

استفاده از روش خوشه بندی ترکیبی کی - مینز و الگوریتم ژنتیک برای کنترل کیفیت نخ ریسیده شده در صنعت نساجی

مرضیه نظریان^۱؛ پدram پیوندی^۲

چکیده

امروزه با گسترش سیستم های پایگاهی و حجم بالای داده های ذخیره شده در این سیستم ها، نیاز به ابزاری است تا بتوان داده های ذخیره شده پردازش کرد. داده کاوی یکی از مهمترین این روشها است که به وسیله آن الگوهای مفید در داده ها با حداقل دخالت کاربران شناخته می شوند و اطلاعاتی را در اختیار کاربران و تحلیل گران قرار می دهند تا براساس آنها تصمیمات مهم و حیاتی اتخاذ شوند. یک رهیافت متداول در این زمینه روش خوشه بندی است که برای تصمیم گیری یا طبقه بندی می تواند تصمیمات نمادینی را به نمونه های جدید با استفاده از نمونه های موجود منتسب کند.

هدف از این مقاله استفاده از روش های داده کاوی به عنوان یک تکنیک به منظور بهبود کیفیت نخ ریسیده شده در صنعت نساجی می باشد. بدین منظور با استفاده از داده های موجود در یک کارخانه ریسندگی الیاف کوتاه و همچنین با بکارگیری روش داده کاوی ترکیبی کی - مینز و الگوریتم ژنتیک روابط بین متغیرهای ریسندگی و پارامترهای کیفی نخ تولیدی آنالیز و بررسی شد. نتایج نشان داد که متغیرهای ریسندگی تاثیر بسزایی در کیفیت نخ تولیدی دارد.

کلمات کلیدی

خوشه بندی، کی - مینز، الگوریتم ژنتیک، کنترل کیفیت، نخ ریسیده شده.

Using Of K-Means Hybrid Clustering Method And Genetic Algorithm For Quality Control Of Staple Yarn In Textile Industry

Marzieh Nazarian; Pedram Payvandy

Master of Science in Textile engineering, Yazd University, Yazd, Iran

Professor Assistant, Department of Textile Engineering, Yazd University, Yazd, Iran

ABSTRACT

Today with the high volume of database systems and stored data in these systems, there is a need for processing the stored data and. Data mining is one of the most important methods that by using of it, useful patterns are diagnosed in data with minimum users interferences and is given information to users and analyzers until base on are taken important decisions. One usual solution in this course is clustering method that can be related symbol Decisions to new samples by using of exiting samples for classification.

The purpose of this paper is using of data mining methods as a technique for improving quality of staple yarn in textile industry. In order to, by using of exiting data in the spinning factory of short staple fiber and also by using of K-Mean hybrid method and Genetic algorithm, relations between spinning variables and quality parameters of producing yarn are analyzed. The results show that spinning variables have important influence on quality of producing yarn.

KEYWORDS

Clustering, K-Means, Genetic algorithm, Quality control, Spun yarn^۲

۱. دانشجوی کارشناسی ارشد مهندسی تکنولوژی نساجی، دانشگاه یزد. nazari_techno@yahoo.com

۲. استادیار دانشکده مهندسی نساجی دانشگاه یزد. دانشگاه یزد. Peivandi@yazduni.ac.ir

حجم انبوه داده ها از یک طرف و نیاز به تبدیل این داده ها به اطلاعات از طرف دیگر وجود ابزاری قدرتمند برای تحلیل داده ها را لازم می سازد. در این میان روش های داده کاوی مطرح می شوند که به هر نوع استخراج دانش یا الگو از داده های موجود در یک پایگاه داده اطلاق می گردد. بطور معمول این دانش و الگوهای مورد نظر در داده ها مستتر بوده و نیز اطلاعات بی ارزشی نیستند. از داده کاوی می توان جهت دسته بندی و تخمین و پیش بینی، خوشه بندی و همچنین یافتن روابط میان داده ها استفاده کرد [۹]. چنانچه می دانیم روش های آماری به تنهایی قادر به انجام این کارها نیستند و بدین ترتیب نیاز به کاوش در میان داده ها بیش از پیش احساس می شود و بدین ترتیب شاخه داده کاوی در طی تحقیقات گسترده ای در زمینه یادگیری ماشین پا به عرصه وجود نهاد.

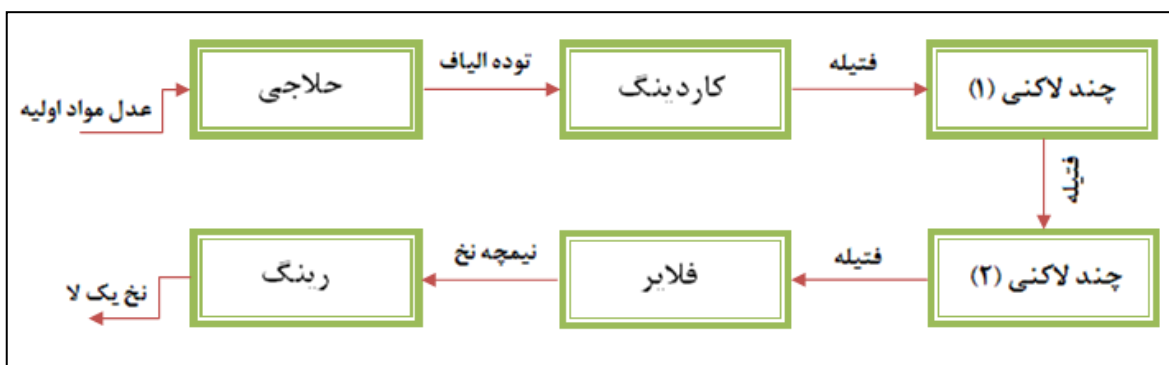
داده کاوی ترکیبی از مفاهیم و الگوریتم ها از آمار، هوش مصنوعی، یادگیری ماشینی و مدیریت داده است که در حوزه های مختلفی همچون امور حاکمیتی، بانکداری، مالی، توزیع و فروش، بازاریابی، بیمه، علوم، مهندسی و... برای کشف روابط و الگوهای پنهان بکار می رود. صنعت نیز یکی از حوزه هایی است که داده کاوی می تواند در آن مورد استفاده قرار گیرد و باعث بدست آمدن مزایای فراوان برای سازمان ها شود.

در زمینه صنعت نساجی نیز تحقیقاتی در مورد تکنیک ها و روش های داده کاوی انجام شده است. Hodge و همکارانش در سال ۲۰۰۱ از تکنیک داده کاوی به منظور پیش بینی عوامل موثر بر میزان عایق بودن پوشاک استفاده کردند [۷]. در سال ۲۰۰۲ با استفاده از روش داده کاوی در صنعت نساجی به طور خاص کارخانه ریسندگی پنبه مورد تجزیه و تحلیل قرار گرفته شد [۱۷]. در سال ۲۰۰۶ به منظور نشان دادن کاربرد مهم روش داده کاوی در صنعت از داده های جمع آوری شده در یک کارخانه ریسندگی الیاف کوتاه برای بررسی عوامل موثر بر کیفیت نخ تولیدی استفاده شد [۱]. در سال ۲۰۰۸ برای کنترل کیفیت نخ ریسیده شده از ابزارهای داده کاوی، خوشه بندی و در آخر از تجزیه و تحلیل های آماری استفاده شد [۴]. در سال ۲۰۱۰ با استفاده از روش داده کاوی یک کارخانه ریسندگی الیاف کوتاه مورد تجزیه و تحلیل قرار داده شد [۱۰]. SHAHRABI و همکارش در سال ۲۰۱۰ با استفاده از روش داده کاوی یک سیستم جدید سایزبندی توسعه داد [۱۸]. در سال ۲۰۱۰ با استفاده از تکنیک داده کاوی یک سیستم سایز بندی برای پسران ایرانی دارای سنین بین ۱۶ و ۲۲، بر اساس نوع بدن مشخص گردید [۲]. نتایج نشان داد که سه نوع بدن و سیستم سایزبندی دارای عملکرد مناسبی بودند. در تحقیقی که در سال ۲۰۱۱ انجام گرفت کارایی و مانیتورینگ دیگ بخار در صنعت نساجی با استفاده از تکنیک های داده کاوی مورد ارزیابی قرار گرفت [۲۱]. S.K. Tyagi و همکارش در سال ۲۰۱۱ از ابزارها و تکنیک های داده کاوی برای کنترل کیفیت صنعت نساجی به منظور تصمیم گیری استراتژیکی و مدیریتی استفاده کردند [۱۹]. در سال ۲۰۱۲ از تکنیک داده کاوی به منظور بهبود مدیریت، ظرفیت تولید و کیفیت در صنعت نساجی استفاده شد [۱۴]. در سال ۲۰۱۲ یک کار تحقیقاتی در زمینه تجزیه و تحلیل سایز بندی پیراهن برای مردان هندی ۲۵ تا ۶۶ ساله با استفاده از تکنیک خوشه بندی صورت گرفت [۱۳].

هدف از انجام این مقاله ارائه یک روش موثر ترکیبی کی - مینز و الگوریتم ژنتیک برای روش خوشه بندی داده ها می باشد. از الگوریتم پیشنهادی برای کنترل کیفیت نخ ریسیده شده در صنعت نساجی استفاده شد. در این مطالعه به ارزیابی تاثیر پارامتر های مربوط به مواد خام و پارامترهای فرآیند تولید بر روی کیفیت نخ ریسیده شده پرداخته شد. نتایج نشان داد که الگوریتم پیشنهادی از نظر کیفیت جواب های حاصل شده کاملاً برتر می باشد.

۲. ریسندگی الیاف کوتاه

ریسندگی الیاف کوتاه شامل کلیه عملیاتی است که بر روی الیاف پنبه والیاف مصنوعی مانند پلی استر انجام می شود تا آن را تبدیل به نخ یک لای ریسیده شده کند. کارخانه مورد بررسی دارای خط ریسندگی الیاف کوتاه می باشد. تولید این کارخانه نخ صد در صد پنبه ای، مخلوط الیاف پنبه و پلی استر، مخلوط الیاف پنبه و نایلون، مخلوط الیاف ویسکوز و پلی استر و مخلوط الیاف پنبه و ویسکوز با درصد های متفاوت و نمرات متفاوت می باشد. مراحل تولید نخ در این کارخانه در شکل ۱ نشان داده شده است.



شکل (۱) مراحل فرایند تولید نخ در کارخانه ریسندگی الیاف کوتاه

درجه مرغوبیت الیاف، غالباً نسبت به طول و میزان درصد مواد خارجی در آنها تعیین می شود و واضح است که هر اندازه الیاف تمیز تر و دارای طول بلندتری باشند، مرغوبترند. بخصوص الیاف بلندتر از لحاظ نساجی اهمیت ویژه ای دارد و می توان از آنها نخهای خوب و ظریف تهیه کرد. به طوری که هر اندازه طول الیاف بلندتر باشد، قطر آن کمتر بوده و این یک خاصیت بسیار خوب دیگری است که در الیاف پنبه بلند وجود دارد و در مورد ریسندگی برای تهیه نخهای ظریف و بادوام حائز اهمیت است. قسمت حلاجی، اولین مرحله از عملیات ریسندگی محسوب می شود و توده الیافی که به صورت بسته هایی به نام عدل وارد کارخانه می شوند، در این قسمت به وسیله ماشین آلات مخصوصی باز و مخلوط شده و ضمن انجام این عملیات به مقدار قابل توجهی مواد خارجی از الیاف جدا و تمیز می شوند. توده الیافی که به صورت بالش درآمد و یا به طریقه سیستم شوت فید عمل شده به ماشین کارد تغذیه می شود. هدف از استفاده این ماشین باز کردن و تمیز کردن الیاف، جدا کردن ناخالصیهایی که در قسمت حلاجی، جدا کردن الیاف کوتاه، موازی کردن الیاف گرفته نشده، مخلوط کردن الیاف، تولید فتیله کارد می باشد. فتیله تولید شده توسط ماشین کارد، دارای الیافی نایکناخت و غیر موازی است که لازم است توسط مکانیزم مخصوصی موازی، یکنواخت و نیز مخلوط شوند، بنابراین، از ماشین چند لاکنی فتیله می توان برای هدفهای فوق استفاده کرد. معمولاً بعد از تولید فتیله نمی توان آن را مستقیماً به ماشین رینگ تغذیه کرد و برای تقلیل وزن مخصوص خطی فتیله و تولید محصولی به نام نیمچه نخ (Roving) باید مرحله نیم تاب انجام گیرد. ضمناً در این ماشین به نیمچه نخ مقدار معینی تاب داده می شود تا استحکام نیمچه نخ در اثر اصطکاک بین الیاف افزایش پیدا کند. نیمچه نخ تولید شده به وسیله فلایر باید توسط ماشین رینگ تبدیل به نخ بشود. در ماشین رینگ عملیاتی مانند تقلیل وزن مخصوص خطی نیمچه نخ به وسیله کشش دادن الیاف، تابیدن فتیله کشش داده شده و تبدیل به نخ، پیچیدن نخ تولید شده بر روی ماسوره انجام می گیرد.

۳. الگوریتم K-Means

یکی از روش های حیاتی کنترل و مدیریت داده ها، خوشه بندی داده ها با خواص مشابه، درون مجموعه ای از دسته ها می باشد. خوشه بندی یک روش دسته بندی کردن داده بر اساس شباهت می باشد [۲۲].

خوشه بندی داده یکی از رایج ترین تکنیک های داده کاوی است. خوشه بندی معمولاً به عنوان پیش درآمدی برای بکارگیری سایر تحلیل های داده کاوی یا مدل سازی به کار می رود. آنالیز خوشه بندی مجموعه داده را به نحوی گروه بندی می کند که شباهت درون خوشه ای بیشترین و شباهت بین خوشه ای کمترین باشد، به عبارت دیگر، داده ها موجود در یک خوشه بر اساس یک معیار شباهت از پیش تعیین شده، شباهت بیشتری نسبت به یکدیگر دارند تا نسبت به داده های موجود در سایر خوشه ها. در دو دهه اخیر تلاش های فراوانی برای یافتن بهترین الگوریتم خوشه بندی برای مجموعه داده های بزرگ صورت گرفته است. علاوه بر این تحقیقاتی زیادی در مورد این موضوع صورت گرفته است [۱۲]. در خوشه بندی هیچ دسته از پیش تعیین شده ای وجود ندارد و داده ها صرفاً براساس تشابه گروه بندی می شوند و عناوین هر گروه نیز توسط کاربر تعیین می گردد. الگوریتم K-Means [۸] یکی از مشهورترین الگوریتم های خوشه بندی داده می باشد که تعداد خوشه ها به عنوان پارامتر مرودی K توسط کاربر از پیش تعریف می شود. این الگوریتم با تعداد K مرکز اولیه که به طور تصادفی از میان داده ها انتخاب شده اند شروع می شود و سپس به طور مرتب مراکز هر یک از خوشه ها به روز رسانی می شود تا زمانی که دیگر تغییری در مراکز خوشه ها ایجاد نشود [۱۵]. در حال حاضر روش های متعددی برای خوشه بندی داده ها وجود دارد که بر اساس نوع داده ها، شکل خوشه ها، فاصله داده ها و غیره عمل خوشه بندی را انجام می دهند [۶].

۴. الگوریتم ژنتیک و استفاده آن در روش K-Means

الگوریتم ژنتیک یکی از الگوریتم‌های جستجو می باشد که به صورت تصادفی جواب مسئله را پیدا می کند. این الگوریتم که در دسته الگوریتم های آزمون و خطا جای می گیرد، اولین بار توسط هالند ارائه شد و بر اساس ژنتیک موجودات زنده و عوامل و شرایطی که برای زندگی و ادامه حیات آنها مؤثر است عمل می کند. پایه این روش نظریه داروین است که بر اصل حیات قویترین موجود در جامعه تأکید می کند [۵]. از این الگوریتم، در حل مسائل بهینه سازی پیچیده که نمی توان برای آن قوانین خاصی در نظر گرفت استفاده می شود. البته محققان اخیراً "روش هایی را توسعه داده اند که می توان از الگوریتم ژنتیک برای داده های کیفی نیز استفاده برد [۲۰]. برای حل یک مسئله با استفاده از این روش ابتدا باید پاسخ های فرضی مسئله را به صورت خاصی نمایش داد که برای ادامه کار و ارزیابی پاسخ ها مشکلی ایجاد نگردد. روشهای متعددی برای نمایش و کدگذاری وجود دارد که از مهمترین و معمولترین آنها روش دودویی و نمایش اعشاری شناور می باشد. در آغاز، جمعیت اولیه که جواب ها را نشان می دهند به صورت تصادفی انتخاب می گردد. هر یک از اعضای این جمعیت که کروموزم نامیده می شود یکی از پاسخ های مسئله می باشد. هر یک از این کروموزمها از رشته های از اعداد با طول برابر انتخاب می گردد که هر یک از این اعداد ژن نام دارد. الگوریتم ژنتیک بر اساس تکرار عمل می کند که به جمعیت در هر مرحله، نسل گفته می شود. هر یک از اعضای این نسل بر اساس تابع ارزش، ارزیابی می شوند.

در این الگوریتم ها نسل جدید سعی دارد که ارزش بیشتری از تابع ارزش را به خود اختصاص دهد و با این عملکرد به تابع هدف نزدیک تر گردد. در هر مرحله از تکرار، هر یک از کروموزم ها با احتمال خاصی با یکدیگر عمل تقاطعی انجام می دهند یا به اصطلاح مزدوج می شوند که پیامد آن یک یا چند کروموزم جدید به نام فرزند می باشد. در این فرزندان ممکن است طبق احتمال خاصی عمل جهش ژنی اتفاق افتد، به این صورت که مقدار یک یا چند ژن از کروموزم تغییر کند [۵]. در مرحله آخر فرزندان طبق تابع ارزش، ارزیابی شده و بر اساس ارزش آنها و ارزش والدین یعنی نسل اولیه که این فرزندان را تولید کرده اند نسل جدید تولید خواهد شد. این مراحل تا جایی تکرار می شود که نسل حاضر به جواب بهینه یا یکی از زیر جوابهای بهینه همگرا گردد. روشهای مختلفی برای عملگرهای تقاطعی و جهشی وجود دارد که با توجه به مسئله و پیچیدگی مربوطه یکی از آنها انتخاب می گردد.

۱.۴. نمایش کروموزم

قدم ابتدایی در شروع الگوریتم ژنتیک نمایش جواب ها به صورت رشته ایی به نام کروموزم ها می باشد. برای شروع باید از متغیرهای موجود، کروموزم تشکیل شود و موتورالگوریتم ژنتیک، یک جمعیت اولیه ناهمگن از کروموزمها را ایجاد نماید. سپس هر کروموزم مورد آزمایش قرار می گیرد. مناسب ترین کروموزم ها شانس بیشتری برای زنده ماندن در طول دوره های دیگر و دوباره تولید شدن را دارا بوده و ضعیف ترها محکوم به نابودی هستند. نوع کروموزم ها در این مقاله از نوع صحیح می باشد که تنها می توانند مقادیر ۱ الی K را به خود تخصیص دهند.

برای ایجاد نسل اول، از پارامترهایی شروع می شود که به آنها کروموزم یا ژنوم می گویند. کروموزم ها عموماً "به صورت یک رشته ساده از داده ها نمایش داده می شوند. البته انواع ساختمان داده های دیگر هم می توانند مورد استفاده قرار گیرند. در ابتدا چندین مشخصه به صورت تصادفی برای ایجاد نسل اول تولید می شوند. در طول هر نسل، هر مشخصه ارزیابی می شود و ارزش تناسب توسط تابع تناسب اندازه گیری می شود. گام بعدی ایجاد دومین نسل از جامعه است که بر پایه فرآیندهای انتخاب می باشد. از مشخصه های انتخاب شده همراه با عملگرهای ژنتیکی استفاده می شود: اتصال کروموزم ها به سر یکدیگر و تغییر برای هر فرد، یک جفت والد انتخاب می شود. انتخاب ها به گونه ای می باشد که مناسب ترین عناصر انتخاب می شوند تا حتی ضعیف ترین عناصر هم شانس انتخاب داشته باشند و از نزدیک شدن به جواب محلی جلوگیری شود.

از ترکیب دو کروموزم، فرزند ایجاد شده و به نسل بعدی اضافه می شود این مراحل طی می شود تا کاندید های مناسبی برای جواب، در نسل بعدی پیدا شوند. مرحله بعدی تغییر دادن فرزندان جدید است. کروموزم های فرزند به طور تصادفی تغییر می کنند یا جهش می یابند. به خصوص با جهش بیت ها در کروموزم های ساختمان داده ها این تغییرات ایجاد می شوند. این فرآیند باعث به وجود آمدن نسل جدیدی از کروموزم هایی می شود که با نسل قبلی متفاوت است. کل فرآیند برای نسل بعدی هم تکرار می شود. دوباره جفت ها برای ترکیب انتخاب می شوند، جمعیت نسل سوم به وجود می آید و ... این فرآیند تکرار می شود تا این که به آخرین مرحله برسیم.

۲.۴. استراتژی تولید جمعیت اولیه

جهت تولید جمعیت اولیه عموماً از روش تولید تصادفی کروموزم ها استفاده می شود. در روش تصادفی از آنجائیکه کروموزم ها متعلق به نواحی مختلف فضای جواب می باشند لذا تنوع کروموزم ها بالا است؛ در نتیجه در تکرارهای اولیه الگوریتم؛ تکامل نسل ها سریعتر انجام می گیرد ولی با افزایش تکرار، تشابه کروموزم ها نیز افزایش یافته تا اینکه در نهایت الگوریتم به یک یا چند حل شاخص همگرا گردد. در برخی از تحقیقات، از تکنیک

های ابتکاری یا فراابتکاری دیگر همانند SA یا TS نیز جهت بدست آوردن یک جمعیت اولیه با کیفیت بالا استفاده شده است. اگرچه اشکال عمده روش فوق افزایش احتمال همگرایی زودرس (Premature Convergence) یا کاهش تنوع در جمعیت می باشد.

۳,۴. روش انتخاب

یکی دیگر از عملگرهای اصلی الگوریتم ژنتیک، عملگر یا اپراتور انتخاب می باشد که به منظور گزینش کروموزوم های والدین از جمعیت کروموزوم ها و وارد نمودن آنها به مرحله تولید مثل و تولید کروموزوم های جدید می باشد. کروموزومی که دارای مقدار برازندگی مناسب تری باشد، احتمال بیشتری را برای ورود به مرحله تولید مثل دارد. بنابراین در هر تکرار با توجه به احتمال انتخاب هر کروموزوم، یک جفت کروموزوم انتخاب می گردند و با ورود به مرحله تولید مثل تحت اپراتورهایی مانند باز ترکیبی و جهش قرار می گیرند.

۴,۴. اپراتور باز ترکیبی

مهمترین عملگر در الگوریتم ژنتیک، عملگر باز ترکیبی است. ترکیب فرآیندی است که در آن نسل قدیمی کروموزومها با یکدیگر مخلوط و ترکیب می شوند تا نسل تازه ای از کروموزومها بوجود بیاید. جفت هایی که در قسمت انتخاب به عنوان والد در نظر گرفته شدند در این قسمت ژن هایشان را با هم مبادله می کنند و اعضای جدید بوجود می آورند. ترکیب در الگوریتم ژنتیک باعث از بین رفتن پراکندگی یا تنوع ژنتیکی جمعیت می شود زیرا اجازه می دهد ژن های خوب یکدیگر را بیابند.

۵,۴. اپراتور جهش

پس از اتمام عمل آمیزش، عملگر جهش بر روی کروموزومها اثر داده می شود. این عملگر یک ژن از یک کروموزوم را به طور تصادفی انتخاب نموده و سپس محتوای آن ژن را تغییر می دهد. اگر ژن از جنس اعداد دودویی باشد، آن را به وارونش تبدیل می کند و چنانچه متعلق به یک مجموعه باشد، مقدار یا عنصر دیگری از آن مجموعه را به جای آن ژن قرار می دهد. پس از اتمام عمل جهش، کروموزومهای تولید شده به عنوان نسل جدید شناخته شده و برای دور بعد اجرای الگوریتم ارسال می شوند.

۵. روش ارزیابی

۱,۵. مجموعه داده ها و کارهای انجام شده

داده ها از یک کارخانه ریسندگی الیاف کوتاه جمع آوری شدند. داده ها مربوط به مواد اولیه (شامل طول الیاف، ظرافت الیاف، نپ، نایکنواختی، نمره)، فرآیند تولید (شامل درصد پنبه و پلی استر مصرفی) و کیفیت (شامل نایکنواختی در ماشین چند لاکنی ۱ و ۲، فلایر و رینگ، نمره فتیله کاردینگ، چند لاکنی ۱ و ۲، فلایر و رینگ، ازدیاد طول نخ یک لا، استحکام نخ یک لا، تاب نخ یک لا، ضریب نایکنواختی ازدیاد طول نخ یک لا، ضریب نایکنواختی استحکام نخ یک لا، میزان نپ نخ یک لا، نقاط ضخیم نخ یک لا، نقاط نازک نخ یک لا، نایکنواختی نخ یک لا) می باشد.

۲,۵. الگوریتم ژنتیک ترکیبی با الگوریتم k-means

الگوریتم های ترکیبی زیادی براساس تئوری تکامل جهت بهبود خوشه بندی با روش k-means تاکنون ارائه شده اند. یکی از این روشها که همگرایی آن در بهینه کلی به صورت تئوری ثابت شده است، الگوریتم GMK می باشد که توسط K.Kirishna در سال ۱۹۹۹ معرفی گردیده است. عملکرد این الگوریتم براساس بهینه نمودن پارامتر مجموع واریانس درون خوشه ها (TWCV (total within cluster variation) می باشد که به آن مجموع مربعات خطای کلی (SE) نیز گفته می شود [۱۱]. که روش محاسبه آن به شرح زیر است.

(۱)

در معادله (۱):

n : تعداد داده ها (patterns)

d : تعداد پارامترهای هر داده (dimension)

k : تعداد خوشه ها

$\{x_i, i=1, 2, 3, \dots, n\}$: مجموعه داده ها می باشد.

در این صورت می توان ماتریس W را به صورت زیر تعریف نمود:

$$W \quad (2)$$

ماتریس $W = [w_{i,z}]$ دارای خصوصیات زیر می باشد:

$$w \quad (3)$$

اگر $C_k = (C_{k1}, C_{k2}, C_{k3}, \dots, C_{kd})$ مراکز پارامترهای مرکز خوشه C_k باشد در این صورت برای محاسبه مراکز خوشه ها خواهیم داشت:

$$C \quad (4)$$

با استفاده از معادله (۳) میزان واریانس درون هر خوشه برابر است با:

$$S \quad (5)$$

و در نهایت مجموع واریانس درون خوشه ها $TWCV$ برابر خواهد بود با:

$$S \quad (6)$$

که هدف پیدا کردن W^* می باشد که در آن $S(W^*)$ کمترین شود یعنی:

$$S \quad (7)$$

یکی از روش های معمول برای پیدا کردن خوشه با کمترین مربع خطا (SE)، استفاده از روش K-means می باشد (با محاسبه فاصله اقلیدسی). اگر نقاط اولیه در الگوریتم K-means خوب انتخاب نشده باشند الگوریتم در یک بهینه محلی خواهد افتاد و SW کوچکترین SW ممکن نخواهد بود. GKA شامل مجموعه ای از جواب های کد شده می باشد که جمعیت اولیه بر اساس انتخاب متغیرها به صورت تصادفی مشخص شده است و در نسل های بعدی بر اساس عملگرهای الگوریتم ژنتیک صورت می پذیرد و تولید نسل ها تا زمان رسیدن به شرایط خواسته شده ادامه پیدا می کند. عملگرهای ژنتیکی که در GKA استفاده شده اند عبارت اند از:

۱ - عملگر انتخاب

۲ - عملگر جهش بر اساس فاصله

۳ - عملگر K-means

در این بخش به نحوه کد کردن و عملگر عملگرهای ژنتیکی پرداخته می شود:

۱.۲.۵. کد گذاری

هدف، کدگذاری ماتریس W می باشد. برای این کار رشته ای به طول n (تعداد داده ها) در نظر گرفته می شود که اعداد داخل رشته نمایانگر خوشه ای می باشند که داده به آن تعلق دارد. که مقادیر داخل کروموزم $\{1, 2, 3, \dots, K\}$ می باشند. به این روش کدگذاری string-of-group-numbers encoding گفته می شود [۳].

۲,۲,۵. جمعیت اولیه

جمعیت اولیه $P(0)$ به صورت تصادفی مشخص می گردد. رشته کروموزوم از اعداد بین ۱ تا K به صورت تصادف با توزیع نرمال پر می شوند. برای اینکه احتمال دارد در فرآیند انتخاب تصادفی یک خوشه اصلاً انتخاب نشود برای جلوگیری از این امر یک عدد صحیح کوچکتر از n/K انتخاب شده و به این تعداد ژن یک گروه اختصاص داده می شود. این عمل برای K گروه تکرار شده و مابقی ژن ها به صورت دلخواه پر می شوند.

۳,۲,۵. انتخاب

عملگر انتخاب به صورت تصادفی و بر اساس توزیع زیر از جمعیت اقدام به انتخاب کروموزوم می نماید.

(۸)

که در آن $F(S_i)$ معرف مقدار تابع برازندگی کروموزوم S_i می باشد. جهت ارزیابی میزان خوب بودن کروموزوم های موجود، نیاز به تابع برازندگی می باشد که در اینجا مجموع واریانس درون خوشه ها به عنوان معیار خوب یا بد بودن کروموزوم ها در نظر گرفته شده است. هر چه $S(W)$ کمتر باشد به این معنی است که کروموزوم دارای برازندگی بیشتری می باشد. جهت تعریف تابع برازندگی بر اساس $S(W)$ روش های زیادی موجود می باشد که در این جا از روش زیر استفاده شده است:

$$f(\quad) \quad (9)$$

و

$$g \quad (10)$$

که در آن \bar{f} و σ معرف میانگین و انحراف معیار $f(S_W)$ در نسل فعلی می باشد و ثابت C عددی مابین ۱ و ۳ می باشد. در نهایت تابع برازندگی به شرح زیر خواهد بود:

(۱۱)

۴,۲,۵. جهش

عملگر جهش بر پایه فاصله معرفی شده است. در جهش، مقدار یک ژن با توجه به فاصله نقطه انتخابی از مراکز خوشه ها تغییر پیدا می کند. در واقع هر ژن معرف یک داده و مقدار آن معرف خوشه ای است که آن داده به آن اختصاص دارد، می باشد. عملگر جهش به گونه ای تعریف می گردد که تمایل دارد مقدار ژن را به گونه ای تغییر دهد که داده به خوشه ای که به مرکز آن نزدیک تر است تعلق گیرد.

جهت انجام این عمل بر روی آن $S_W(i)$ که معرف ژن انتخابی در کروموزوم S_W می باشد و با توجه به این که ژن i ام در ارتباط با داده x_i می باشد می توان فاصله ژن از مرکز خوشه J به صورت زیر تعریف نمود:

$$d \quad (12)$$

و سپس با احتمال زیر خوشه داده با همان مقدار ژن جایگزین می گردد.

$$P_j \quad (13)$$

که در آن c_m عددی بزرگتر از یک و $d_{\max} = \max_j \{d_j\}$ است. باید توجه داشت احتمال این که خوشه خالی در این عملگر به وجود آید، وجود دارد.

مثلاً اگر چهار خوشه داشته باشیم توسط عملگر جهش فقط خوشه های ۱ و ۲ انتخاب شده باشند و خوشه ۳ اصلاً انتخاب نشده باشد و با توجه به این که هر چه تعداد خوشه کمتر باشد SE طبیعتاً بیشتر خواهد بود. باید از خالی شدن خوشه ها جلوگیری به عمل آید. یک روش سریع برای بررسی نمودن اینکه احتمال تشکیل خوشه خالی وجود دارد این هست که آیا جامعه داده x_i از خوشه خودش $C_{SW}(i)$ بزرگتر از صفر هست یا نه. در

صورتی که $d_{sw}(i) = 0$ بدین منظور است که داده و مرکز خوشه یکی می باشد. بنابراین فقط در مواردی که فاصله بزرگتر از صفر است عملگر جهش می تواند بر روی داده اعمال گردد. کروموزمی که دارای K خوشه غیر خالی باشد کروموزم معتبر شناخته می شود و در غیر اینصورت کروموزم معتبر نیست.

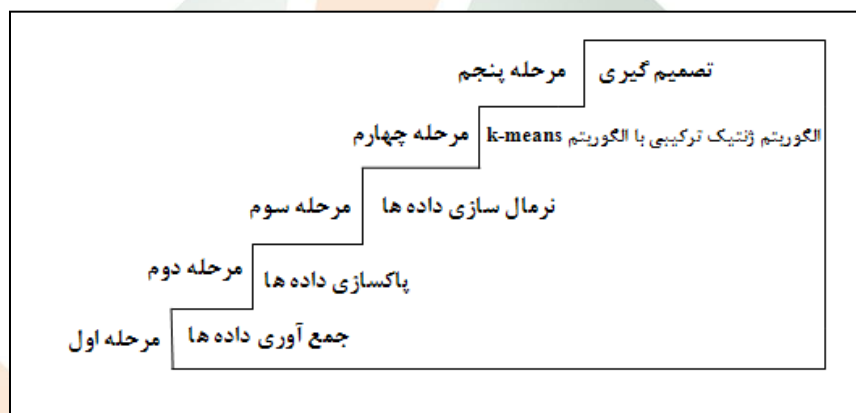
۵,۲,۵. عملگر دورگه شدن K-means

در الگوریتم ترکیبی به جای عمل دو رگه شدن، عملگر K-means قرار داده شده است جهت افزایش سرعت همگرایی از الگوریتم K-means استفاده شد که اگر sw کروموزم انتخابی باشد مراحل زیر توسط عملگر K-means بر آن اعمال می گردد.

- ۱ - مراکز خوشه با توجه با فرمول (۴) برای ماتریس W محاسبه می گردد.
 - ۲ - هر داده به نزدیکترین خوشه نسبت داده می شود و ماتریس \bar{W} تشکیل می گردد.
- احتمال دارد که حاصل این عملگر منجر به کروموزم نامعتبر گردد. در این حالت کروموزم نامعتبر به روش زیر به کروموزم معتبر تبدیل می گردد:
- ۱ - خوشه C که دارای بیشترین واریانس درون خوشه ای است انتخاب می گردد.
 - ۲ - به تعداد مورد نظر از دورترین داده ها به مرکز خوشه C جهت پر کردن خوشه خالی استفاده می گردد.

۵,۳,۵. مدل سازی

مراحل داده کاوی در این تحقیق در شکل ۲ نشان داده شده است.



شکل (۲) مراحل فرآیند داده کاوی در کارخانه ریسندگی الیاف کوتاه

برای انجام مدل سازی ابتدا تعداد ۱۰۰۰ داده که شامل ۲۵ نوع متغیر می باشد، را از آزمایشگاه کنترل کیفیت کارخانه مورد نظر مورد مطالعه قرار گرفت. متغیرها در جدول ۱ نشان داده شده است.

جدول (۱) پارامترهای مورد استفاده در فرآیند داده کاوی

شماره	نام متغیر	شماره	نام متغیر
۱	درصد ترکیب پنبه	۱۴	نمره نخ
۲	درصد ترکیب پلی استر	۱۵	Cv% نخ
۳	نمره فتیله کاردینگ	۱۶	نقاط نازک نخ
۴	نپ فتیله کاردینگ	۱۷	نقاط ضخیم نخ
۵	نمره فتیله چند لاکنی ۱	۱۸	نپ نخ
۶	Cv% فتیله چند لاکنی ۱	۱۹	IPI نخ
۷	U% فتیله چند لاکنی ۱	۲۰	تاب نخ
۸	نمره فتیله چند لاکنی ۲	۲۱	میانگین استحکام نخ
۹	Cv% فتیله چند لاکنی ۲	۲۲	میانگین ازدیاد طول نخ
۱۰	U% فتیله چند لاکنی ۲	۲۳	Cv% استحکام نخ
۱۱	نمره نیمچه نخ	۲۴	Cv% ازدیاد طول نخ
۱۲	Cv% نیمچه نخ	۲۵	میانگین RKM نخ
۱۳	U% نیمچه نخ		

داده های جمع آوری شده در برنامه اکسل مرتب گردید. سپس بر روی داده های خام متغیرها پردازش صورت گرفته و به صورت داده هایی که قابلیت تحلیلی بیشتری داشته باشند در آمدند. برای عمل پاک سازی و تمیز کردن داده ها چون بعضی از آزمایشات در روز یک دفعه و برخی دیگر سه دفعه انجام می شدند ، به همین علت مقادیر مربوط به تعدادی از ردیف ها خالی می ماند، برای رفع این مشکل آزمایشاتی که روزانه انجام می شد در دو ردیف بعد تکرار گردید. برخی از داده ها نیز به علت عدم ثبت، بجای آنها، میانگین داده قرار داده شد. برای ادامه ی کار داده هایی که درصد کمی از جامعه ی آماری را شامل می شدند حذف گردیدند. هنگامی که داده های مورد کاوش مشخص گردیدند، معمولاً به تبدیل های خاصی روی داده ها نیاز است. تبدیل های ساده همچون تبدیل داده ها به داده های نرمال تعریف شدند. برای نرمال سازی داده ها از رابطه ۱۴ استفاده شد.

(۱۴)

P_i

که در رابطه ۱۴:

P_{ij}^* : مقدار متغیر بعد از نرمال سازی

P_{ij} : مقدار متغیر قبل از نرمال سازی

\bar{P}_j : مقدار میانگین

S_j : واریانس Z ام متغیر از همه نمونه ها می باشد.

از متغیرهای نرمال شده به عنوان ورودی الگوریتم ژنتیک ترکیبی با الگوریتم k-means استفاده شد. اصولاً وارد کردن داده ها به صورت خام باعث کاهش سرعت و دقت الگوریتم انتخابی می شود.

در مرحله بعد، متغیر ها خوشه بندی شدند. در این تحقیق برای خوشه بندی گروه ها از تکنیک الگوریتم ژنتیک ترکیبی با الگوریتم k-means استفاده شد. برای نوشتن الگوریتم مورد نظر از نرم افزار Matlab استفاده شد. در این الگوریتم متغیرها بر اساس شباهت خود در خوشه های مختلفی قرار می گیرند.

پارامترهای تنظیم شده برای الگوریتم ژنتیک ترکیبی با الگوریتم k-means در جدول ۲ ارائه شده است.

جدول (۲) پارامترهای الگوریتم ژنتیک ترکیبی با الگوریتم k-means

تعداد نسل	تعداد جمعیت	احتمال جهش	ثابت C	ثابت C_m
۱۰۰	۵۰	۰.۰۵	۲	۱

برای تعیین بهترین دسته بندی، تعداد خوشه های مختلف مورد بررسی قرار گرفت. نتایج حاصل از الگوریتم ژنتیک ترکیبی با الگوریتم k-means در جدول ۳ نشان داده شده است. جهت ارزیابی خوشه ها از معیار دیویس - بولدین (DB) استفاده شد. این معیار از شباهت بین دو خوشه استفاده می کند که بر اساس پراکندگی بین یک خوشه و عدم شباهت بین دو خوشه تعریف می شود. این شاخص میانگین شباهت بین هر خوشه با شبیه ترین خوشه به آن را محاسبه می کند. می توان دریافت که هر چه مقدار این شاخص کمتر باشد، خوشه های بهتری تولید شده است [۱۶].

جدول (۳) تعداد داده ها در هر خوشه و معیار دیویس - بولدین برای هر خوشه

تعداد خوشه (K)	تعداد داده ها در هر خوشه	معیار دیویس - بولدین (DB)
۳	۴۲۲ ۲۶۶ ۳۱۲	۲.۱۵
۴	۱۸۶ ۳۱۲ ۲۶۶ ۲۳۶	۲.۱۸۳
۵	۱۳۰ ۱۸۶ ۲۳۶ ۳۱۲ ۱۳۶	۲.۰۲۹
۶	۱۳۰ ۱۸۶ ۲۱۹ ۲۰۸ ۱۲۱ ۱۳۶	۲.۱۵۴
۷	۱۹۱ ۵۸ ۱۸۳ ۱۲۹ ۱۳۳ ۲۰۶ ۱۰۰	۲.۱۷۹
۸	۵۶ ۱۱۰ ۱۳۶ ۱۴۶ ۱۳۰ ۱۷۳ ۱۷۳ ۷۶	۲.۱۳۵
۹	۱۲۴ ۱۷۳ ۷۵ ۱۳۵ ۹۵ ۱۰۸ ۱۰۳ ۵۸ ۱۲۹	۲.۱۶۲

۶. نتایج و بحث

هدف از این تحقیق استفاده از روش خوشه بندی ترکیبی کی - مینز و الگوریتم ژنتیک برای کنترل کیفیت نخ ریسیده شده در صنعت نساجی می باشد. این روش یکی از تکنیک های کشف الگوها از پایگاه داده های مشخص در داده کاوی است. ایده اصلی در این الگوریتم بهره مندی از مزایای هر دو الگوریتم k-means و ژنتیک می باشد و در واقع این ترکیب، سرعت همگرایی را بهبود می بخشد.

برای بررسی و تایید روش انتخاب شده از شاخص دیویس - بولدین (DB) استفاده شد که طبق بررسی های انجام شده و مطابق با جدول ۳ بهترین حالت برای داده ها خوشه بندی با $K=5$ انتخاب شد. زیرا دارای کمترین مقدار DB است. با در نظر گرفتن $K=5$ به عنوان بهترین تعداد خوشه، داده های متعلق به هر گروه مشخص شد.

مراکز گروه ها برای خوشه $K=5$ در جدول ۴ نشان داده شده است.

جدول (۴) مراکز گروه ها برای بهترین خوشه ($K=5$)

شماره گروه					
شماره متغیر	۱	۲	۳	۴	۵
۱	۶۵,۷۵	۶۵,۴۱	۶۵,۴۱	۶۵,۰۸	۶۵,۷۵
۲	۳۴,۹۲	۳۵,۲۵	۳۵,۲۵	۳۵,۵۹	۳۴,۹۲
۳	۵,۵۴	۵,۵۴	۵,۲۷	۵,۴۳	۵,۲۷
۴	۲,۱۵	۲,۱۴	۲,۱۵	۲,۱۵	۲,۱۲
۵	۱۷,۳۹	۱۷,۵۸	۱۷,۴۸	۱۷,۹۰	۱۷,۶۲
۶	۳,۲۸	۳,۲۷	۳,۲۸	۳,۲۷	۳,۲۸
۷	۲,۵۶	۲,۵۶	۲,۵۷	۲,۵۷	۲,۵۷
۸	۴,۶۱	۴,۶۱	۴,۶۱	۴,۶۱	۴,۶۱
۹	۲,۸۰	۲,۸۰	۲,۸۱	۲,۸۰	۲,۸۰
۱۰	۲,۱۹	۲,۱۹	۲,۱۹	۲,۱۸	۲,۱۹
۱۱	۱,۲۰	۱,۲۰	۱,۲۰	۱,۲۰	۱,۲۰
۱۲	۵,۳۴	۵,۳۴	۵,۳۵	۵,۳۵	۵,۳۴
۱۳	۴,۲۹	۴,۳۰	۴,۳۰	۴,۳۰	۴,۳۰
۱۴	۳۲,۸۵	۳۲,۸۶	۳۲,۸۶	۳۲,۸۷	۳۲,۸۸
۱۵	۱۴,۸۲۶	۱۴,۸۳۴	۱۴,۸۳۷	۱۴,۸۳	۱۴,۸۲۸
۱۶	۸,۲۳	۸,۱۸	۸,۲۵	۸,۲۰	۸,۱۴
۱۷	۳۶,۳۵	۳۶,۳۵	۳۶,۳۵	۳۶,۷۳	۳۶,۵۴
۱۸	۵۰,۲۳	۵۰,۳۱	۵۰,۲۳	۵۰,۳۶	۵۰,۲۵
۱۹	۹۴,۶۴	۹۴,۸۴	۹۴,۷۹	۹۵,۰۹	۹۴,۷۴
۲۰	۲۱,۶۶	۲۱,۶۸	۲۱,۶۵	۲۱,۶۶	۲۱,۶۸
۲۱	۴۴۶,۷۸	۴۴۶,۴۸	۴۴۶,۲۴	۴۴۶,۷۲	۴۴۷,۰۸
۲۲	۱۱,۲۳	۱۱,۱۹	۱۱,۲۱	۱۱,۲۲	۱۱,۲۲
۲۳	۹,۵۴	۹,۵۲	۹,۵۳	۹,۴۷	۹,۵۲
۲۴	۱۲,۰۸	۱۲,۰۵	۱۲,۱۰	۱۲,۲۱	۱۲,۲۰
۲۵	۲۴,۲۱	۲۴,۲۰	۲۴,۱۸	۲۴,۲۱	۲۴,۲۳

روش پیشنهادی در این مقاله که خوشه بندی ترکیبی کی- مینز و الگوریتم ژنتیک می باشد، بر اساس ۲۵ متغیر مشخص شده، ۱۰۰۰ داده را به صورت کلی و موردی ارزیابی کرد. با استفاده از این روش پیشنهادی می توان یک سری داده های جدید را مورد ارزیابی و بررسی قرار داد. به عنوان مثال با داشتن طول و ظرافت مواد اولیه جدید می توان فهمید که متعلق به کدام گروه است و خصوصیات محصول نهایی را پیش بینی کرد.

۷. نتیجه گیری

در این تحقیق با توجه به ویژگی های الگوریتم ژنتیک، از یک الگوریتم ژنتیک ترکیبی با الگوریتم **k-means** جهت دسته بندی داده ها استفاده شد. جهت بررسی روش پیشنهادی پس از انتخاب مجموعه داده ها با استفاده از برنامه متلب، الگوریتم مورد نظر نوشته شد. داده های مورد پردازش مربوط به آزمایشگاه کنترل کیفیت یک کارخانه ریسندگی الیاف کوتاه شامل پارامترهای مواد اولیه، پارامترهای فرآیند و پارامترهای کیفی نخ تولیدی بود. دقت الگوریتم از پارامترهای مهم تحلیل روش پیشنهادی بود. نتایج بدست آمده نشان دادند که استفاده از الگوریتم های ژنتیک همراه با تکنیک های داده کاوی می تواند کارایی مناسبی را به همراه داشته باشد. نتایج نشان داد که خوشه بندی با استفاده از روش ترکیبی کی - مینز و الگوریتم ژنتیک به عنوان یک تکنیک داده کاوی، یک ابزار موثر برای پیش پردازش کیفیت محصول نهایی قبل از ورود مواد اولیه به خط تولید یک کارخانه و شروع عملیات تولید می باشد.

۸. تقدیر و تشکر

از کارخانه یزدباف ایران که ما را در تهیه این تحقیق یاری نموده اند، صمیمانه تشکر می نمایم.

۹. مراجع

- ANDERSON, M. E., "APPLICABILITY OF DATA MINING IN YARN MANUFACTURING", Msc_Thesis, North Carolina State University , ۲۰۰۶ [۱]
- Bagherzadeh, R., Latifi, M. and Faramarzi A.R., "Employing a Three-Stage Data Mining Procedure to Develop Sizing System" World Applied Sciences Journal, Vol. ۸, p.p.۹۲۳-۹۲۹, ۲۰۱۰ [۲]
- D. Goldberg., Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley, ۱۹۸۹ [۳]
- Daley, C. M., "Application of Data Mining Tools for Exploring Data: Yarn Quality Case Study". Msc_Thesis , North Carolina State University, ۲۰۰۸ [۴]
- Eiben, A.E. and Smit, J.E., "Introduction to Evolutionary Computing", ۱st edition, Springer, ۲۰۰۸ [۵]
- Glimcher, L., Jin, R. and Agrawal, G., "Middleware for data mining applications on clusters and grids", Journal of Parallel and Distributed Computing, vol. ۱۸, p.p.۳۷-۵۳, ۲۰۰۸ [۶]
- Hodge, G., Oxenham, W., Jones, M., Karpe, Yatin., Uncu. S. and Schertel, S., "INFORMATION ENGINEERING: TEXTILE INDUSTRY'S VALUE-ADDING KEY TO EFFECTIVE DECISION-MAKING" Institute of Textile Technology, p.p.۱-۱۰, ۲۰۰۱ [۷]
- Hammerly, G. and Elkan, C., "Alternatives to the k-means algorithm that find better clusterings", Proceedings of the ۱۱th international conference on information and knowledge management, p.p.۶۰۰-۶۰۷, ۲۰۰۲ [۸]
- Han, J. and Kamber, M., Data Mining Concepts and Techniques, ۲st edition, Morgan Kaufmann, ۲۰۰۶ [۹]
- Hamilton, B. John., "Process and Product Data Management for Staple Yarn Manufacturing", Msc_Thesis, North Carolina State University , ۲۰۱۰ [۱۰]
- Jones D. R. and Beltramo, M. A., "Solving partitioning problems with genetic algorithms" , in Proc. 4th Int. Conf. Genetic Algorithms. San Mateo, CA: Morgan Kaufman, ۱۹۹۱ [۱۱]
- Jain, A. K., Murty, M. N. and Flynn, P.J., "Data clustering: a review", ACM Computing Surveys, Vol. ۳۱, No. ۳, p.p.۲۶۴-۳۲۳, ۱۹۹۹ [۱۲]
- Jeyasingh, M.M., "Mining the Shirt Sizes for Indian Men by Clustered Classification" I.J. Information Technology and Computer Science, Vol. ۶, p.p.۱۲-۱۷, ۲۰۱۲ [۱۳]
- Jeyasingh, M.M., Appavoo, K. and Sakthivel, P., "Data Mining for Prediction of Clothing Insulation" International Journal of Modern Engineering Research (IJMER), Vol. ۲, p.p.۱-۵, ۲۰۱۲ [۱۴]
- MacQueen J.B., "Some methods for classification and analysis of multivariate observations", In Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. ۱, p.p.۲۸۱-۲۹۷, ۱۹۶۷ [۱۵]
- Mitchell. Tom., "Machine Learning", ۱st edition, McGraw-Hill, ۱۹۹۷ [۱۶]
- Schertel L., "DATA MINING AND ITS POTENTIAL USE IN TEXTILES: A Spinning Mill", Phd_Thesis, North Carolina State University, ۲۰۰۲ [۱۷]
- SHAHRABI, J. and Salehi, M., "DEVELOPMENT OF A NEW SIZING SYSTEM BASED ON DATA MINING APPROACHES" ۷th International Conference – TEXSCI, ۲۰۱۰ [۱۸]
- Tyagi, S.K., and Sharma, B.K. "Data Mining Tools and Techniques to Manage the Textile Quality Control Data for Strategic Decision Making" International Journal of Computer Applications, Vol. ۱۳, No. ۴, p.p.۲۶-۲۹, ۲۰۱۱ [۱۹]
- Shengchun, D.; Zengyou, He. and Xiaofei, Xu.; "G-ANMI: A mutual information based genetic clustering algorithm for categorical data", Knowledge-Based Systems, Vol. ۲۳, p.p.۱۴۴-۱۴۹, ۲۰۱۰ [۲۰]
- TYAGI, S.K., SHARMA, B.K. and SAXENA, S. "BOILER EFFICIENCY EVALUATION AND MONITORING THROUGH DATA MINING TECHNIQUES IN TEXTILE INDUSTRY" International Journal on Computer Science and Engineering (IJCSE), Vol. ۳, No. ۴, p.p.۱۷۲۷-۱۷۳۳, ۲۰۱۱ [۲۱]

Yang, F., Sun, T. and Zhang, C., "An efficient hybrid data clustering method based on K-harmonic means and Particle Swarm Optimization", Expert Systems with Applications, Vol. ۳۶, p.p.۸۴۷-۸۵۲, ۲۰۰۹ [۲۲]



کنفرانس داده کاوی ایران