خوشه‌بندی شرکت‌های دانش‌بنیان کشاورزی با الگوریتم‌های فراابتکاری

الهام مهرپرورحسینی، حامد رفیعی[1]
hamedrafiee@ut.ac.ir

چکیده

خوشه‌بندی یکی از عملیات مهم در داده‌کاوی بشمار می‌رود که نتایج آن برای محققان و سیاست‌گذاران در زمینه‌های مختلف به منظور تحلیل و برنامه‌ریزی مفید است. از آنجا که در سال‌های اخیر، اقتصاد دانش‌بنیان با حمایت از شرکت‌های دانش‌بنیان در ایران در حال توسعه است، تحلیل ویژگی‌های این شرکت‌ها و خوشه‌بندی آنها به منظور برنامه‌ریزی اثربخش، چشم‌انداز مناسبی برای سیاست‌گذاران فراهم می‌آورد. از این رو، در این مقاله شرکت‌های دانش‌بنیان کشاورزی ایران براساس حوزه فناوری، تعداد محصولات و ارزش فروش محصول با استفاده از روش‌های خوشه‌بندی K میانگین ساده و K میانگین در ترکیب با الگوریتم‌های سلسله‌مراتبی و کلونی مورچه‌ها خوشه‌بندی شدند. براساس نتایج، خوشه‌بندی K میانگین در ترکیب با الگوریتم کلونی مورچه‌ها در مقایسه با دو روش دیگر توزیع متوازن‌تری از شرکت‌ها بین شش خوشه داشته است و میانگین ضریب پهنای نیمرخ ۰/۷ اعتبار این خوشه‌بندی را تأیید می‌کند. پیشنهاد می‌شود، به منظور اثربخشی و پرهیز از هدررفت بودجه حمایت از شرکت‌های دانش‌بنیان، برای شرکت‌هایی که به دلیل ویژگی‌های مشابه در یک خوشه قرار گرفته‌اند، برنامه‌های حمایتی مشابه طرح‌ریزی شود. برای شرکت‌های دانش‌بنیان با میانگین فروش اندک که در یک خوشه قرار گرفته‌اند برنامه‌های حمایتی ورود به بازار و برای شرکت‌های با تنوع محصول و درآمد متوسط بالا برنامه‌های توسعه بازار صادراتی توصیه می‌شود.

طبقه بندی JEL: C81, L2, M13, O31, Q16

کلیدواژگان: اقتصاد دانش‌بنیان، داده‌کاوی، خوشه‌بندی K میانگین، الگوریتم سلسله مراتبی، الگوریتم کلونی مورچه‌ها

۱ـ دانشجوی دکتری و استادیار گروه اقتصاد کشاورزی دانشگاه تهران

مقدمه

در دهۀ ۱۹۷۰ میلادی با پیشـرفت سـریع فناوری‌های برتر در جهان، به ویژه در زمینه‌های ارتباطات و رایانه، الگوی رشـد اقتصـادی جهان به طور اسـاسـی تغییر کرد و به دنبال آن، از دهۀ ۱۹۹۰ میلادی، دانش به عنوان مهم‌ترین سـرمایه، جایگزین سـرمایه‌های پولی و فیزیکی شـد(چن و همکاران ، ۲۰۰۴). دانش از منابع اصـلی ایجاد اشـتغال و ثروت و عامل دسـتیابی به توان رقابتی در بازارهای بین‌المللی در عصـر جهانی شـدن اسـت که فرایندهای توسـعۀ اقتصادی را عمق و شتاب می‌بخشد. تأثیر چشمگیر دانش بر عملکرد کلان اقتصادی کشورهای پیشرفته در دهه‌های گذشته، موجب پیدایش اصطلاح اقتصاد دانش‌بنیان شده است(سلمانی، ۱۳۹۴). علی‌رغم تعاریف متعدد در مجامع اقتصادی یک تعریف عمده از اقتصاد دانش بنیان وجود دارد. براساس تعریف بانک جهانی (۲۰۰۷)، اقتصاد دانش‌بنیان اقتصادی است که در آن دانش به عنوان موتور اصلی رشد اقتصادی عمل می‌کند و تولید و توزیع، مبتنی بر آن شکل گرفته و سرمایه‌گذاری در صنایع با محوریت دانش مورد توجه خاصی قرار گرفته است. سرمایه‌گذاری در دانش نیز عبارت است از مجموع هزینه‌های انجام شده در واحدهای تحقیق و توسعه و هر هزینه‌ای که بابت فعالیت‌های ارتقای دانش انجام گیرد.

در این راسـتا، کشـورهای در حال توسـعه برای دسـتیابی به اهدافی مانند رشـد اقتصـادی و بهبود مزیت رقابتی در بازارهای جهانی، اقتصـاد دانش‌بنیان را در دسـتور کار قرار داده‌اند، اما آنها در تبدیل دانش به ثروت با تنگناهای مختلفی روبه‌رو هستند. یکی از مشکلاتی که در زمینه حمایت از شرکت‌های دانش‌بنیان در ایران وجود دارد، استفاده از ابزارهای حمایتی مشـابه بدون در نظر گرفتن تفاوت ویژگی‌های شـرکت‌ها اسـت که موجب کاهش انگیزه و کارایی شرکت‌ها و اتلاف منابع می‌شود. این در حالی است که شرکت‌ها در حوزه‌های مختلف، تنوع محصولات، سطح فروش و سـایر ویژگی‌ها برای توسـعه بازار و بهبود بهره‌وری نیازمند سـطوح و انواع متفاوتی از حمایت‌ها هسـتند. از این رو دسـته‌بندی شـرکت‌های دانش‌بنیان با رویکردی همه‌جانبه نگر، با در نظر گرفتن ویژگی‌های مختلف آنها می‌تواند چشم‌انداز مناسبی را برای سیاست‌گذاران به منظور طرح‌ریزی ابزارهای حمایتی متناسب هر گروه از شرکت‌ها فراهم آورد و شـرکت‌ها با بهره‌مندی از حمایت‌های ویژۀ خود از کارایی و قدرت رقابت‌پذیری بیشـتری برخوردار شـوند. این موضوع از آنجا که سازوکار حمایت از کسب و کارهای دانش‌بنیان در کشور نوپاست از اهمیت بیشتری برخوردار است. زیرا شکل‌دهی و اصلاح سازوکارها در گام‌های نخست با سهولت بیشتری همراه است.

براساس آیین‌نامۀ حمایت از شرکت‌های دانش‌بنیان مصوب ۱۳۹۱ هیئت وزیران، شرکت‌ها در صورت دارا بودن شرایط ذکرشـده در قانون، مجوز دانش‌بنیانی را از معاونت علمی و فناوری ریاسـت جمهوری دریافت می‌کنند و همۀ مراحل اجرایی این قانون توسط این معاونت انجام می شود. در این راستا، معاونت علمی ریاست جمهوری از سال ۱۳۹۱ به

طور ادواری اقدام به انتشار فهرست محصولات دانش‌بنیان مورد پذیرش در ۱۴ حوزه می‌کند که در سال ۱۳۹۵ تعداد کل شرکت‌های دانش‌بنیان ثبت شده ۲۷۹۷ بوده است. از این میان ۱۰۲ شرکت در حوزه زیست‌فناوری به تولید ۲۷۲ محصول در زمینه کشاورزی، دامی و گیاهی پرداخته‌اند(معاونت علمی و فناوری ریاست جمهوری، ۱۳۹۶). با توجه به تأکید سند چشم‌انداز ۱۴۰۴ کشور و بند اول سیاست‌های اقتصاد مقاومتی بر تقویت اقتصاد دانش‌بنیان و همچنین اهمیت نفوذ دانش در بخش کشاورزی به منظور تحول آن از یک بخش سنتی به مدرن، بررسی و برنامه‌ریزی برای شرکت‌های دانش‌بنیان فعال در زمینه کشاورزی، دامی و گیاهی از اهمیت بیشتری برخوردار است. از این رو، در این مقاله به بررسی و خوشه‌بندی شرکت‌های دانش‌بنیان کشاورزی با استفاده از روش Kمیانگین ساده و الگوریتم‌های فراابتکاری ترکیبی با K میانگین و مقایسه این روش‌ها به منظور دستیابی به مناسب‌ترین خوشه‌بندی پرداخته می‌شود. در ادامه به برخی مطالعات مربوط به خوشه‌بندی در علوم اقتصادی اشاره خواهد شد.

امروزه خوشه‌بندی از روش های رایج در مطالعات علوم مختلف از جمله اقتصاد به شمار می‌رود. لینارس و همکاران(۲۰۱۸) با استفاده از نسبت‌های مالی استاندارد شده، شرکت‌های عضو بورس اوراق بهادار را با روش‌های مختلف محاسبه فاصله خوشه‌بندی کردند. به طوری که شرکت‌های با ساختار مالی مشابه در یک گروه قرار گرفتند. رحمان و همکاران(۲۰۱۷) مشتریان یک شرکت بیمه‌ای در بنگلادش را برای سال‌های ۲۰۱۴-۲۰۱۱ با روش‌های مختلف خوشه‌بندی کردند. علی حیدری و خادم زارع (۱۳۹۴) به خوشه‌بندی مشتریان اعتباری بانک‌ها براساس نسبت‌های مالی با استفاده از روش تحلیل پوششی داده‌ها پرداختند. زارع احمدآبادی و همکاران(۱۳۹۳) بازار هدف یک کارخانه کاشی را به طور استانی با دو روش kمیانگین و الگوریتم کلونی مورچه‌ها خوشه‌بندی کردند که براساس شاخص اعتبار خوشه‌بندی، روش الگوریتم کلونی مورچه‌ها نتایج بهتری داشت. در مطالعهٔ شیرکوند و همکاران(۱۳۹۳) برای خوشه‌بندی مشتریان صنعت بیمه، نخست ۳۰ شاخص به عنوان ارزش‌های مورد انتظار مشتریان این صنعت انتخاب و در غالب پرسشنامه با سوالات دارای طیف لیکرت نظرات مشتریان جمع‌آوری شد. سپس با استفاده از روش کی‌میانگین مشتریان صنعت به چهار خوشهٔ قیمت‌گرا، خدمات‌گرا، سهولت‌گرا و رابطه‌گرا تقسیم شدند. در نهایت نتایج آزمون کای دو پیرسون نشان داد، خوشه‌های مشتریان از نظر متغیرهای سن، تأهل، سطح تحصیلات و درآمد تفاوت معناداری با یکدیگر ندارند. نیل‌ساز و همکارانش(۱۳۸۶)، اپتیمهین(۲۰۱۱)، برنث (۲۰۰۵)، الفنسی و سارگنت(۲۰۰۰) پژوهش‌هایی در زمینه خوشه‌بندی داده‌های اقتصادی به انجام رسانده‌اند. در بخش بعدی روش پژوهش با توجه به مطالعات انجام شده و سپس نتایج و پیشنهادها بیان خواهد شد.

مواد و روش ها

داده‌کاوی به فرایند اســـتخراج الگوهای پنهان و یا ویژگی‌های جالب و مفید از مجموعه داده‌ها گفته می‌شـــود که با اســـتفاده از آن می‌توان به تصـــمیم‌گیری و پیش‌بینی رفتار آینده پرداخت. خوشـــه‌بندی در داده‌کاوی یکی از عملیات مهم در نتیجه‌گیری داده‌کاوی بر روی داده‌ها به حســـاب می‌آید. خوشـــه‌بندی افرازبندی یک گروه متنوع به تعدادی زیرگروه مشابه یا گروه‌بندی مجموعه‌ای از اشیاء به کلاسی از اشیاء مشابه می‌باشد، در هر خوشه باید داده‌هایی شبیه به هم قرار گیرند و کمترین شباهت را با داده‌های موجود در دیگر خوشه‌ها دارا باشند(صادقیان، ۱۳۹۲).

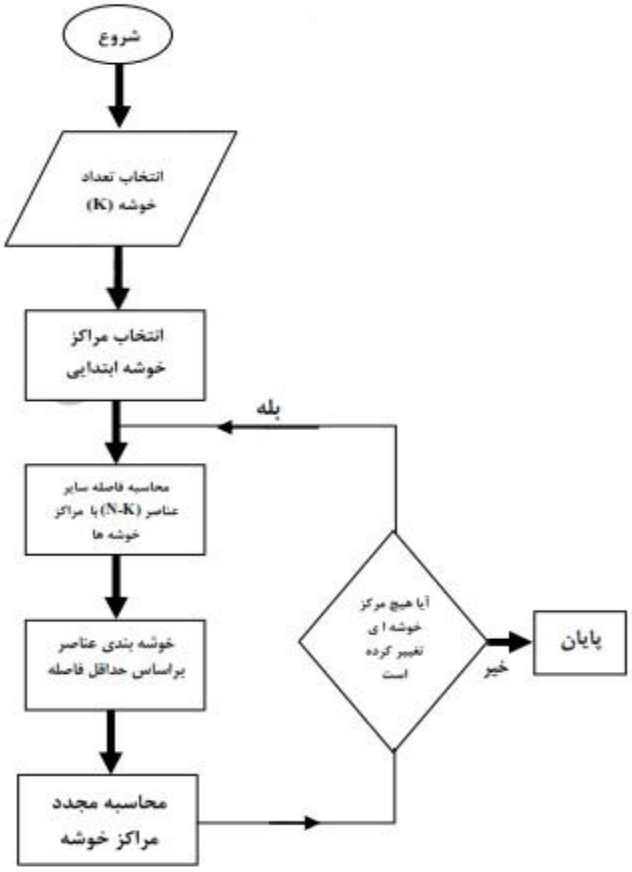
الگوریتم Kمیانگین یکی از روش‌های خوشه‌بندی است که اولین بار توسط مک‌کوئین(۱۹۶۷) ارایه شد. این روش برای خوشه‌بندی داده‌هایی طراحی شده است که به صورت کمّی باشند و خوشه دارای مرکزی به نام میانگین است. در این روش ابتدا اشیا به صورت تصادفی به k خوشه تقسیم می شوند و در گام بعدی فاصلهٔ هر یک از اشیا از مرکز خوشه خود محاسبه می شود. در صورتی که فاصلهٔ شی مورد نظر از میانگین خوشهٔ خود زیاد و به خوشه دیگری نزدیک‌تر باشد، این شی به خوشه‌ای که نزدیک‌تر است اختصاص می‌یابد. این کار آنقدر تکرار می‌شود تا تابع خطا حداقل شود و یا اعضای خوشه تغییر نیابد. در شکل (۱) مراحل اجرای الگوریتم kمیانگین به طور خلاصه نشان داده شده است.

اگر D مجموعهٔ داده‌ها با n شـی باشـد و C_i بیانگر خوشـه‌ها به تعداد k خوشـه باشـد، در این صـورت تابع خطا(EF) مجموع فواصل هر شی از مرکز خوشهٔ خودش به شکل رابطه (۱) تعریف می‌شود:

$$EF = \sum_{i=1}^{k} \sum_{X \in C_i} d(X, \mu(C_i)) \qquad (۱)$$

که در آن μ نشان‌دهندهٔ مرکز خوشه و $d(X,\mu(C_i))$ فاصلهٔ هر شی از مرکز خوشه است. فاصلهٔ هر شی از مرکز خوشهٔ خود می‌تواند بر پایهٔ اقلیدسی محاسبه شود. به دلیل آن که در خوشه‌بندی‌های مرکزگرا، تابع خطایی وجود دارد که لازم است حداقل شود، می‌توان به مسایل خوشه‌بندی مرکزگرا به دید مسایل بهینه‌سازی نگریست(مومنی، ۱۳۹۳).

از آنجا که در روش kمیانگین ساده، مراکز خوشه‌ها به صورت تصادفی انتخاب می شود، نتیجهٔ خوشه‌بندی وابسته به انتخاب مراکز اولیه است که از نقاط ضعف این روش به شمار می‌رود. برای اصلاح این موضوع می‌توان از الگوریتم‌های دیگری در ترکیب با kمیانگین استفاده کرد که با انتخاب مناسب‌ترین مراکز خوشه‌ها، تابع خطای فاصله رابطهٔ (۱) را حداقل کند.



شکل ۱. الگوریتم خوشه‌بندی kمیانگین

خوشه بندی سلسله مراتبی یکی از پرکاربردترین روش‌هـــای خوشـــه بندی اســـت. در ایـــن روش ابتدا فاصله دو به دوی مشاهدات از هم محاسبه می شود. پس از تعیین فاصله دو بـــه دوی مشاهدات، با توجه به نزدیکی مشاهدات نسبت به یکدیگر، مشاهدات با هم تشکیل یـــک خوشهٔ جدید می‌دهند. این کار تا جایی پیش مـــی‌رود کـــه تمـــام مشـــاهدات در یـــک خوشـــه قرار می‌گیرند(مؤمنی، ۱۳۹۵). در این مقاله از این الگوریتم برای تعیین مراکز خوشه‌ها استفاده شده است. به طوری که ابتدا فاصله هر یک از مشاهدات به صورت دو به دو با سایر مشاهدات اندازه‌گیری شده و براساس تعداد خوشه‌های تعیین شده، مشاهداتی که کمترین فاصله را با بیشترین تعداد مشاهدات دارند به عنوان مراکز ابتدایی انتخاب می‌شـوند و سـایر مراحل مانند روش kمیانگین انجام می‌شـود. در واقع در این روش تابع خطا برای تمام حالات به عنوان مرکز انتخاب شدن همهٔ مشاهدات محاسبه می‌شود. این روش به دلیل محاسبه فاصله برای تک‌تک مشاهدات برای زمانی که تعداد مشاهدات بسیار زیاد به دلیل حجم بالای محاسبات توصیه نمی شود. یکی از ابتکارات این مقاله استفاده از الگوریتم سلسله مراتبی در ترکیب با روش kمیانگین است.

روش دیگری که برای بهبود عملکرد خوشـه‌بندی به روش kمیانگین در مطالعات پیشـنهاد شـده اسـت، ترکیب آن با الگوریتم فراابتکاری کلونی مورچه‌ها اســـت. الگوریتم‌های فراابتکاری به الگوریتم‌هایی گفته می‌شـــود که معمولا از طبیعت الهام گرفته شده‌اند و به طور متو سط بهترین تقابل کیفیت و زمان حل م سئله را به همراه دارند. اما به دلیل سـاختار هوشـمندانه، محاسـبات زیادی دارند که بسـیار زمان‌بر اسـت. از این رو، برای حل مشـکل سـرعت پایین و همگرایی دیررس این الگوریتم‌ها از ترکیب آنها با الگوریتم kمیانگین استفاده می‌شود که موجب حل مشکل همگرایی زودرس و بیرون آمدن از بهینهٔ محلی به بهینهٔ فراگیر در الگوریتم kمیانگین ســـاده می‌شـــود و نیاز به تعیین تعداد خوشـــه‌ها از ابتدا نیز ندارد(گلی‌چنار و صـنیعی آباده، ۱۳۹۱). در این روش هر مورچه در یک گره از گراف مربوط به م سئله قرار می‌گیرد و سپس به صورت احتمالی به گرهٔ دیگری حرکت کرده و به همین ترتیب ادامه می‌دهد تا یک جواب کامل ایجاد گردد. در این روش مراکز خوشه‌ها براساس مقدار فرومونی که مورچه‌ها هنگام حرکت در مسیر بین مشـــاهدات برای جای می‌گذارند تعیین می‌شـــود(زارع احمدآبادی و همکاران، ۱۳۹۵). پس از تعیین مراکز و تعداد خوشه‌های بهینه سایر مراحل مانند روش kمیانگین خواهد بود.

در ادامه با ا ستفاده از سه روش بیان شده، شرکت‌های دانش‌بنیان بخش ک شاورزی با ا ستفاده از داده‌های معاونت علمی ریا ست جمهوری در سال ۱۳۹۵ خو شه‌بندی و برا ساس مقای سهٔ نتایج کارآمدترین نوع خو شه‌بندی برای این داده‌ها انتخاب خواهد شد. برای انجام خوشه‌بندی و اجرای الگوریتم‌ها از بستهٔ نرم‌افزاری متلب استفاده شده است.

نتایج

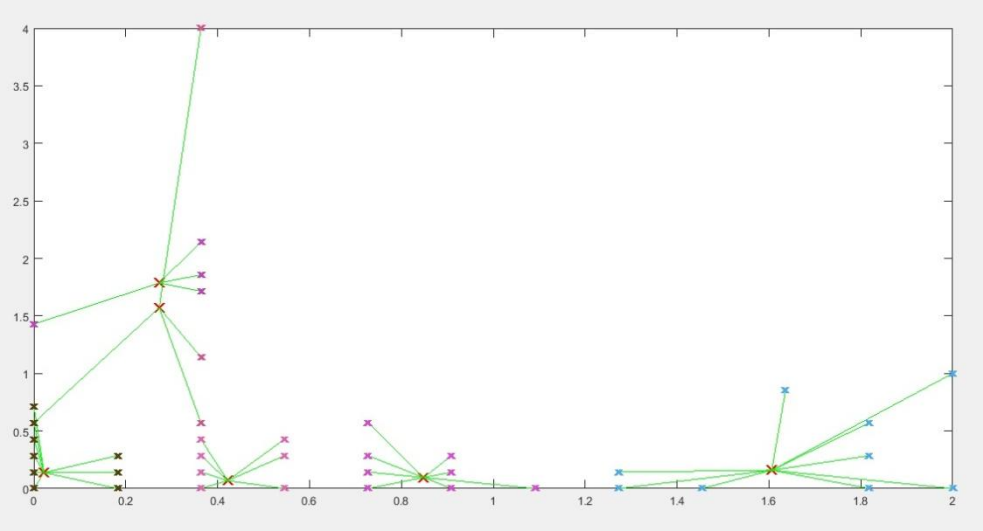
همان‌طور که پیش از این بیان شد، هدف این مقاله خوشه‌بندی شرکت‌های دانش‌بنیان بخش کشاورزی است، به صورتی که شرکت‌های دارای بیشترین وجوه اشتراک در یک گروه قرار بگیرند. برای این منظور، نخست از روش خوشه‌بندی kمیانگین ساده، از رایج‌ترین ابزارهای خوشه‌بندی در مطالعات مختلف، استفاده شد؛ اما از آنجا که یکی از نقاط ضعف این روش تصادفی بودن انتخاب مراکز خوشه‌هاست، از این رو، دو روش kمیانگین ترکیبی با الگوریتم‌های سلسله مراتبی و کلونی مورچه‌ها نیز به کار گرفته شدند که نتایج خوشه‌بندی مربوط به سه روش در جدول (۱) قابل مشاهده است.

جدول۱. مقایسه انواع روش‌های خوشه‌بندی و اعتبار آنها

نوع خوشه‌بندی	شماره خوشه‌ها	۱	۲	۳	۴	۵	۶	۷	مجموعه گروه‌ها
k-means ساده	تعداد شرکت‌ها	۹۱	۱	۱	۳	۲	۳	۱	۱۰۲
	میانگین ضریب نیمرخ	۰/۹۹	۱	۱	۰/۸۵	۱	۱	۱	۰/۹۹
k-means ترکیبی با الگوریتم سلسله مراتبی	تعداد شرکت‌ها	۵۰	۲۶	۱۸	۵	۳			۱۰۲
	میانگین ضریب نیمرخ	۰/۷۴	۰/۷۹	۰/۶۴	۰/۲۵	۰/۶۱			۰/۷۱
k-means ترکیبی با الگوریتم کلونی مورچگان	تعداد شرکت‌ها	۳۳	۲۵	۱۸	۴	۴	۱۸		۱۰۲
	میانگین ضریب نیمرخ	۰/۶۱	۰/۷۶	۰/۷۵	۰/۵۶	۰/۷۴	۰/۷۶		۰/۷۰

ماخذ: یافته‌های پژوهش

در روش kمیانگین ساده، براساس ضریب شاخص نیمرخ ۹۹ درصدی، مناسب‌ترین تعداد خوشه‌ها ۷ بود، در حالی که تعداد بهینهٔ خوشه‌ها در روش kمیانگین ترکیبی با الگوریتم سلسله مراتبی و کلونی مورچه‌ها به ترتیب ۵ و ۶ با ضرایب نیمرخ ۷۱ و ۷۰ درصد بود. با وجود آنکه، ضریب نیمرخ در روش kمیانگین ساده در مقایسه با دو روش دیگر بزرگ‌تر است، اما توزیع نامتوازن شرکت‌ها بین خوشه‌ها و وجود سه خوشهٔ تک‌عضوی موجب می‌شود، دو روش دیگر که هم از نظر آماری معتبر هستند و هم توزیع متوازن‌تری از شرکت‌ها در خوشه‌ها نشان می‌دهند، مناسب‌تر تشخیص داده شوند. در بین دو روش kمیانگین ترکیبی با الگوریتم سلسله مراتبی و کلونی مورچه‌ها نیز روش kمیانگین ترکیبی با الگوریتم کلونی مورچه‌ها که میانگین ضرایب نیمرخ در تمام گروه‌های آن بیشتر از ۵۰ درصد است و بیان می‌کند، اعضای هر خوشه بیشترین تعلق را به آن خوشه دارند و توزیع متوازن‌تری را نشان می‌دهد، انتخاب می‌شود. در نمودار (۱) نگاشت گرافیکی نقاط در خوشه‌بندی kمیانگین ترکیبی با الگوریتم کلونی مورچه‌ها در انتخاب مراکز خوشه‌ها نشان داده شده است.



نمودار ۱. نگاشت گرافیکی نقاط در در خوشه‌بندی kمیانگین ترکیبی با کلونی مورچه‌ها

همان‌طور که در جدول (۲) مشاهده می‌شود، براساس روش kمیانگین ترکیبی با کلونی مورچه‌ها، بیشترین تعداد شرکت‌ها و محصولات در خوشه یک و کمترین تعداد در خوشه‌های چهار و پنج که به ترتیب با ۲۷۲۹۳ و ۵۴۰۴ میلیون ریال دارای بیشترین میانگین فروش محصول هستند قرار دارند. دلیل پایین بودن میانگین فروش در خوشهٔ سوم آن است که بیشتر اعضای این خوشه، شرکت‌های دانش‌بنیان نوپا هستند که هنگام دریافت گواهی دانش‌بنیان، محصول آنها به مرحلهٔ فروش نرسیده بوده است. علت اندک بودن تعداد اعضای این خوشه‌ها آن است که تعداد کمی

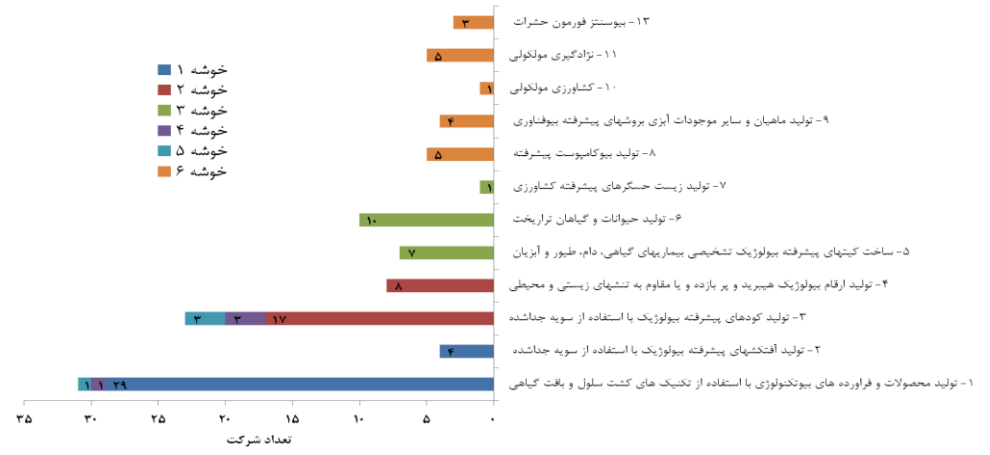
از شرکت‌ها دارای تعداد زیادی محصول دانش‌بنیان با فروش بالا هستند و اغلب شرکت‌ها با تعداد کمی محصول بازار کوچکی را کسب کرده‌اند که بالطبع لازم است، شرایط این شرکت‌ها در استفاده از ابزار حمایتی، متفاوت با دیگر شرکت‌ها باشد.

جدول ۲. وضعیت توزیع شرکت‌ها، محصولات و فروش آنها در خوشه‌بندی Kمیانگین ترکیبی با کلونی مورچه‌ها

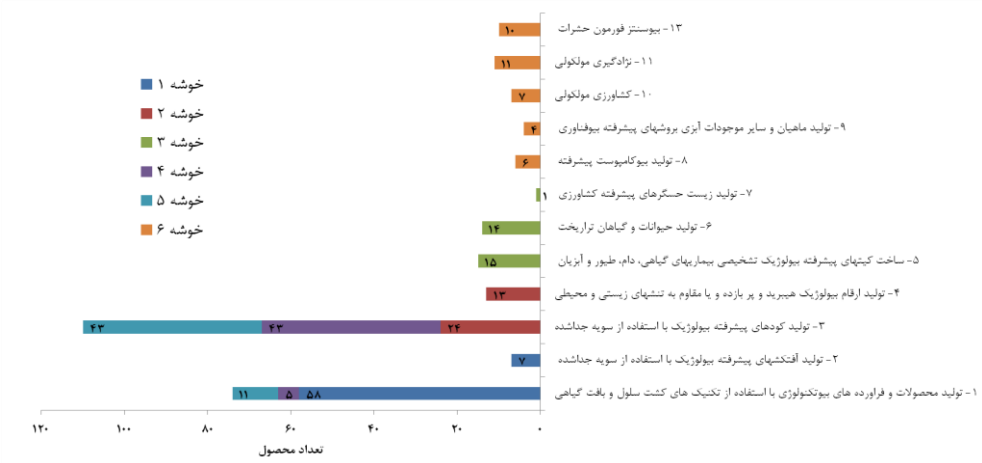
شماره خوشه	۱	۲	۳	۴	۵	۶	مجموعه گروه‌ها
تعداد شرکت	۳۳	۲۵	۱۸	۴	۴	۱۸	۱۰۲
تعداد محصولات	۶۵	۳۷	۳۰	۴۸	۵۴	۳۸	۲۷۲
مجموع مقدار فروش (میلیون ریال)	۷۳۵۴	۷۸۶۴	۵۰۲	۱۰۹۱۷۲	۲۱۶۱۵	۹۲۱۲	۱۵۵۷۱۹
میانگین مقدار فروش (میلیون ریال)	۲۲۳	۳۱۵	۲۸	۲۷۲۹۳	۵۴۰۴	۵۱۲	۱۵۲۷

ماخذ: یافته‌های پژوهش

همان‌طور که بیان شد، براساس دسته‌بندی محصولات دانش‌بنیان در معاونت علمی و فناوری ریاست جمهوری، دستهٔ محصولات کشاورزی، دامی و گیاهی خود دارای سیزده زیردسته است که تاکنون در زیردستهٔ دوازده، یعنی تولید دارو و واکسن بیولوژیک برای دام و طیور، محصولاتی ثبت نشده است. توزیع شرکت‌ها و محصولات براساس خوشه‌بندی انجام شده، در هر یک از دوازده دستهٔ دارای محصول به ترتیب در نمودارهای (۲) و (۳) نشان داده شده است. از ۳۳ شرکت عضو خوشهٔ یک، ۲۹ شرکت عضو زیردستهٔ یک، تولید محصولات و فراورده‌های بیوتکنولوژی با استفاده از تکنیک‌های کشت سلول و بافت گیاهی، هستند. بیشتر شرکت‌ها و محصولات عضو دسته‌های سه و چهار، تولید کودهای پیشرفتهٔ بیولوژیک و ارقام بیولوژیک هیبرید و پربازده، نیز به خوشهٔ دو تعلق دارند. خوشه‌های سه و شش به ترتیب با ۳۰ و ۳۸ محصول و ۱۸ شرکت بیشترین پراکندگی را در زیردسته‌ها دارند. همهٔ شرکت‌های زیردسته‌های پنج تا هفت که اغلب شرکت‌های نوپا هستند، در خوشهٔ سه و زیردسته‌های هشت تا سیزده در خوشهٔ شش قرار دارند. اعضای خوشه‌های چهار و پنج که اغلب شرکت‌های پرسابقه با تنوع محصول و درآمد بالا هستند در زیردسته‌های اول و سوم قرار دارند. در دستهٔ کودهای پیشرفته، سه شرکت خوشه‌های چهار و پنج، هریک ۴۳ محصول دارند.



نمودار ۲. توزیع شرکت‌ها در خوشه‌ها به تفکیک زیردسته‌های فناوری



نمودار ۳. توزیع محصولات در خوشه‌ها به تفکیک زیردسته‌های فناوری

نتیجه‌گیری و پیشنهادها

هدف این مقاله خوشه‌بندی شرکت‌های دانش‌بنیان کشاورزی براساس سه ویژگی زیردستهٔ فناوری، تعداد محصولات دانش‌بنیان شرکت و مقدار فروش این محصولات در شرکت بود.

برای د ستیابی به این هدف، از سه روش خو شه‌بندی تفکیکی kمیانگین ساده و kمیانگین در ترکیب با الگوریتم‌های سلسله مراتبی و کلونی مورچه‌ها استفاده شد و براساس شاخص ضریب نیمرخ و توزیع متوازن شرکت‌ها در خوشه‌ها، روش kمیانگین ترکیبی با الگوریتم کلونی مورچه ها به عنوان بهترین روش برای این مطالعه انتخاب شـــد که در آن

شرکت‌های دانش‌بنیان کشاورزی ایران به شش خوشه تقسیم‌بندی شدند. در ادامه با توجه به نتایج به دست آمده، پیشنهادهایی ارایه می‌شود.

همان‌طور که مشاهده شد، اعضای هر یک از خوشه‌ها ویژگی‌های مشابهی دارند که آنها را از سایر خوشه‌ها متمایز می‌کند. از این رو، شرکت‌های درون هر خوشه، حمایت‌های ویژهٔ خود را نیاز دارند.

شرکت‌های خوشهٔ سه که دارای میانگین درآمد پایین‌تری بوده و اغلب جزو شرکت‌های دانش‌بنیان نوپا هستند نیازمند حمایت‌هایی در جهت ورود محصول به بازار دارند و در مقابل، شرکت‌های خوشه‌های چهار و پنج که تنوع محصول و مقدار فروش بیشتری دارند، با ریسک کمتری برای ادامهٔ فعالیت خود روبه‌رو هستند و پیشنهاد می‌شود با حمایت‌های حقوقی و بازرگانی به ورود این شرکت‌ها به بازار صادراتی کمک شود. از آنجا که شرکت‌های خوشه‌های یک، دو، چهار و پنج بیشترین عضو را در زیردسته‌های یک تا چهار دارند که محصولات آنها اغلب نهاده‌های مصرفی بخش کشاورزی مانند بذر، کود و آفت‌کش‌ها هستند، پیشنهاد می‌شود، شرایطی فراهم شود تا ارتباط این شرکت‌ها با بازار هدف که کشاورزان هستند تسهیل گردد. برای تحقق بخشیدن این امر، برگزاری نمایشگاه‌هایی برای معرفی محصولات در مناطق مختلف کشور پیشنهاد می شود. برای اعضای خوشهٔ شش که زیردستهٔ فناوری محصولات آن اغلب دورهٔ آزمایشگاهی پیش از تولید طولانی‌تری دارند، تقویت حمایت‌های فنی از جمله تسهیلات آزمایشگاهی و ایجاد ارتباط با سرمایه‌گذاران برای پروژه‌های تحقیق و توسعه محصولات این شرکت‌ها از اهمیت فراوانی برخوردار است.

از آنجا که محصولات شرکت‌های دانش‌بنیان، فناورانه است و بخش کشاورزی در ایران یک بخش سنتی است، با وجود نیاز مبرم بخش به محصولات این شرکت‌ها، ممکن است این شرکت‌ها با چالش تقاضا در بازار روبه‌رو شوند که لازم است خود شرکت‌ها با فراگیری راهبردهای بازاریابی و دولت با کمک به آنها در زمینهٔ معرفی محصول و توسعهٔ بازار طرف تقاضا در بازار محصولات دانش‌بنیان بخش کشاورزی را تقویت کند. یکی از راهبردهای پیشنهادی برای این امر، تشکیل نهادی تخصصی برای توسعهٔ بازار این محصولات است که فعالیتی شبیه به سازمان توسعه و تجارت در بازار داخلی داشته باشد.

منابع

زارع احمدآبادی، ح.، امام، م، ناصر صدرآبادی، ع. (۱۳۹۵). تبیین الگوی خوشه‌بندی بازار هدف، شماره ۱۶، صفحه ۱۷-۳۶.

سلمانی، ب. (۱۳۹۴). «پارک‌های علم و فناوری سازوکاری برای اقتصاد دانش‌بنیان»، نشر نور علم، تهران.

شیرکوند، س.، جوکار، ق. و مزیدی، ع. (۱۳۹۳). خوشه‌بندی مشتریان بیمه براساس ارزش‌های مورد انتظار، پژوهشنامه بیمه، سال ۲۹، شماره ۴، صفحه ۱۳۷-۱۶۳.

صادقیان، م. (۱۳۹۲). یافتن مرکز بهینه برای خوشه‌ها در الگوریتم K-means، پایان‌نامه کارشناسی ارشد، دانشگاه شیخ بهایی.

علی حیدری بیوکی، ط. و خادمی زارع، ح. (۱۳۹۴). توسعه روش تحلیل پوششی داده‌ها به منظور خوشه‌بندی مشتریان اعتباری بانک‌ها، مجله مدل‌سازی در مهندسی، سال ۱۳، شماره ۴۱، ص ۵۹-۷۴.

گلی‌چنار، ف. و صنیعی آباده،م. (۱۳۹۱). مروری بر روش‌های ترکیبی الگوریتم خوشه بندی k-means و الگوریتم های فرا مکاشفه ای جهت رویارویی با معایب الگوریتم k-means، همایش منطقه ای علوم کامپیوتر، مهندسی کامپیورتر و فناوری اطلاعات، فارس.

معاونت علمی و فناوری ریاست جمهوری. (۱۳۹۶). آمار شرکت‌های دانش‌بنیان، تهران.

معاونت علمی و فناوری ریاست جمهوری. (۱۳۹۶). فهرست کالاها و خدمات دانش‌بنیان، ویرایش چهارم، تهران.

مؤمنی، م. (۱۳۹۳). خوشه‌بندی داده‌ها، تهران، مؤلف.

Alfansi, L. and Sargeant, A, (2000). Market segmentation in Indonesian banking sector: The relationship between demographics and desired customer benefits. International Journal of Bank Marketing, 18(2), pp 86-98.

Brandt, A., (2005). Cluster analysis for market segmentation. Department of Curriculum Teaching and Learning, Toronto, University of Toronto. Master of Art, p.60.

Chen, J., Z. Zhu and H. Y. Xie (2004). Measuring intellectual capital: a new model and empirical study, Journal of Intellectual Capital, 5(1): 195-212.

Epetimehin, F.M., (2011). Market segmentation: A tool for improving customer satisfaction and retention in insurance service delivery .Journal of Emerging Trends in Economics and Management Sciences (JETEMS), 2(1), pp.62-67.

Linares-Mustarós, S., Coenders, G., Vives-Mestres, M. (2018). “Financial performance and distress profiles. From classification according to financial ratios to compositional classification”. Advances in Accounting, DOI: 10.1016/j.adiac.2017.10.003

OECD. (2007). Competitive Regional Clusters. Paris: OECD.

Rahman Md., K.Z. Arefin · S. Masud · S. Sultana · R.M. Rahman. (2017). Analyzing Life Insurance Data with Different Classification Techniques for Customers’ Behavior Analysis, Advanced Topics in Intelligent Information and Database Systems, Studies in Computational Intelligence 710, DOI 10.1007/978-3-319-56660-3_2

Clustering of Agricultural Knowledge Based Companies with Metaheuristic Algorithms

Abstract

Clustering is one of the most important operations in data mining, and its results are useful for researchers and policy makers in various fields for analysis and planning. Since in recent years, the knowledge based economy has been developing with the support of knowledge-based companies in Iran, the analysis of the characteristics of these companies and their clustering for effective planning provides an appropriate perspective for policy makers. Hence, in this article, Iranian knowledge based companies of agriculture has been clustered based on the field of technology, the number of products and value of product sales, using three clustering methods, simple K-Means and K-Means in combination with hierarchical and ant colony algorithms. According to the results, the clustering using K-Means in combination with the ant colony algorithm, in comparison with the other two methods, has a more balanced distribution of the firms among six clusters, and the average Silhouette width value of the 0.7 confirms the validity of this clustering. It is suggested that similar support programs be designed for companies that are due to similar characteristics in a cluster in order to be effective and avoid losing budget. For knowledge based companies with low average sales that are in a cluster, market support programs, and for companies with product diversification and high average incomes, export market development programs are recommended.

JEL Classification: ***C81, L2, M13, O31, Q16***

Keywords: Knowledge-based Economics, Data Mining, K-Means Clustering, Hierarchical Algorithm, Ants colony Algorithm