به فرایند نیم خودکار تجزیه‌وتحلیل پایگاه داده‌های بزرگ بـه‌منظـور یـافتن الگوهای مفید اطلاق می‌شود(Sander, 2003) .

۲) داده‌کاوی عبارت از فرایند استخراج اطلاعات معتبر از پیش ناشناخته، قابل‌فهم و قابل‌اعتمـاد از پایگاه داده‌های بزرگ و استفاده از آن در تصمیم‌گیری در فعالیت‌های تجاری مهـم اسـت (Two, 1999)

۳) داده‌کاوی به معنای جستجو در یک پایگاه داده‌ها بـرای یـافتن الگوهـایی میـان داده‌هـا اسـت (David, 2001).

در سال‌های اخیر به دلیل اهمیت موضوع روابـط در بـازار صـنعتی، شـناخت عـواملی کـه بـه شکل‌گیری و تداوم روابط صنعتی منجر می‌شود بسیار موردتوجه بازاریابـان قـرار گرفتـه اسـت. در زنجیره تأمین، رابطه بین تأمین‌کنندگان و شـرکت بـه‌عنـوان خریـدار صـنعتی از اهمیـت خاصـی برخوردار است. اعتماد متقابل، تعهد، ارتباط و همکاری بین این دو موجب رضایت، وفاداری و سـود بیشتر برای هر دو گروه شده است و درنهایت رضایت مصرف‌کننده نهایی را در پی خواهـد داشـت (عبدالوند و شماعی،۱۳۸۸).

در این میان شرکت‌ها با انبوهی از اطلاعات روبرو هستند که از طرف تأمین‌کنندگان به‌سـوی شرکت سرازیر می‌شود مدیران بایستی بتوانند با استفاده از تکنیـک‌هـای خـاص از ایـن اطلاعـات استفاده و تأمین‌کنندگان را ارزیابی کنند و فعالیت‌های آنان را زیر نظر داشـته باشـند و در صـورت لزوم با درجه‌بندی تأمین‌کنندگان و دسته‌بندی آنان شرکت را همیشه در موقعیت برتر قرار دهند.

یکی از این تکنیک‌ها داده‌کاوی است که شرکت با استفاده از ابزارهای موجود در آن مـی‌توانـد در این دسته‌بندی و رتبه‌بندی از آن استفاده کند.

۲. پیشینه تحقیق

ارزیابی و انتخاب تأمین‌کنندگان فرآیندی است که طـی آن تأمین‌کنندگان، به‌عنوان جزیـی از زنجیره تأمین، مورد تحلیل، ارزیـابی و انتخـاب قـرار مـی‌گیرنـد. در ایـن مقالـه وبـر و همکـاران (Weber, & Ellram, 1992) به استناد بررسی ۷۴ مقاله که در ارتبـاط بـا روش‌هـا و معیارهـای انتخاب تأمین‌کنندگان بین سال‌های ۱۹۶۶–۱۹۹۰ تهیه کرده بودند دریافتند بیش از ۶۳٪ مقـالات (۴۷ مقاله) موضوع انتخاب تأمین‌کننده را در محیط چندمعیـاره مـدنظر قـرارداده‌انـد و بـه بررسـی معیارها و روش‌های تحلیلی در فرایند انتخاب تأمین‌کننـده پرداختـه و معیارهـا را بـه‌صـورت زیـر معیارها را رتبه‌بندی نموده‌اند:

۱-قیمت (۶۱ مقاله/ ۸۲٪) ۲-موعد تحویل (۴۴ مقاله /۵۹٪) ۳-کیفیت (۴۰ مقاله/ ۵۴٪) ۴-ظرفیت و تجهیزات (۲۳ مقاله/ ۳۱٪) ۵-موقعیت جغرافیایی (۱۶ مقاله/ ۲۱٪) ۶- توانایی فنی (۱۵ مقاله/ ۲۰٪) ۷- مدیریت و سازمان (۱۰ مقاله/ ۱۳٪)۸-شهرت و موقعیت در صنعت (۸ مقاله/ ۱۱٪) ۹-وضع مالی (۷ مقاله/ ۹٪)۱۰-تاریخچه عملکرد (۷ مقاله/ ۹٪) ۱۱-سرویس و خدمات (۷ مقاله/ ۹٪) ۱۲-رویکردها (۶ مقاله/ ۸٪) ۱۳- توانای بسته‌بندی (۳ مقاله/ ۴٪) ۱۴- کنترل کیفی (۳ مقاله/ ۴٪) ۱۵- برنامه‌های آموزشی (۶ مقاله/ ۳٪) ۱۶- سابقه کار (۲ مقاله/ ۳٪) ۱۷- سیستم ارتباطات (۲ مقاله/ ۳٪) ۱۸- توافق دوجانبه (۲ مقاله/ ۳٪) ۱۹- تأثیرپذیری (۲ مقاله ۳٪) ۲۰- میل به تجارت و (۱ مقاله ۱٪) ۲۱- تضمین و گارانتی (۱ مقاله ۱٪) می‌توان نتیجه گرفت که معیارهای با رتبه پایین که در دهه ۶۰ میلادی در اواخر قرن بیستم و اوایل قرن بیست یکم از اهمیت بیشتری برخوردار شده‌اند؛ مانند عوامل موقعیت جغرافیایی که در تحقیق دیکسون در رتبه بیستم قرار داشته در تحقیق وبر به رتبه ۵ آمده است.

ونگ و یانگ در سال ۲۰۰۹ از یک مدل ترکیبی استفاده کردند مدل ترکیبی آن‌ها شامل فرایند ترکیبی سلسله مراتبی و برنامه‌ریزی چندهدفه خطی مختلط در شرایط تخفیف کلی بود (Angus & et al, 2009).

جامعه آماری: انتخاب تعداد ۲۷۳ عدد از شرکت‌های تأمین‌کننده کلیه قطعات خودرو طرف قرارداد با ساپکو (ایران‌خودرو) سازه‌گستر (سایپا) است شرکت‌های انتخاب‌شده همگی دارای گریداز دو شرکت مذکور می‌باشند.

فرضیه‌های تحقیق: شاخص‌هایی چون تحویل به‌موقع، قیمت، کیفیت، و... از بهترین شاخص‌ها برای خوشه‌بندی تأمین‌کنندگان می‌توان در نظر گرفت.

- اهمیت هریک از شاخص‌ها در خوشه‌بندی متفاوت است.
- با توجه به شاخص‌ها الگوریتم کامینز در خوشه‌بندی بهتر عمل می‌کند.

روش تجزیه تحلیل داده‌ها: در این تحقیق با استفاده از داده‌کاوی و خوشه‌بندی داده‌های تأمین‌کننده مورد تجزیه‌وتحلیل قرار گرفت.

تعاریف عملیاتی مفاهیم

تأمین‌کنندگان: به شرکت‌ها و مؤسساتی اطلاق می‌شود که یکی از نیازهای شرکت را از چه از جهت کالای واسطه ساخته‌شده به‌منظور ایجاد و تولید محصول نهایی بر اساس شاخصه‌های

ارائه‌شده برای شرکت تهیه و تولید می‌نمایند. مدیریت زنجیره تأمین با به وجود آوردن یک سازمان مجازی متشکل از نهادها و ارگان‌های مستقل به دنبال یک هدف است و اثربخشی و کارایی و مدیریت همه نهادها و عملیات‌ها از قبیل یکپارچه‌سازی خرید، مدیریت تقاضا، طراحی و توسعه محصول و برنامه‌ریزی و کنترل تولید است (Harwick, 1997)

خوشه‌بندی[1]: خوشه‌بندی یکی از شاخه‌های داده‌کاوی و نوعی آموزش بدون نظارت[2] است. خوشه‌بندی، تقسیم داده‌ها به گروه‌های مشابه است؛ داده‌ها بر اساس اصل حداکثر کردن شباهت داخل گروه‌ها و به حداقل رساندن شباهت بین گروه‌ها، خوشه‌بندی می‌شوند .خوشه‌بندی یک روش متداول توصیفی است که در جستجوی تشخیص تعداد محدودی خوشه برای توصیف داده‌ها است (Kaufman & Rousseeuw, 2005). در این مرحله با استفاده دو روش خوشه‌بندی: دومرحله‌ای و کامینز به خوشه‌بندی تأمین‌کنندگان پرداخته شده است:

خوشه کامینز: در روش مبتنی بر مرکز (k-means) تشابه خوشه‌ها بر اساس تشابه میانگین اشیاء درون خوشه‌ها محاسبه می‌رود که به‌نوعی بیانگر مرکز ثقل[3] خوشه است. روش کار به این صورت است: در ابتدا به‌صورت تصادفی kشیء که هر کدام بیانگر میانگین یا مرکز یک خوشه هستند انتخاب می‌شوند، هر یک از اشیاء به خوشه‌ای منسوب می‌شوند که کمترین فاصله را با میانگین آن داشته باشند، سپس میانگین جدید برای هر خوشه محاسبه می‌رود و این فرایند تخصیص اشیا به خوشه‌های مناسب ادامه می‌یابد تا تابع هدف همگرا شود. معمولاً برای تابع هدف از معیار مربع خطا استفاده می‌رود: $E = \sum_{i=1}^{k} \sum_{x \in C_i} |x - m_i|^2$ که در آن x نقطه‌ای در فضا بیانگر هر شیء و mi میانگین خوشه Ci است. این معیار به دنبال خوشه‌هایی است که تا حد ممکن فشرده و از هم مجزا باشند. پیچیدگی زمانی الگوریتم از درجه O(nkt) است (t تعداد کل تکرارها است) و معمولاً الگوریتم در یک بهینه محلی متوقف می‌رود (Guha, Rastogi, Shim. 1998).

انتخاب ویژگی: انتخاب ویژگی‌های مناسب برای رسیدن به بهترین نتیجه در طبقه‌بندی داده‌ها از مباحث چالش‌برانگیز دهه‌های اخیراست. اگرچه از دیدگاه تئوری، یادگیری از تعداد ویژگی‌های بیشتر موجب افزایش دقت پیش‌بینی می‌شود، اما شواهد علمی بیانگر آن است که این امر همواره صادق نیست، زیرا تمام ویژگی‌ها، برای تشخیص برچسب کلاس داده‌ها مهم نیستند و یا برخی از آن‌ها اساساً با برچسب داده‌ها بی‌ارتباط هستند (Das et al,1990; Lindenbaum et al, 2004).

1. Cluster Detection
2. Unsupervised learning
3. Center of Gravity

ازآنجاکه عامل‌های بسیاری، ازجمله کیفیت داده‌ها در موفقیت یک الگوریتم یادگیری مؤثر است؛ اگر داده‌ها حاوی ویژگی‌ها و یا اطلاعات تکراری، نامربوط، نویزی و یا نامطمئن باشند، اخذ دانش از داده‌ها مشکل می‌گردد (Hall et al, 2000). هدف از انتخاب ویژگی این است که یک زیرمجموعه کوچک از ویژگی‌ها انتخاب شوند، توزیع ویژگی‌هایی که انتخاب می‌شوند، باید تا حد امکان به توزیع کلاس اصلی با توجه به تمام مقادیر ویژگی‌های انتخاب‌شده نزدیک باشد (John et al, 1994). علاوه بر آن وجود تعداد زیادی از ویژگی‌ها موجب افزایش نامطلوب زمان اجرای الگوریتم یادگیری نیز می‌شود (Bellman, 1961).

برای اینکه بتوان تقسیم‌بندی درستی از روش‌های مختلف انتخاب ویژگی داشت، به این صورت عمل می‌شود که فرایند انتخاب ویژگی در تمامی روش‌ها به این بخش‌ها تقسیم می‌شود:

تابع تولیدکننده: این تابع زیرمجموعه‌های کاندیدا را برای روش موردنظر پیدا می‌کند.

تابع ارزیابی: زیرمجموعه موردنظر را بر اساس روش داده شده، ارزیابی و یک عدد را به‌عنوان میزان خوبی روش باز می‌کند. روش‌های مختلف سعی در یافتن زیرمجموعه‌ای دارند که این مقدار را بهینه کند.

شرط خاتمه: تصمیم‌گیری در مورد زمان توقف الگوریتم است.

تابع تعیین اعتبار: تصمیم می‌گیرد که آیا زیرمجموعه انتخاب‌شده معتبر است یا خیر؟

در این بخش، با استفاده از الگوریتم Relief به محاسبه انتخاب ویژگی پرداخته‌ایم.

خوشه دومرحله‌ای[1]: این روش در دو مرحله خوشه‌بندی را انجام می‌دهد. در مرحله نخست، رکوردها به تعداد زیادی از خوشه‌ها تقسیم می‌شوند و سپس در مرحله بعد، خوشه‌های منتج شده از گام اول به تعداد بهینه شکل داده می‌رود. خروجی این الگوریتم تعداد بهینه خوشه‌ها را معین می‌کند. این الگوریتم یک تحلیل خوشه‌یی قابل مقیاس‌بندی است که برای میزان داده زیاد ابداع شده است. با این روش می‌توان متغیرهای پیوسته و Categorical یا Attributes را بررسی نمود (جمشید ناظمی و همکاران، بهار ۱۳۹۱)

همان‌گونه که از نام این روش مشخص است این الگوریتم شامل دو مرحله است:

۱- پیش خوشه‌بندی کردن حالت‌ها (یا اطلاعات) به تعدادی زیرخوشه این گام از یک روش خوشه‌بندی متوالی استفاده می‌کند. آن داده‌ها را یک‌به‌یک بررسی می‌کند و تصمیم می‌گیرد که آیا

1. Two -step Cluster

داده تحت بررسی می‌تواند در دسته‌بندی‌های قبلی قرار گیرد یا یک دسته‌بندی جدید را بر اساس معیار فاصله که در زیر توضیح داده شده، آغاز کند.

۲-خوشه‌بندی کردن زیر خوشه‌های به‌دست‌آمده از مرحله اول. تعداد خوشه‌ها به‌صورت خودکار در این مرحله تعیین می‌رود. زمانی یک داده در حال بررسی به خوشه‌یی اختصاص داده می‌رود که به آن نزدیک‌تر است. سپس کل خوشه‌ها مقایسه و هر جفت خوشه با کمترین فاصله مابینشان انتخاب و در یک خوشه ادغام می‌شوند (بشری وفایی و همکاران، ۱۳۹۲).

مدل مفهومی تحقیق:

۳. روش‌شناسی تحقیق

برای اجرای پروژه، شناسایی ابزار و داده‌های در اختیار و بررسی وضعیت انجام‌شده به شرح زیر ارائه می‌شود: داده‌های گرید خودروسازان شامل مشخصات ۲۷۲ شرکت تأمین‌کننده و ۴۸ متغیر تأمین‌کننده از تاریخ ۱۳۹۱/۰۱/۱۵ تا ۱۳۹۲/۱۲/۲۸ است که با متغیرهای از قبیل گرید خودروسازان، سنوات همکاری تأمین‌کننده با خودروساز و... ارائه شده است.

برای انتخاب شاخص‌ها از گره Feature Selection استفاده می‌شود تا اندازه اهمیت شاخص‌ها را مشخص کند و با کمک آن مهم‌ترین شاخص‌ها را برای خوشه‌بندی تأمین‌کنندگان انتخاب شود؛ و داده‌های پرت را هم می‌توان با استفاده از گره Filter پاک‌سازی نمود و در ادامه گره مذکور را به گره موسوم به Missing Value imputation برای برخورد با داده‌های مفقود وصل می‌شود و سپس گره مذکور به گره Auto Cluster برای خوشه‌بندی وصل می‌گردد

درنهایت با کمک گره Means که مقایسه مقدار میانگین را به‌صورت زوجی بین رشته‌ها و یا بین گروه‌ها مطابق با یک رشته را انجام می‌دهد، مقدار اهمیت متغیر، درجه آزادی، مقدار، خطای استاندارد انحراف معیار و میانگین در شکل ذیل نشان داده می‌شود. گره فوق میانگین، مقدار، خطای استاندارد و انحراف معیار را نیز محاسبه می‌کند.

تعیین اهمیت هریک از شاخص‌ها در خوشه‌بندی: گره مذکور شاخص‌ها را به ترتیب اندازه اهمیت مشخص می‌کند که مهم‌ترین آن قیمت تحویلی سازنده به خودروساز با درجه اهمیت ۱ است و پایین‌ترین آن به مقدار بدهی اختصاص دارد که تأمین‌کننده می‌تواند بابت بدهی شرکت به او متحمل شود با مقدار ۰۷٪ قرار دارد؛ که مقدار پایین‌تر از این را گره Means در حد وسط یا مرز Marginal قرار می‌دهد. البته گره Meansبعضی از متغیرهای مهم را در Marginal قرارداده بود که با استفاده از گره Feature selection این متغیرها را به چرخه محاسبه بازگرداندیم.

جدول ۱. تعیین میزان اهمیت متغیر، درجه آزادی، خطای استاندارد و انحراف معیار آن

*Cells contain: Mean, Standard Deviation, Standard Error, Count

Field	1.000*	2.000*	3.000*	F-Test	df	Importan... ۳
قیمت تحویلی سازنده به خودروساز	34758.276 10566.733 1962.193 29	23555.350 10217.744 1006.784 103	23781.199 11866.219 999.316 141	12.822	2, 270	1.000 Important
مقدار هزینه ای که شرکت بابت قطع رابطه با تامین کننده باید بپردازد	90724137.931 12386330.460 2300083.824 29	97932038.835 14232841.156 1402403.510 103	93390070.922 15559917.852 1310381.471 141	4.065	2, 270	0.982 Important
ضعف در دانش فنی و تکنولوژی تامین کننده	3.448 1.404 0.261 29	2.709 1.419 0.140 103	2.908 1.398 0.118 141	3.150	2, 270	0.956 Important
مقدار مشارکت تامین کننده در بهبود کیفیت قطعه	47.793 5.583 1.037 29	49.699 5.942 0.586 103	50.617 6.109 0.514 141	2.839	2, 270	0.940 Important
تعداد عدم تامین سفارشات در هنگام تغییر آنها از سوی خودرو ساز	64.034 18.423 3.421 29	65.932 19.068 1.879 103	60.617 19.406 1.634 141	2.334	2, 270	0.901 Important
تعداد فسخ قرار دادیا تامین کننده به دلیل توسان عرضه	2.483 1.122 0.208 29	2.602 1.183 0.117 103	2.291 1.137 0.096 141	2.203	2, 270	0.888 Important
زیان مالی خودروساز بابت برگشت از خط تولید از بابت محصولات ...	64379310.345 7766290.681 1442163.974 29	61737864.078 7124882.063 702035.490 103	61510638.298 7586281.812 638880.180 141	1.830	2, 270	0.838 Important
مشارکت در کاهش هزینه حمل	57.414 8.974 1.666 29	54.029 8.371 0.825 103	54.745 8.929 0.752 141	1.702	2, 270	0.816 Important
تعدادپرسنل مجرب وآزموده	151.207 70.748 13.138 29	127.932 74.118 7.303 103	127.390 71.239 5.999 141	1.381	2, 270	0.747 Important
مقدرا فروش سالیانه تامین کننده به شرکت	118965517.241 23662029.195 4393928.513 29	121708737.864 19079108.309 1879920.402 103	124397163.121 20673613.076 1741032.297 141	1.089	2, 270	0.662 Important

شاخص‌های قابل انتخاب برای خوشه‌بندی تأمین‌کنندگان:

بعد از عملیات میانگین، انحراف معیار، خطای استاندارد و تعیین مقدار می‌توان از بین ۲۹ شاخص، شاخص‌های زیر را برای خوشه‌بندی تأمین‌کنندگان انتخاب نمود.

۱- قیمت تحویلی سازنده به خودروساز؛

۲- وضعیت مالی (سرمایه‌گذاری عمده)؛

۳- گزارش‌های توقف خط تولید (بابت عدم‌تحویل محصول از جانب تأمین‌کننده)؛

۴- مقدار فروش سالیانه تأمین‌کننده به شرکت؛

۵- مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده؛

۶- تعداد کارکنان مجرب و آزموده؛

۷- مشارکت در کاهش هزینه حمل؛

۸- تضمین کیفیت و گارانتی محصول؛

۹–فاصله کارخانه تولیدکننده تا خودروساز؛
۱۰–سنوات همکاری سازنده با شرکت (به سال)؛
۱۱–مقدار مبلغی که تأمین‌کننده می‌تواند بابت بدهی شرکت به وی متحمل شود؛
۱۲– تعداد سفارش تأمین‌شده در هنگام تغییر سفارش از سوی خودروساز؛
۱۳– تعداد برگشتی‌ها از خدمات پس از فروش شرکت؛
۱۴ -تعداد قراردادهای فسخ‌شده با تأمین‌کننده به دلیل نوسان عرضه؛
۱۵– انعطاف‌پذیری در زمان‌بندی تحویل؛
۱۶- تعداد گواهی‌نامه‌های کیفی تأمین‌کننده؛
۱۷– خودروساز با توجه شرایط بازار؛
۱۸– تعداد اخطارهای کیفی هر قطعه‌ساز.

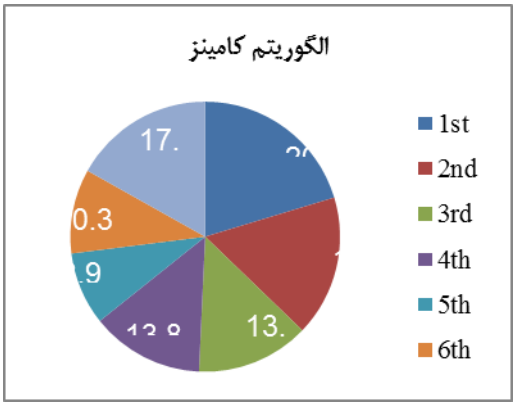
بعد از انتخاب متغیرها اکنون خوشه‌بندی با استفاده از الگوریتم‌های کامینز و دومرحله‌ای انجام می‌شود که در جدول ۲ نشان داده شده است.

جدول ۲: خوشه‌بندی بر اساس الگوریتم کامینز و دومرحله‌ای

مدل	زمان ساخت	ضریب نیمرخ	تعداد خوشه‌ها	کوچک‌ترین خوشه (تعداد)	کوچک‌ترین خوشه (٪)	بزرگ‌ترین خوشه (تعداد)	بزرگ‌ترین خوشه (٪)	کوچک‌ترین/بزرگ‌ترین خوشه	اهمیت
کامینز	۱<	۰/۲۱۶	۷	۲	۲۰/۹	۱۳	۷۹/۱	۱۵/۳	۰/۷۳
دومرحله‌ای	۱<	۰/۱۹۶	۲	۳	۸/۹	۲۰	۲۰/۷	۱۵	۰/۲۶

مشخصات گره k-Means:

خلاصه	مدل
الگوریتم	کامینز
داده‌ها	۲۸
خوشه‌ها	۷



شکل ۱: الگوریتم کامینز

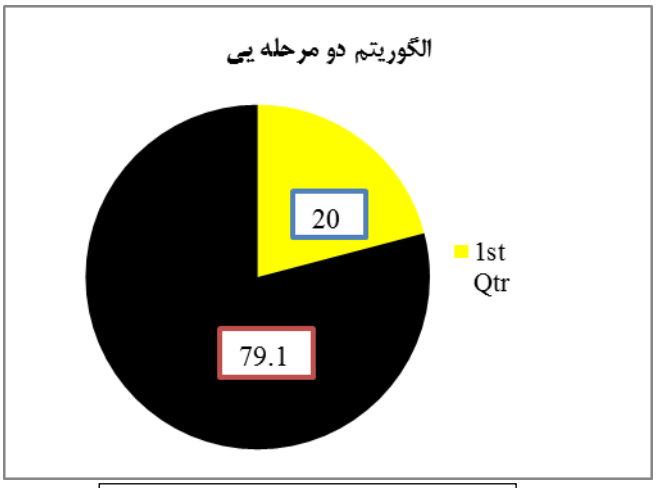
با توجه به جدول ۳ ضریب نیمرخ ۰/۲ به دست می‌آید که در آن بزرگ‌ترین خوشه ۲۰/۰۷٪ و کوچک‌ترین آن ۶/۹٪ قرار دارد.

شکل ۳ بیانگر ضریب اهمیت هریک از متغیرها در خوشه‌بندی است که در آن ضعف تأمین‌کننده در دانش فنی و تکنولوژی با درجه اهمیت ۱ و کمترین آن با مقدار ۱/۴۰ با توان ۴- قرار دارد.

جدول ۳: ضریب اهمیت هر یک از متغیرها در خوشه‌بندی

خوشه	خوشه ۱	خوشه ۲	خوشه ۳	خوشه ٤	خوشه ٥	خوشه ٦	خوشه ۷
اندازه	٪۲۰/۷	٪۱۷/۲	٪۱۳/۸	٪۱۳/۸	٪۶/۹	٪۱۰/۳	٪۱۷/۲
داده‌ها	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده
	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده	مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده
	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی	تعداد گواهی‌نامه‌های کیفی
	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول
	تعداد فسخ قرارداد به دلیل نوسان عرضه	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول	ضمانت و گارانتی محصول
	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده
	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده	رقابتی بودن قیمت سازنده
	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز	تعداد عدم تأمین سفارشات در هنگام تغییر خودروساز

گره دومرحله‌ای: در شکل ۲ ضریب نیمرخ با مقدار میانگین ۰/۱ در دو قسمت که بزرگ‌ترین آن با مقدار ۰/۷۹٪ و کوچک‌ترین آن با مقدار ۲۰/۰۹٪ قرار دارد.



خلاصه مدل	
الگوریتم	دومرحله‌ای
داده‌ها	۲۸
تعداد خوشه‌ها	۲

شکل ۲: الگوریتم دومرحله‌یی

در جدول ۴ (الف، ب و ج) انواع متغیرها برحسب خوشه‌های ۱ و ۲، نشان داده شده است و شامل متغیر مقدار تحویل به‌موقع با درجه اهمیت ۱ در بالاترین و متغیر برگشتی‌ها از خدمات پس از فروش با مقدار ۳/۸۱٪ با توان ۵- است

جدول ٤: انواع متغیرها برحسب خوشه‌های ۱ و ۲

الف

خوشه ۱	خوشه ۲	خوشه
٪۲۰/۹	٪۷۹/۱	اندازه
تعداد تحویل به‌موقع	تعداد تحویل به‌موقع	داده‌ها
تعداد کارکنان مجرب و آزموده	تعداد کارکنان مجرب و آزموده	
تعداد توقف خط تولید بابت عدم‌وجود محصول تأمین‌کننده	تعداد توقف خط تولید بابت محصول تأمین‌کننده	
زیان مالی خودروساز بابت برگشت از تولید محصولات تأمین‌کننده	زیان مالی خودروساز بابت برگشت از تولید محصولات تأمین‌کننده	
ارزش ریالی محصولات تأمین‌کننده	ارزش ریالی محصولات تأمین‌کننده	
سنوات همکاری با شرکت	سنوات همکاری با شرکت	
مقدار مشارکت تأمین‌کننده در بهبود کیفیت قطعه	مقدار مشارکت تأمین‌کننده در بهبود کیفیت قطعه	

ب

تعداد گواهی‌نامه کیفی تأمین‌کننده	تعداد گواهی‌نامه کیفی تتأمین‌کننده
رقابتی بودن قیمت سازنده در بازار	رقابتی بودن قیمت سازنده در بازار
تعداد فسخ قرارداد به دلیل نوسان عرضه	تعداد فسخ قرارداد به دلیل نوسان عرضه
مقدار نرخ بهره‌وری در تأمین‌کننده	مقدار نرخ بهره‌وری در تأمین‌کننده
مشارکت در کاهش هزینه حمل	مشارکت در کاهش هزینه حمل
انعطاف‌پذیری در زمان‌بندی تحویل	انعطاف‌پذیری در زمان‌بندی تحویل
قیمت تحویلی	قیمت تحویلی
تعداد عدم تأمین سفارش در هنگام تغییر از سوی خودروساز	تعداد عدم تأمین سفارش در هنگام تغییر از سوی خودروساز
وضعیت مالی	وضعیت مالی
ضعف در دانش فنی و تکنولوژی تأمین‌کننده	ضعف در دانش فنی و تکنولوژی تأمین‌کننده
مقدار هزینه پرداختی بابت قطع رابطه با تأمین‌کننده	مقدار هزینه پرداختی بابت قطع رابطه با تأمین‌کننده

ج

ضمانت و گارنتی محصول	ضمانت و گارنتی محصول
تعداد اخطارهای کیفی هر قطعه ساز	تعداد اخطارهای کیفی هر قطعه ساز
تعداد محصولات تأمین‌کننده	تعداد محصولات تأمین‌کننده
سطح دانش منابع انسانی	سطح دانش منابع انسانی
ظرفیت ریسک‌پذیری	ظرفیت ریسک‌پذیری
مقدار فروش سالیانه به شرکت	مقدار فروش سالیانه به شرکت
تعداد برگشتی‌ها از خدمات پس از فروش	تعداد برگشتی‌ها از خدمات پس از فروش

٤. نتیجه‌گیری و پیشنهادها

نتیجه‌گیری: مسأله ارزیابی و انتخاب تأمین‌کنندگان به‌طور اساسی یک مسأله تصمیم‌گیری با معیارهای چندگانه است. از ویژگی‌های خاص این مسأله وجود معیارهای کیفی و کمی در ارزیابی است که امر تصمیم‌گیری را با دشواری و پیچیدگی خاصی همراه می‌سازد. داده‌کاوی روش مهمی برای سنجش کارایی مطرح شده است که در این مقاله از این روش برای ارزیابی و انتخاب تأمین‌کنندگان استفاده می‌شود.

در این تحقیق از طریق جمع‌کردن متغیرهای فوق در هریک از الگوریتم‌ها نتایج زیر به دست می‌آید: الگوریتم کامینز برابر با ۲/۹۹ و الگوریتم دومرحله‌ای برابر با ۱/۳۱ می‌شود. با توجه به جواب به‌دست‌آمده الگوریتم کامینز برای خوشه‌بندی تأمین‌کنندگان خودروساز مفیدتر از الگوریتم دومرحله‌یی است از تعداد ۲۷۲ شرکت تعداد ۵۷ شرکت در خوشه ۱، تعداد ۴۷ شرکت در خوشه ۲، تعداد ۳۸ شرکت در خوشه ۳، تعداد ۳۸ شرکت در خوشه ۴، تعداد ۱۸ شرکت در خوشه ۵ و تعداد ۲۸ شرکت در خوشه ۶ و تعداد ۴۷ شرکت در خوشه ۷ قرار دارند.

در این ارزیابی متغیرهای موردنظر عبارتند از: قیمت، کیفیت، تعداد قطعات برگشتی از خط تولید، تعداد قطعات برگشتی از خدمات پس از فروش، تعداد تحویل به‌موقع، تنوع قطعات تولیدی، فاصله کارخانه تأمین‌کننده تا خودروساز، مقدار تمایل خودروساز به ارتباط مجدد با تأمین‌کننده، ضعف در دانش فنی و تکنولوژیک تأمین‌کننده، تعداد اخطارهای کیفی تأمین‌کننده، تضمین کیفیت محصول، ارزش قیمتی محصولات تأمین‌کننده.

پیشنهادها: با توجه نوع فعالیت در این مقاله پیشنهادهای زیر ارائه می‌شود:

۱- استفاده از انواع درخت تصمیم برای طبقه‌بندی صفات تأمین‌کنندگان؛

۲- استفاده از الگوریتم‌های اشاره‌شده برای تأمین‌کنندگان قسمت‌های دیگر صنایع تولیدی؛

۳- استفاده از الگوریتم‌های دیگر برای خوشه‌بندی تأمین‌کنندگان.

منابع

۱.بشری وفایی و همکاران ()۱۳۹۲ «دسته‌بندی توده‌های آب در خلیج چابهار با استفاده از روش خوشه‌بندی»- **اقیانوس‌شناسی**/ سال چهارم/ شماره ۱۳/ بهار.

۲.عبدالوند – محمدعلی، شماعی، آناهیتا (۱۳۸۸) «ارزیابی عوامل مؤثر بر رضایت خریدار در زنجیره تأمین» **مجله مدیریت بازاریابی**، سال ۴، شماره ۷، از صفحه ۶۹ تا ۹۷

۳.جمشید ناظمی و همکاران (۱۳۹۱) «خصوصیات مشتریان بانکداری خرد با استفاده از تکنیک‌های داده‌کاوی». **مجله مدیریت بازاریابی**، شماره ۱۴ –بهار ۱۳۹۱

4. Angus F.M. Huang,John T.H. Wu, Stephen J.H. Yang, Wu-Yuin Hwang(2009); **The success of ePortfolio-based programming learning style diagnosis: Exploring the role of a heuristic fuzzy knowledge fusion**,The journal Expert Systems with Applications Volume 39, Issue 10, August 2012, Pages 8698–8706
5. Bellman, R. Adaptive Control Processes: **A Guided Tour. Princeton University Press,** Princeton, 1961.
6. Choi, J. Bai, S. X., Geunes, J., & Romeijn, H. E (2007). **Manufacturing delivery performance for supply chain management, Mathematical and Computer Modelling**; Vol 45: 11–20.
7. David Hand, Heikki Mannila, Padhraic Smyth. **Principles of Data Mining**. The MIT Press . 2001.
8. .**David J. HAND**, **Data Mining: Statistics and More?**, December 2002.
9. Hall, M. A. Correlation-based **Feature Selection for Discrete and Numeric Class Machine Learning**. In Proceedings of the Seventeenth international Conference on Machine Learning (June 29 - July 02, 2000). P. Langley, Ed. Morgan Kaufmann Publishers, San Francisco, CA, 359-366
10. -Harwick, T. (1997),**optimal decision making for the supply chain APICS, the performance Advantage**, Vol 7, NO1, pp:42-44
11. John, G.H., Kohavi, R., & Pfleger, K. (1994). **Irrelevant features and the subset selection problem. Machine Learning**: Proceedings of the Eleventh International Conference, SanFrancisco, CA: Morgan Kaufman n.
12. J. Sander, ***"Principles of Knowledge Discovery in Data: Clustering I"***, Department of Computing Science University of Alberta, Tutorial Slides, 2003
13. Kaufman, L., Rousseeuw, P.2005. **Finding groups in data: and Introduction to cluster analysis.** 1st ed, John Wiley & Sons.
14. Lindenbaum, M., Markovitch, S., and Rusakov, D. **Selective Sampling for Nearest Neighbor Classifiers**. Machine Learning.number 54, vol. 2 (Feb. 2004), 125-152.
15. Liu, J., Ding, F., Lall, V (2000). **Using data envelopment analysis to compare suppliers for supplier selection and performance improvement**", The Journal of Supply Chain Management; 5(3): 143-150.
16. Mandal, A., Deshmukh, S. G (1994). **Vendor selection using interpretive structural modelling** (ISM), International Journal of Operations and Production Management; 14(6): 52-59.
17. Pan, A.C. (1989). **Allocation of order quantity among suppliers**, Journal Purchasing Materials Management; 39(6): 36-39.

18. Sarkar A., Mohapatra P. K. J (2006). **Evaluation of supplier capability and performance: A method for Supply base reduction**, Journal of Purchasing & Supply Management; 12: 148-163.
19. S.Guha, R. Rastogi, K.Shim, *CURE:* **An Efficient Clustering Algorithm for Large Databases.** Proceedings of the ACM SIGMOD Conference, pp. 73-74 1998.
20. «.**Two Crows Corporation,Introduction to Data Mining and Knowledge Discovery**»,1999
21. Weber, C.A., Ellram, L.M., 1992.« **Supplier selection using multi-objective programming: a decision support system**». International Journal of Physical Distribution & Logistics Management 23(2), 3}14