اولین کنفرانس بین المللی جهش علوم مدیریت اقتصاد و حسابداری
1st international conference on the mutation of management science, economics and accounting
www.mmea.ir
موسسه آموزش عالی ادیب مازندران
ISC کد اختصاصی: ۰۰۲۰۱-۹۱۰۴۶

کنترل خط تولید با داده کاوی نمودارهای کنترل کیفیت با استفاده از الگوریتم ماشین بردار پشتیبان (SVM)

شهرام کاظم نژاد واقفی

عضو هیئت علمی مربی گروه مدیریت دانشکده علوم انسانی دانشگاه آزاد اسلامی مرند

نویسنده مسئول : شهرام کاظم نژاد واقفی

تلفن :۰۹۱۴۱۱۸۵۳۶۳

ادرس الکترونیک: shkvaghefi@gmail.com

چکیده

داده کاوی، علم استخراج اطلاعات و دانش و کشف الگوهای پنهان از یک پایگاه داده بسیار بزرگ و پیچیده می باشد. فرابینی یکی از مباحث مهم در صنایع برای بالا بردن سطح کیفی محصولات و کاهش هزینه ها به شمار می آید. اگرچه نمودار کنترل یکی از ابزار کنترل کیفیت است اما این نمودار معمولا نمیتواند زمان واقعی تغییر در نمودار را نمایان سازد. همیشه از فرآیندهای مختلف داده های بسیاری دردسترس است که با کاوش این داده ها با استفاده از تکنیکهای تجزیه و تحلیل آماری می توان کیفیت را بهبود و سودآوری را افزایش داد. در نمودارهای مختلف کنترل فرآیند آماری، بواسطه رخداد نوسانات طبیعی در فرآیند تولید، تشخیص صحیح و سریع الگوهای غیرطبیعی و انحرافات معنی دار در بازۀ زمانی کوتاه با چالشهایی مواجه است. از اینرو ارائه مدلهای مبتنی بر کلاسبندی کننده ماشین بردار پشتیبان با هدف شناسایی و تفسیر الگوهای نمودارهای کنترل فرآیند منطقی میباشد. پژوهش جاری مدلی را جهت تشخیص و بررسی الگوهای پایه و همزمان پیشنهاد مینماید .

کلمات کلیدی: داده کاوی؛ کنترل کیفیت؛ کلاسبندی کننده؛ ماشین بردار پشتیبان؛ الگوهای پایه .

www.SID.ir

۱. مقدمه

داده کاوی، علم استخراج اطلاعات و دانش و کشف الگوهای پنهان از یک پایگاه داده بسیار بزرگ و پیچیده می باشد. داده کاوی یک متدولوژی بسیار قوی و با پتانسیل بالا است که به سازمانها کمک می کند تا بر روی مهمترینِ اطلاعات از مخزن داده ها تمرکز نمایند و وقت خود را به منظور تجزیه تحلیل داده هایی که از اهمیت پایین تری برخوردارند، هدر ندهند.داده کاوی کمک می کند تا سازمانها با کاوش بر روی داده های سیستم خود، الگوها، روندها و رفتارهای آینده را پیش بینی کرده و بهتر تصمیم بگیرند. این روش با استفاده از تحلیل وقایع گذشته، یک تحلیل پیش بینانه ارائه می دهد و به سوالاتی جواب می دهد که پاسخ به آنها در گذشته ممکن نبوده و یا به زمان زیادی، نیاز داشته است.از داده کاوی می توان برای انجام کارهایی مثل کلاس بندی، پیش بینی، تخمین و خوشه بندی داده ها استفاده کرد. برای اینکار، تکنینکهای بسیار زیادی توسعه یافته اند که از مهمترین آنها می توان به الگوریتمهای خوشه بندی، شبکه های عصبی، الگوریتم ژنتیک، نزدیکترین همسایگی و درخت تصمیم گیری اشاره نمود. [2]کیفیت محصول تولیدی در بازارهای رقابتی امروز نقش انکار ناپذیری در ایجاد مزیت رقابتی و پیشی گرفتن از رقبا دارد، برای مدیریت و کنترل کیفیت دستیابی به فرایندهای تولیدی پایدار و کارا اجتناب ناپذیر است. از این سو کنترل کیفیت آماری ، مجموعه ای قوی از ابزارهای آماری مفید برای کنترل تغییرات فرایند و بهبود کارایی فرایند است که نقش تعیین کننده ای در پیشبرد اهداف کنترل کیفیت دارد. [3]نمودارهای کنترل به عنوان بخش مهمی از این ابزارها به طور وسیعی در روشهای گوناگون کنترل کیفیت مورد استفاده قرار میگیرند، زمانی که یک فرایند از محدوده کنترل خارج میشود، پس از مدتی منحنی کنترل این نمودارها از حدود کنترل تجاوز کرده و به این وسیله به دلیل خروج فرایند از کنترل اخطار میدهد. البته این اخطار پس از یک تاخیر زمانی، در فرایند ظاهر میشود. پس از اخطار نمودار کنترل مهندسان کیفیت با جستجوی دلیل خطای به وجود آمده به رفع و اصلاح آن پرداخته و فرآیند را دوباره به حالت کنترل باز میگردانند. پر واضح است چنانچه زمان وقوع تغییر فرایند مشخص باشد، یافتن علت خطا بسیار سریعتر و آسانتر انجام خواهد گرفت و به همین دلیل فرایند انجام اقدامات اصلاحی سرعت خواهد یافت. پس با شناخت نقش موثر شناسایی زمان تغییر فرایند در بهبود کیفیت، مطالعات گسترده ای برای یافتن زمان واقعی تغییر در فرایند تحت عنوان نقطه تغییر فرایند انجام گرفته است. هر یک از مطالعات صورت گرفته به سهم خود قدم موثری در تشخیص درست زمان تغییر برداشته اند. عده ای از پژوهشگران با استفاده از روشهایی به براورد نقطه تغییر پرداخته اند. ساموئل و همکاران[12] با استفاده از روش تخمین زننده بیشینه درست نمایی زمان تغییر را در میاگین فرایندهای نرمال، هنگام اخطار نمودار کنترل را براورد نوده اند، آنها همچنین کارایی براورد خود را برای نمودار کنترل $\bar{x}$ شوهارت مورد بررسی قرار دادند و نتیجه گرفتند که مدل آنها در مقایسه با زمان اخطار نمودار کنترل، براورد بهتری از نقطه تغییر ارائه میدهد[13] و حتی در شرایطی که میزان تغییر به وجد آمده در میانگین کم و ARL نمودار کنترل بزرگ است، براورد آنها مطلوب تر است. هاوکینز [18]در مقاله خود به این موضوع اشاره میکند که روشهای رایج برای شناسایی تغییر فرایند شوهارت مانند CUSUM و EWMA نیازمند داشتن اطلاعاتی از پارامترهای فرایند در حالت کنترل است. در حالتی که این اطلاعات به ندرت به صورت دقیق قابل دستیابی هستند، که آنها برای آشکارسازی و تشخیص تغییرات پله ای براساس اطلاعات ناقص از فرایند، فرنول نقطه تغییر با پارامتر نامعلوم را ارئه کردند. ساموئل و همکاران در مطالعه ای دیگر نقطه تغییر را در فرایندی که در آن تغییر پله ای در واریانس توزیع نرمال رخ میدهد را به روش MLE براورد کردند. در این مطالعه نیز زمان نقطه براورد شده با زمان اخطار نمودار کنترل S شوهارت مقایسه شده است. [17] پیگناتیلو و سیمسون ابزاری برای نظارت در فرایند پیشنهاد کردند که نه تنها توانایی آشکارسازی تغییرات فرایند را صرف نظر از مقدار تغییرات دارد بلکه آمارهای مفیدی برای نقطه تغییر ارائه میکند، در این نمودار روش نسبت راستی نمایی برای توسعه نمودارهای کنترل برای تغییرات دائمی پله ای در میانگین فرایندهای نرمال به کار رفته است. [14]و[15]

۲. کنترل کیفیت آماری

این علم، به سه بخش اصلی شامل کنترل فرایند آماری[1](SPC)، طراحی آزمایشها و نمونه گیری به منظور پذیرش یا رد محموله، تقسیم می شود.

- در نمونه گیری به منظور پذیرش محموله،که یک روش بازرسی است، با در نظر گرفتن تعادل بین ریسک تولید کننده و مصرف کننده، و با توجه به هزینه های نمونه گیری، طرحی مناسب انتخاب شده و بر اساس آن نسبت به پذیرش یا رد یک انباشته اقدام می شود.

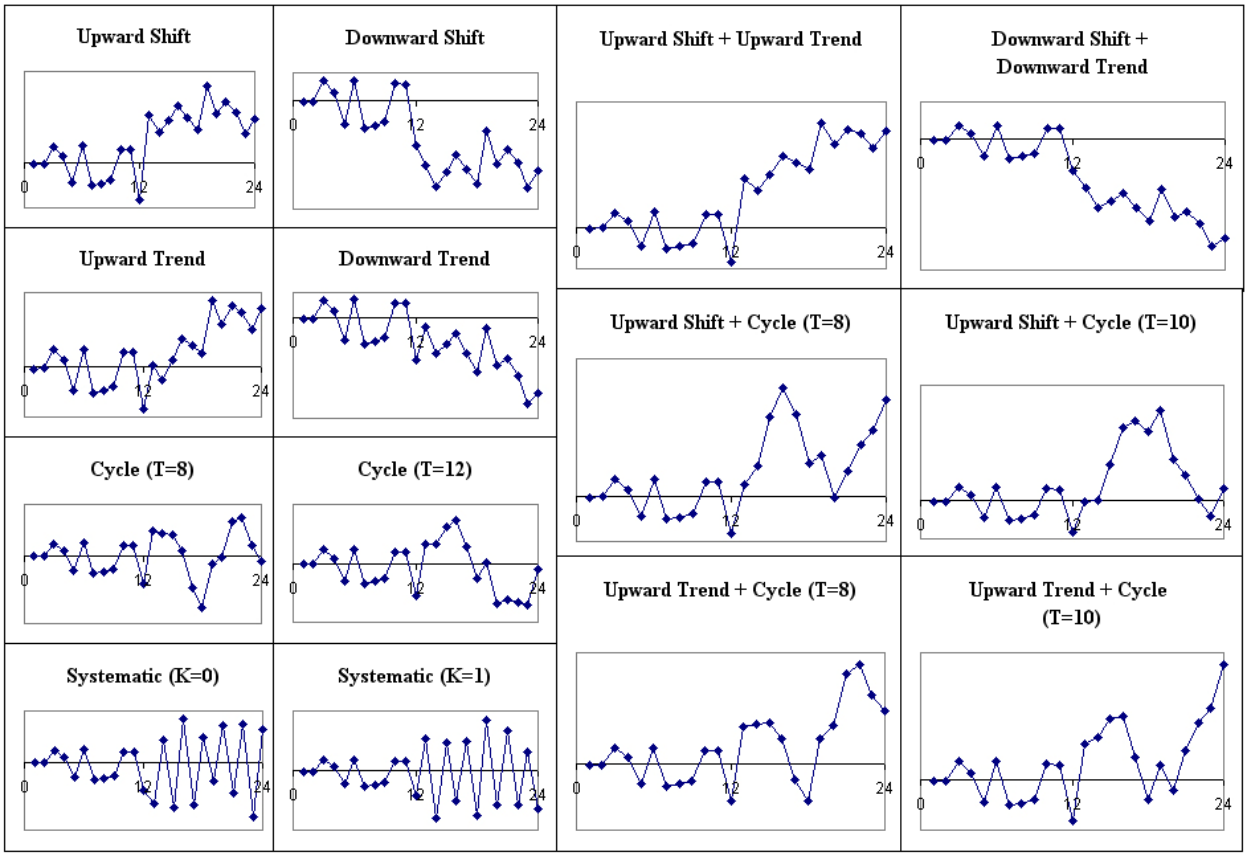
[1] statistic process control

www.SID.ir

- طراحیِ آزمایش، یک روش فعال می باشد، که با تغییر آگاهانه و هدفمند در سطوح فاکتورها، اثر آنها بر روی متغیرِ پاسخ، اندازه گیری و بررسی شده و سطوح بهینه فاکتورهای تاثیر گذار،تعیین می شود.
- کنترل فرایند آماری، یک تکنیک آماری است که برای کاهش پراکندگی و در نتیجه، بهبود کیفیت به کار می رود.از ابزارهای این روش می توان به هیستوگرام، برگه کنترل،نمودار پارتو، نمودار علت و معلول،نمودار پراکندگی، نمودار تمرکز نقصها و نمودار های کنترلی اشاره نمود. مهمترین و قدرتمندترین ابزار برای کنترل آماریِ یک فرایند، نمودارهای کنترلی می باشند که به منظور کنترل میزان تغییرات در یک مشخصه کیفی یا چندین مشخصه کیفی، کاربرد دارند. مشخصه های کیفی در دو قالب کلیِ قابل اندازه گیری و وصفی تقسیم بندی می شوند که هریک، نمودارهای کنترلیِ مختص خود را دارا می باشند[4].

در واقع بروز رفتارهای غیرتصادفی وجود اختلالات غیرقابل قبول را در فرآیند تولید هشدار میدهد و چون نمودارهای کنترل سنتی صرفا روی اطلاعات آخرین نمونه تمرکز داشته و توجهی به اطلاعات مشترک بدست آمده از نمونه‌های متوالی ندارند، تجزیه و تحلیل صحیح و سریع الگوهای رفتاری کاملا ضروری خواهد بود. بر این اساس تعاریف مشخصی از الگوهای رفتاری نمودارهای کنترل ارائه شده است.

در این میان الگوهای Shift ، Trend ، Cycle و Systematic به عنوان الگوهای پایه معرفی می‌شوند زیرا عموما ریشه‌های فرایندی دارند و بطور معمول در اکثر نمودارهای کنترل مشخصه های کیفی ظاهر می‌شوند. الگوهای پایه ای میتوانند بصورت منفرد یا ترکیبی نمایان گردند .[5]

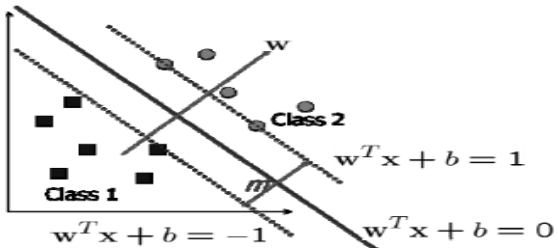


شکل ۱. الگوهای Shift ، Trend ، Cycle و Systematic

۳. ماشین بردار پشتیبان (SVM)

www.SID.ir

SVM دسته بندی کننده ای است که جزو شاخه روش هسته[۲]در یادگیری ماشین محسوب می شود. SVM در سال۱۹۹۸ توسط وپنیک[۳]معرفی گردیده است. [16]شهرت SVM به خاطر موفقیت آن در تشخیص حروف دست نویس است که با شبکه های عصبی به دقت تنظیم شده برابری می کند (۱/۱٪ خطا). هدف این دسته الگوریتم ها تشخیص و متمایز کردن الگوهای پیچیده در داده ها است (از طریق کلاسترینگ، دسته بندی، رنکینگ، پاکسازی و ...). در SVM با فرض اینکه دته ها به صورت خطی جداپذیر باشند، ابر صفحه هائی با حداکثر حاشیه[۴]را به دست می آورد که دسته ها را جدا کنند. در شکل ۱ مرز تصمیم گیری بهینه و بردار های پشتیبان نشان داده شده است. مشخص است که در بین نمونه های آموزشی همواره زیر مجموعه ای وجود دارد که بتواند بهترین مرز تصمیم گیری را تعریف کند. برای مثال نقاط آموزشی دارای کمترین فاصله از مر زهای تصمیم گیری می توانند به عنوان زیر مجموعه ای برای تعریف مرزهای تصمیم گیری در نظر گرفته شوند. یکی از مزایای بسیار مهم ماشین بردار پشتیبان بهبود بازده این طبقه بندی کننده با افزایش بعد داده ها است. در واقع این طبقه بندی کننده یک طبقه بندی کننده آماری نیست و مستقیماً با استفاده از نقاط آموزشی مرز تصمیم گیری بین دو کلاس را به دست می آورد. توجه داشته باشید که منظور از غیر آماری بودن ماشین بردار پشتیبان این است که در این طبقه بندی کننده از ممان های آماری استفاده نمی شود و بنابراین خطای حاصل از تخمین ممانهای آماری که باعث پدیده هیوز می شود در اینجا وجود ندارد. از طرفی هر چه بعد داده ها بیشتر شود فاصله کلاس ها از یکدیگر بیشتر شده و امکان جدا سازی آنها با تعداد محدودی نقاط آموزشی افزایش می یابد. ایده SVM برای جداسازی دسته ها این است که دو صفحه مرزی بسازد: دو صفحه مرزی موازی با صفحه دسته بندی رسم کرده و آندو را آنقدر از هم دور می کنیم که به داده ها برخورد کنند. صفحه دسته بندی که بیشترین فاصله را از صفحات مرزی داشته باشد، بهترین جدا کننده خواهد بود. نزدیک ترین داده های آموزشی به ابر صفحه های جداکننده بردار پشتیبان نامیده می شوند. در صورت استفاده مناسب از SVM این الگوریتم قدرت تعمیم خوبی خواهد داشت، زیرا علیرغم داشتن ابعاد زیاد از بیشبرازش[۵]پرهیز می کند. [6]این خاصیت ناشی از بهینه سازی این الگوریتم است که در فشرده سازی اطلاعات استفاده می شود. در واقع به جای داده های آموزشی از بردارهای پشتیبان استفاده می کند.



شکل ۲. ایده SVM برای جداسازی دسته ها

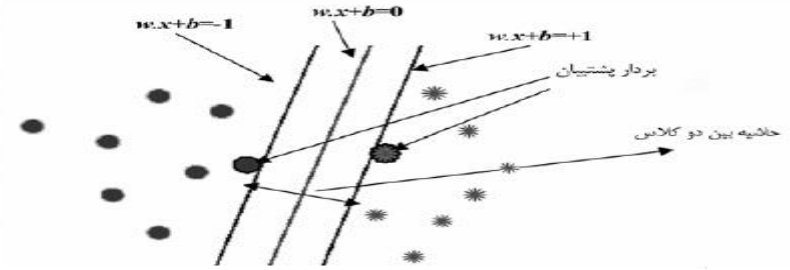
[۲] Kernel Methods

[۳] Vapnik

[۴] maximum margin

[۵] Overfitting

www.SID.ir



شکل ۳.مرز خطی بهینه برای حالتی که دو کلاس کاملاً از یکدیگر جدا هستند

در این روش با استفاده از تمامی باندها و یک الگوریتم بهینه سازی، نمونه هایی که مرزهای کلاس ها را تشکیل می دهند به دست می آورند. این نمونه ها را بردارهای پشتیبان گویند. تعدادی از نقاط آموزشی که کمترین فاصله تا مرز تصمیم گیری را دارند می توانند به عنوان زیر مجموعه ای برای تعریف مرزهای تصمیم گیری و به عنوان بردار پشتیبان در نظر گرفته می شود.[7]

در شکل (۳) دو کلاس و بردار های پشتیبان مربوط به آنها نشان داده شده است. فرض کنید داده ها از دو کلاس تشکیل شده و کلاس ها در مجموع دارای $\mathbf{x}_i$ *نقطه* آموزشی باشند که x_i یک بردار است. این دو کلاس با $y_i = \pm 1$ برچسب زده می شوند. برای محاسبه مرز تصمیم گیری دو کلاس کاملاً جدا از هم، از روش حاشیه بهینه استفاده می شود. در این روش مرز خطی بین دو کلاس به گونه ای محاسبه می شود که:

- تمام نمونه های کلاس +1 در یک طرف مرز و تمام نمونه های کلاس -1 در طرف دیگر مرز واقع شوند.
- مرز تصمیم گیری به گونه ای باشد که فاصله نزدیکترین نمونه های آموزشی هر دو کلاس از یکدیگر در راستای عمود بر مرز تصمیم گیری تا جایی که ممکن است حداکثر شود.

یک مرز تصمیم گیری خطی را در حالت کلی می توان به صورت زیر نوشت:

$$\mathbf{w}.\mathbf{x} + b = 0 \quad (1)$$

$\mathbf{x}$ یک نقطه روی مرز تصمیم گیری و $\mathbf{w}$ یک بردار n بعدی عمود بر مرز تصمیم گیری است. $b / \mathbf{w}$ فاصله مبدأ تا مرز تصمیم گیری و $\mathbf{w}.\mathbf{x}$ بیانگر ضرب داخلی دو بردار $\mathbf{x}$ و $\mathbf{w}$ است. از آنجا که با ضرب یک ثابت در دو طرف (1) باز هم تساوی برقرار خواهد بود. برای تعریف یکتای مقدار b و $\mathbf{w}$ شرایط زیر زیر بر روی آنها اعمال می شود.

اگر x_i باشد پشتیبان بردار یک $\Rightarrow y_i(wx_i + b) = 1$ (2)

اگر x_i نباشد پشتیبان بردار یک $\Rightarrow y_i(wx_i + b) > 1$ (3)

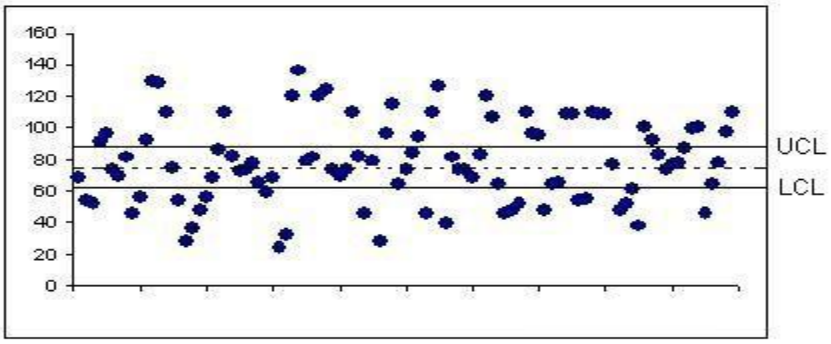
اولین مرحله برای محاسبه مرز تصمیم گیری بهینه، پیدا کردن نزدیکترین نمونه های آموزشی دو کلاس است. در مرحله بعد فاصله آن نقاط از هم در راستای عمود بر مرز هایی که دو کلاس را به طور کامل جدا می کنند محاسبه می شود. مرز تصمیم گیری بهینه، مرزی است که حداکثر حاشیه را داشته باشد. مرز تصمیم گیری بهینه با حل مسئله بهینه سازی زیر محاسبه می شود.

$$\max_{w,b} \min_{i=1,\dots,L} \left[y_i \frac{(w.x_i + b)}{|w|} \right] \quad (4)$$

www.SID.ir

۴. استفاده از ابزارهای داده کاوی برای حل مسائل کیفیت

استفاده از داده کاوی و ابزارهای آن باعث بهبود نحوه تصمیم گیری و در نتیجه بهبود عملکرد سازمان خواهد شد . همچنین همیشه از فرآیندهای مختلف داده های بسیاری دردسترس است که با کاوش این داده ها با استفاده از تکنیکهای تجزیه و تحلیل آماری می توان کیفیت را بهبود و سود آوری را افزایش داد . تفکر بسیاری از مهندسین کیفیت بر این اساس استوار بوده است که داده کاوی به عنوان یک علم نوپا می تواند در کنار ابزارهای سنتی کنترل کیفیت آماری استفاده و سازمان را در جهت تعالی سوق دهد. مهندسین کیفیت بر این باورند که داده کاوی و کنترل کیفیت آماری نه تنها مغایر یکدیگر نیستند بلکه یکدیگر را تکمیل می کنند. هر کدام از آنها دارای نقاط ضعفی می باشند که بوسیله دیگری این نقاط ضعف جبران می شود. به عنوان مثال ابزارهای کنترل کیفیت آماری برای تحلیل حجم کوچکی از داده ها استفاده می‌شوند. در صورت وجود حجم بزرگی از داده ها می توان از تکنیک های داده کاوی برای حل مشکلات کیفی استفاده نمود. بر عکس این موضوع نیز صادق است یعنی بسیاری از ابزارهای داده کاوی برای تحلیل حجم بزرگی از داده ها طراحی شده اند و برای تحلیل حجم کوچک از داده ها نمی توان از آنها استفاده کرد یا در صورت استفاده از آنها عملکرد ضعیفی از خود نشان می دهند.[8] به دلیل اینکه ابزارهای داده کاوی به گونه طراحی شده اند که از درصدی از داده ها برای ساخت مدل و از درصدی دیگر برای تست مدل استفاده می شود لذا وجود تعداد قابل توجهی داده برای عملکرد بهینه آن امری اجتناب ناپذیر است. هاوکینز [18] همچنین روشی با یک نمودار کنترل برای آشکارسازی تغییر در واریانس یک فرایند در شرایطی که مقدار اسمی واریانس نامعلوم باشد پیشنهاد کردند. غضنفری و همکاران[1] با استفاده از روش خوشه بندی به براورد نقطه تغییر میانگین در نمودار کنترل $\bar{X}$ شوهارت پرداختند. آنها با ارائه سه شاخص خوشه بندی نتایج خود را برای فازهای یک و دو و برای توزیعهای نرمال و غیر نرمال با روشهای پیشنهادی ساموئل و هاوکینز و همکاران مقایسه کردند و نتیجه گرفتند نه تنها روش پیشنهادی آنها بسیاری از محدودیت های دو روش اخیر را ندارد بلکه نتایجی مشابه و حتی در برخی از موارد بهتر از دو روش اخیر هم به دنبال دارد. روش خوشه بندی ارائه شده از طریق غضنفری و همکاران جواب های قابل قبولی را برای براورد نقطه تغییر ارائه داده است. با این حال روش خوشه بندی استفاده شده این مشکل را نیز دارد که جواب به دست آمده ممکن است در دام بهینه محلی قرار گرفته و الگوریتم ارائه شده قادر به بهبود جواب نداشته باشد، برای رفع این مشکل از منطق جستجوی ممنوعه استفاده کرده و الگوریتم جدید را طراحی و اجرا نموده که با درصد بالایی جوابهای بدست آمده از روش غضنفری و همکاران را بهبود بخشیده است. استفاده از داده کاوی به منظور بهبود کاربرد نمودارهای کنترلی را با یک مثال توضیح می دهیم. فرض کنید ۴ پارامتر در یک مشخصه کیفیِ مرتبط با یک قطعه ، تاثیر گذار باشند و هدف، بهبود کیفیت آن مشخصه مورد نظر باشد. با استفاده از اطلاعات موجود که از حجم بالایی برخوردار می باشد، نمودارهای کنترلی X/R مربوط به مشخصه یاد شده طبق شکل زیر، حاصل شده است.



شکل ۴. نمودارهای کنترلیX/R

طبق اصول مرتبط با کنترل کیفیت آماری، لازم است علل مربوط به انحرافات بادلیل که باعث خروج مشخصه کیفی از محدوده کنترلی شده است را شناسایی نموده و پس از رفع آن، نمودار را دوباره ترسیم نمود. وضعیت موجودِ نمودار، حاکی از آن است که در صورت استفاده از روش یاد شده، عملا

www.SID.ir

زمان بسیار زیادی در این راه، صرف خواهد شد. لذا این تکنیک مقرون به صرف نمی باشد. با توجه به این شرایط، تکنیک داده کاوی می تواند راهگشا باشد. بدین شکل، می توان با بررسی اطلاعات موجود در مورد ۴ پارامتر کیفیِ یاد شده، و با رسم نمودار میله ای برای آنها، به جای استفاده از روشهای قبلی، آن پارامتری را که بیشترین انحراف، نسبت به میانگین خود را دارا می باشد، ملاک بررسی قرار داده و با تحت کنترل در آوردن آن ، عملا دامنه بررسی را به سه پارامتر محدود نمود.

کاربرد داده کاوی در آنجاست که به جای بررسی حجم بالایی از پارامترهای تاثیر گذار، با درجه بندی و خوشه بندی اطلاعات بر اساس اهمیت آنها و تاثیر گذاریِ آنها بر پارامتر کیفیِ مورد نظر، بتوان دامنه عمل را محدود نموده و در کمترین زمان و با حداقل هزینه به هدف رسید.

۴-۱ نمودارهای کنترل در فاز 1 به عنوان یک ابزار داده کاوی

شش سیگما یک موضوع جالب و داغ بین متخصصان کیفیت است و داده کاوی یک موضوع داغ میان متخصصان سیستمهای اطلاعاتی است .شش سیگما یک فرآیند سیستماتیک است که با کاهش پراکندگی سعی در بهبود کیفیت دارد .داده کاوی از انبار داده استفاده می کند ، اطلاعاتی را استخراج می کند تا عملکرد تجاری سازمان را بهبود دهد .این دو رویکرد همدیگر را تکمیل می کنند اما تفاوتهایی با هم دارند[10].

داده کاوی از حجم وسیعی از داده ها برای استخراج اطلاعات و کشف دانش استفاده می کند در صورتی که مهندس کیفیت با استفاده از داده های کمی ، تجزیه و تحلیل خود را انجام می دهد .داده کاوی از ابزارهای اتوماتیک یا نیمه اتوماتیک برای استخراج اطلاعات استفاده می کند که این در تضاد با کاربردهای سنتی کیفیت است که کار استخراج اطلاعات به صورت دستی انجام می شود. نام برد .در واقع ما یک off-line کاوش داده ها برای رسیدن به کیفیت بیشتر نزدیک به چیزی است که تاگوچی [11]از آن به عنوان کیفیت انبار داده داریم در داخل این انبار جستجو می کنیم تا الگویی را استخراج کنیم و بر اساس آن راهی را پیدا کنیم تا بتوانیم کارها را بهتر و داده کاوی همدیگر را On-Line انجام دهیم .تکنیک های داده کاوی پیچیده تر از ابزار ساده کنترل فرآیند آماری هستند .پردازش تکمیل می کنند .تکنیک های پیشرفته کیفیت مانند طراح آزمایشها و آنالیز قابلیت اطمینان در زیر شاخه داده کاوی قرار می گیرند.

متخصصان کیفیت باید با متخصصان سیستمهای اطلاعاتی همکاری کنند و اطلاعتشان را در اختیار هم بگذارند. با استفاده از تکنیکهای داده کاوی فقط می توان الگوها را کشف کرد ولی شاید علت اینکه این الگو ایجاد شده است را پیدا نکنیم .بنابراین بایستی از متخصصان کیفیت و قابلیت اطمینان که دارای انباری از ابزارها برای کشف علتها می باشند استفاده کرد .مثالی از یک بانک در مقاله ای مطرح شد که مدیر بانک در صدد است با استفاده از داده های بسیاری که در دست دارد بفهمد که چه مدت طول می کشد تا یک مشتری حساب خود را ببندد و چه عاملی سبب می شود که مشتریان پس از مدتی حساب خود را ببندند .با استفاده از تکنیک داده کاوی نشان داده شد که مشتریانی که از طریق وب حساب باز کرده اند و از آن استفاده می کنند بیشتر حساب خود را در بانک می گذارند تا مشتریانی که مستقیماً به بانک مراجعه می کنند .اما شاید این امر به خاطر خبرگی آنها باشد و نه راحتی استفاده از وب یا خوب بودن و راحتی آن .لذا بایستی از یک متخصص کیفیت برای پیدا کردن علت این امر استفاده کرد.[9]

ابزارهای کنترل فرایند آماری و مدلهای نقطه تغییر شامل ویژگیهایی هستند که این ویژگیها آنها را برای روشهای کلاسبندی مناسب میگرداند. این ویژگیها عبارتند از:

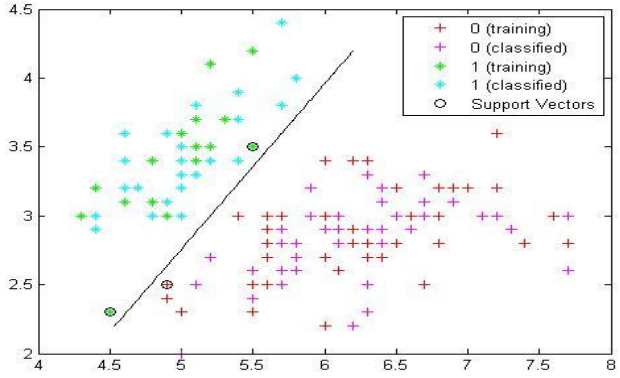
۱. در کنترل فرایند آماری دو حالت ممکن و متمایز در هر فرایند وجود دارد، حالت تحت کنترل و حالت خارج از کنترل، این دو حالت میتوانند به عنوان دو کلاس احتمالی برای فرایند کلاسبندی در نظر گرفته شوند.
۲. مشابه روشهای کلاسبندی، یک نمودار کنترل ابزاری برای طبقه بندی الگوهاست. تقریبا همه مشاهدات تحت کنترل به عنوان الگوی غیرتصادفی در نظر گرفته شوند.
۳. مشابه روشهای کلاسبندی در کنترل فرایند آماری یک مشاهده براساس براساس میزان مجاورتش با پارامترهای خارج کنترل و اندازه فاصله اش از پارامترهای تحت کنترل به عنوان یک مشاهده خارج از کنترل یا تحت کنترل قلمداد میشود.

www.SID.ir

همانگونه که مطرح شد در فاز 1 ما با مجموعه ای از داده های قدیمی سر و کار داریم .در این فاز هیچ اطلاعاتی از فرآیند در دسترس نمی باشد لذا بایستی با استفاده از این داده های قدیمی فرآیند را با ثبات پارامترهای فرآیند را تخمین زد .نحوه کار بدین صورت است که مشاهدات بر روی نمودار کنترل رسم و اگر هشداری توسط نمودار کنترل داده شود انحراف با دلیل مرتبط با آن نقطه پیدا و نقطه مربوطه از روی نمودار کنترل حذف می شود و محاسبات بر اساس نقاط باقی مانده تکرار می شود .آن قدر این عمل انجام می شود تا اینکه فرآیند تحت کنترل آماری درآید .بر اساس فرآیندی که تحت کنترل آماری است می توان پارامترهای فرآیند را تخمین زد و نمودار کنترلی برای کنترل مشاهدات آتی را طراحی کرد .این مجموعه داده های قدیمی می تواند شامل هزاران رکورد باشد لذا نمونه گیری از این انبار داده 2 و سپس استفاده از آنها برای طراحی یک نمودار کنترل و برآورد پارامترهای مجهول فرآیند به نوعی در تعریف داده کاوی می گنجد .نمودارهای کنترل طراحی شده بر اساس داده های یک انبار داده ، نحوه قرار گیری مشاهدات آتی روی نمودار را نشان می دهند و برای کنترل و بهبود فرآیند استفاده می شوند لذا با توجه به تعریف داده کاوی می توان نمودارهای کنترل را نیز به عنوان یک ابزار داده کاوی مدنظر قرار داد.

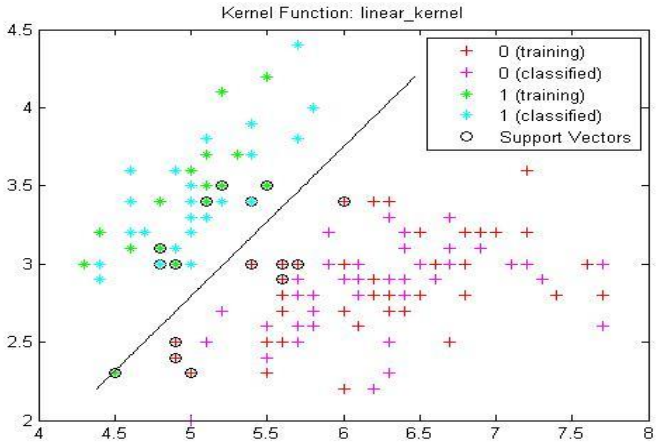
۴-۲ شبیه سازی

در بردار های آموزشی زیر اعداد مربوط به کلاسها یا حالت تحت کنترل و خارج از کنترل براساس داده کاوی و تعریف بردار ویژگی است، در ستون آخر بردار قرار دارند و مولفه های بردار تعریف شده براساس میزان فاصله از میانگین تعریف میگردند. شکل۴ کلاسبندی ورودی با توجه به تعریف کلاسها پس از آموزش براساس کلاسهای مربوط به کلاس اول و کلاس دوم با استفاده از ماشین بردار پشتیبان (SVM) میباشد. شکل۵ کلاسبندی دیتاست ورودی با توجه به تعریف کلاسها ب پس از آموزش براساس کلاسهای مربوط به کلاس اول و کلاس دوم با ماشین بردار پشتیبان (SVM) و استفاده از تابع کرنل می‌باشد.



شکل۵ - کلاسبندی کلاسهای اول و دوم با استفاده از ماشین بردار پشتیبان (SVM)

www.SID.ir



شکل۶. خوشه بندی کلاسهای اول و دوم با استفاده از ماشین بردار پشتیبان (SVM) با تابع کرنل

۵. نتیجه گیری

در نمودارهای مختلف کنترل فرآیند آماری، بواسطه رخداد نوسانات طبیعی در فرآیند تولید، تشخیص صحیح و سریع الگوهای غیرطبیعی و انحرافات معنی دار در بازهٔ زمانی کوتاه با چالشهایی مواجه است. از اینرو ارائه مدلهای مبتنی بر کلاسبندی کننده ماشین بردار پشتیبان با هدف شناسایی و تفسیر الگوهای نمودارهای کنترل فرآیند منطقی میباشد. پژوهش جاری مدلی را جهت تشخیص و بررسی الگوهای پایه و همزمان پیشنهاد مینماید .این مدل با توجه نمونه های و شرایط کنونی خط تولید بطور هوشمند حالت تحت کنترل و حالت خارج از کنترل را تشخیص میدهد.

منابع

[1] غضنفری م،ملک محمدی س،علیزاده س.(۱۳۸۹). بخش بندی مشتریان در صادرات پوشاک بر پایه الگوریتم های خوشه بندی. نشریه پژوهشنامه بازرگانی.(۱۴) ۵۶ . ۵۹ - ۸۶

[2] Chang, S. I. and Aw, C. (1996). "A neural fuzzy control chart for detecting and classifying process mean shifts." *International Journal of Production Research*, Vol. 34, No. 8, PP. 2265–2278.
[3] Cheng, C. S. (1995). "A multi-layer neural network model for detecting changes in the process mean." *Computers & Industrial Engineering*, Vol. 28, No. 1, PP. 51–61.
[4] Cheng, C. S. (1997). "A neural network approach for the analysis of control chart patterns." *International Journal of Production Research*, Vol. 35, No. 3, PP. 667–697.
[5] Chiu, C., Chen, M. and Lee, K. (2001). "Shifts recognition in correlated process data using a neural network." *International Journal of Systems Science*, Vol. 32, No.2, PP. 137–143.
[6] Guh, R. S. and Hsieh, Y. (1999). "A neural network based model for abnormal pattern recognition of control charts." *Computers & Industrial Engineering*, Vol. 36, PP. 97–108.
[7] Guh, R. S., Zorriassatine, F. and O'Brien, C. (1999). "On line control chart pattern detection and discrimination _ a neural network approach." *Artificial Intelligence in Engineering*, Vol. 13, PP. 413–425.
[8] Guh, R. S. and Tannock, J. (1999). "Recognition of control chart concurrent patterns using a neural network approach." *International Journal of Production Research*, Vol. 37, No. 8, PP. 1743–1765. 8 –
[9] Guh, R. S. (2004). "Optimizing feed forward neural networks for control chart pattern recognition through genetic algorithms." *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 18, No. 2, PP. 75–99.
[10] Guh, R. S. (2005). "A hybrid learning-based model for on line detection and analysis of control chart patterns." *Computers & Industrial Engineering*, Vol. 49, PP. 35–62.
[11]Genichi Taghchi ,(1985), **Quality engineering in japan**, Communication in statistics -Theory and Methods , Volume 14, 1985 – Issue 11, Pages 2785-2801
[12] Michael D. Samuel and Edward O. Garton.(1985). A Weighted Normal Estimate and Tests of Underlying Assumptions. The Journal of Wildlife Management, Vol. 49, No. 2 , pp. 513-519
[13] Morton klain, (2018)Two Alternatives to the Shewhart $\bar{X}$ Control Chart,jornal of quality technology, Pages 427-431
[14sean L. simpson,(2010), An adjusted likelihood ratio test for separability in unbalanced multivariate repeated measures data, Statistical Methodology,Volume 7,Issue 5 ,Pages 511-519

www.SID.ir

[15]Joseph J.Pignatiellojr.MarcusB.peery,(2006),Estimation of the change Point of aNormal Process Mean with a Linear Trend Disturbance in SPC ,Quality Technology&Quanitative Management ,Volume 3,Issue3,Pages325-334

[16] Vladimir Vapnik,(1998), The Support Vector Method of Function Estimation, **Nonleanear Modeling ,pp 55-85**

[17] Thomas .R.samuel, Joseph J. Pignatiello, James A. Calvin(1998). IDENTIFYING THE TIME OF A STEP CHANGE IN A NORMAL PROCESS VARIANCE.Quality engineering , volume 10 1998- Issue .3Pages 529-538 .

[18] **Willie E. Hopkins**,**(2002) ,** Effects of Cultural Recomposition on Group Interaction Processes, **,** Academy of Management ReviewVol. 27, No. 4

www.SID.ir