مهندسی صنایع و مدیریت شریف، (تابستان ۱۳۹۵)
دوره‌ی ۱ - ۳۲، شماره‌ی ۱/۲، ص. ۱۳-۲۱

Original Article

ارائه‌ی رویکردی برای کاهش هزینه‌های کیفیت در اندازه‌گیری پارامترهای شیمیایی با استفاده از تکنیک‌های داده‌کاوی

الهام آخوندزاده نوقابی (دکترا)
دانشکده‌ی مهندسی صنایع، دانشگاه تربیت مدرس
مهدی دانشمندی* (کارشناس ارشد)
دانشکده‌ی مهندسی صنایع، مؤسسه آموزش عالی ابوانکی
بهروز مینایی بیدگلی (استادیار)
دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت

در بازارهای رقابتی امروزی، مقوله‌ی کیفیت بیش از پیش مورد توجه تولیدکنندگان قرار گرفته است. در این راستا، یکی از موضوعاتی که از اهمیت زیادی برخوردار است، بحث پیرامون هزینه‌های کیفیت است. تولیدکنندگان همواره برسر دوراهی بین کیفیت بالای محصولات و هزینه‌های آن قرار می‌گیرند. در این نوشتار سعی شده تا با استفاده از تکنیک‌های داده‌کاوی، داده‌های مربوط به پارامترهای شیمیایی رنگ‌آمیزی به‌روش موقعیت‌زدایی الکتریکی (الکترودیپوزیشن)[1] که برای کنترل فرایند رنگ‌آمیزی کاربرد دارد، مورد کاوش قرار گیرد و از این طریق دانشی استخراج شود تا بتوان از آن در زمینه کاهش تعداد نمونه‌گیری محصولات استفاده کرد، به‌گونه‌یی که از کیفیت محصولات نکاهد و ریسک خرابی محصولات نیز افزایش پیدا نکند. برای این منظور از تکنیک قواعد باهم‌آیی استفاده شده است و تعداد پارامترهای نمونه‌گیری با استفاده از نتایج به دست آمده کاهش یافته است. این رویکرد می‌تواند در موارد مشابه نیز به‌منظور کاهش ابعاد استفاده شود.

e.akhondzadeh. n@gmail.com
m.daneshmandi@yahoo.com
b_minaei@iust.ac.ir

واژگان کلیدی: کنترل کیفیت، الکتودیپوزیشن ، داده‌کاوی، قواعد باهم‌آیی.

۱. مقدمه

امروزه دانش داده‌کاوی[2] و کشف دانش از حجم عظیمی از داده‌ها در بسیاری از علوم کاربرد دارد. پایگاه‌های داده‌یی حجیم که در دل خود دانش بسیار سودمندی دارند به این دلیل که در نگاه اول فقط دارای اعدادی بی‌معنا هستند، هرگز مورد مطالعه قرار نمی‌گیرند. حال آن که با بررسی مناسب می‌توان قواعدی جذاب و مفید از داخل همین داده‌های بی‌معنا استخراج کرد و این مهم با داده‌کاوی میسر خواهد شد.

از طرفی علوم بسیاری نظیر تولید و کنترل کیفیت نیز از داده‌کاوی بهره‌های مفیدی برده‌اند. تحقیقات انجام شده در این زمینه بیان می‌دارد که این حوزه به‌عنوان یکی از حوزه‌های کاربردی داده‌کاوی مورد توجه قرارگرفته است. در این مورد به‌طور خاص می‌توان اشاره کرد به کنترل فرایند آماری (SPC)[3] که اعداد و ابعاد زیادی از محصولات را اندازه‌گیری و ثبت می‌کند.

الگوریتم‌های رویکرد کنترل فرایند آماری (SPC) در مواجهه با فرایندهای پیچیده و حجیم دارای محدودیت‌هایی است؛ و این مسئله‌یی است که الگوریتم‌های داده‌کاوی، به‌دلیل تجزیه و تحلیل مقادیر زیادی از داده‌ها و نیز قابلیت‌های اثبات شده‌ی خود، از آن مبرا هستند.[1]

همچنین استفاده از روش‌ها و الگوریتم‌های داده‌کاوی به‌دلیل پیشرفته بودن در سیستم جمع‌آوری داده‌ها و ابزار آنالیز و روش‌های شرح قوی[2] در فرایندهای بهبود مستمر (نظیر 9-Steps, 8D, 7-Step, PDCA, Six Sigma-DMAICS, Kaizen- PDCA) بسیار مفید و مناسب است.[3]

هم‌اینک بیشتر محققین و حتی مدیران عملیاتی علاقه مشترکی برای به خدمت گرفتن و توسعه‌ی توانایی‌ها و الگوریتم‌های غیر نظارتی دارند تا بتوانند با صرفه‌جویی در زمان و محصولات معیوب، حتی دانش فعالیت‌های به شدت پیچیده و مهندسی را تحت کنترل بگیرند.[4] این‌گونه رویکردهای سیستمی به درک علل اصلی و ریشه‌های بنیادین مشکلات و همچنین پیاده‌سازی و ایجاد دانش‌های پیچیده نیزکمک خواهند کرد.[5]

در یکی از مطالعات انجام‌شده برای نظارت بر فرایندهایی با بازدهی بالا[6] که از نمودارهای کنترل اقلام نامنطبق و نمودار C استفاده می‌کنند، با استفاده از شبکه‌های عصبی مصنوعی روشی ارائه شده است که با استفاده از آن به‌جای ترسیم دو نمودار فوق به‌صورت مجزا، بر یک نمودار تجمیعی که شرایط خارج ازکنترل را به‌گونه‌یی مؤثرتر

* نویسنده مسئول
تاریخ: دریافت ۱۳۹۲/۱۱/۵، اصلاحیه ۱۳۹۳/۱۲/۱۳، پذیرش ۱۳۹۴/۴/۲.

تشخیص می‌دهد، تمرکز می‌کند. در مطالعه‌یی دیگر[۷] نیز بر محصولات تولیدی با جوشکاری قوس الکتریکی که از فرایند کنترل کیفیت برخط بهره‌مندند، با استفاده از نمودار شوهارت و درخت تصمیم‌گیری[۴] و نیز با رویکرد داده‌کاوی با سری‌های زمانی، کنترل و نظارت انجام گرفته است.

همچنین برای به دست آوردن زمان دقیق تغییرات در نمودارهای کنترل از رویکرد داده‌کاوی فازی و نیز آمار استفاده شده است[۸] که در آن با در نظر گرفتن هر دو نمونه‌برداری ثابت و متغیر روش ترکیبی جدیدی از خوشه‌بندی فازی و مفاهیم آماری پیشنهاد شده است.

محققین با ادغام الگوریتم‌های داده‌کاوی و الگوریتم‌های سنتی SPC (الگوریتم‌هایی همچون شبکه‌های عصبی مصنوعی، رگرسیون و ماشین بردار پشتیبان) به نظارتی کارآمدتر در فرایندهای همبسته و چندگانه رسیده‌اند.[۱] در اینجا برای ساخت نمودارهای تجمعی کنترل چندمتغیره‌یی که بتوانند بر روند فعالیت نظارت داشته باشند از مدل‌های مذکور بهره‌برداری شده است و در پایان نیز با شبیه‌سازی سناریوهای پیشنهادی توسط الگوریتم‌های داده‌کاوی که در فرایند بهبود مستمر نمودارهای کنترل سنتی مبتنی بر سری‌های زمانی با متدهای قدیمی کار می‌کردند، نتایج بسیار مفیدی ارائه شده است. ساختار الگوریتم مورد استفاده در این نوشتار اخیراً نیز بسیار مورد توجه قرارگرفته و حتی در پاره‌یی از موارد[۴] عملکرد فرایند آن مورد بهبود و تکامل نیز قرار گرفته است.

در نوشتار حاضر نیز کاربرد تکنیک‌های داده‌کاوی در حوزه‌ی کیفیت مورد بررسی قرار می‌گیرد و رویکرد جدیدی از کاربرد قواعد باهم‌آیی در کاهش هزینه‌های اندازه‌گیری ارائه می‌شود. از آنجا که بهبود مستمر همواره مورد تأکید هر خط کنترل کیفیتی است، در اینجا سعی داریم تا الگوی هوشمندی برای این مهم بیابیم. بدین ترتیب که با یافتن رابطه‌های منطقی بین پارامترهای کنترلی می‌توان تعداد پارامترهای مورد بازرسی را کاهش داد و تنها با اندازه‌گیری چند پارامتر مهم سطح کیفیت محصول تولیدی را ارزیابی کرد و سریعاً جهت افزایش کیفیت آن یا انجام اقدام اصلاحی مناسب گام برداشت. به عبارتی در مصرف زمان و هزینه‌های کیفیت[۵] بیشترین صرفه‌جویی انجام خواهد شد.

از آنجا که کاهش آثار منفی و عیوب و کاستی‌های محصول، یا افزایش سود حاصل از تولید محصولات مصداقی برای بهبود مستمر سیستم است،[۹] رویکرد ارائه شده در این مطالعه را می‌توان به‌عنوان جزئی مؤثر در فرایند بهبود مستمر معرفی کرد.

در این مقاله ابتدا به معرفی مختصر داده‌کاوی و خط رنگ‌آمیزی مورد آزمایش می‌پردازیم. سپس به شرح پارامترهای مورد بررسی و چگونگی آماده‌سازی و پایش آنها برای گروه‌بندی‌شان خواهیم پرداخت. سپس قواعدی را با استفاده از روش‌های داده‌کاوی به دست می‌آوریم که میزان وابستگی هریک از این پارامترها به یکدیگر را مستقل از پارامتر دیگر بیان می‌کند. در بخش آخر نیز جمع‌بندی و نتیجه‌گیری بیان شده و پیشنهادهایی برای تحقیقات آتی ارائه می‌شود.

۲. رنگ‌آمیزی به‌روش (الکترودیپوزیشن)[۶]

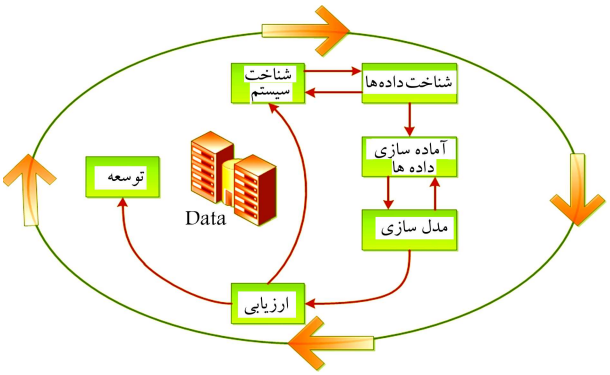
خطوط رنگ‌آمیزی به‌روش الکترودیپوزیشن ــ به‌دلیل کیفیت بالای چسبندگی رنگ و همچنین یک‌دستی رنگ محصول ــ در بسیاری از صنایع پراهمیت مانند خودروسازی استفاده می‌شود. از این خطوط در حال حاضر در شرکت‌های معتبر داخلی نظیر خودروسازان استفاده می‌شود و تمامی رنگ‌های زیرین (آستر) با استفاده از این نوع خطوط رنگ‌آمیزی می‌شود. برای رنگ‌آمیزی ابتدا سطح فلز را کاملاً شست‌وشو می‌دهند تا عاری از هرگونه چربی و آلودگی شود. سپس با برخی کاتیون‌های فعال سطح فلز را می‌پوشانند تا کاملاً آماده‌ی پذیرش رنگ شود. سپس قطعه‌ی مورد نظر را در حوضچه‌یی از رنگ فرو می‌برند و با بهره‌برداری از روش‌های کاتدی و آندی سطح فلز را رنگ‌آمیزی می‌کنند؛ و در پایان رنگ را در حرارت کوره ثابت و خشک خواهند کرد.

۳. داده‌کاوی

داده‌کاوی و کشف دانش (KDD)[۷] حوزه‌یی نسبتاً جدید و رو به رشد است[۱۰] که حوزه‌هایی چون آمار، یادگیری ماشین[۸]، پایگاه داده‌ها، مصورسازی[۹]، محاسبات موازی، الگوریتم‌ها و کسب دانش[۱۰]، و نیز چند رشته‌ی دیگر را با هم تلفیق می‌کند[۱۱و۱۲] تا دانش سودمند اما نهفته در انبوه داده‌های پایگاه‌های داده‌یی بزرگ را استخراج و سپس قواعدی تولید کند که حتی دست‌اندرکاران مرتبط با آن پایگاه را نیز شگفت‌زده کند.[۱۳] در حال حاضر KDD یکی از رو به رشدترین سرفصل‌های تحقیقاتی در دانش‌های رایانه‌یی و ریاضی است، زیرا توانایی بالایی برای تغییرات پیوسته و به دست آوردن فهم جدید و مداوم از داده‌ها دارد.[۱۴] مزیت اساسی این روش بر دیگر روش‌ها (نظیر آمار) وسعت زیاد الگوریتمی و داده‌یی آن است. به‌عبارتی داده‌کاوی مطابق الگوریتم‌ها و روش‌های مختلف خود تقریباً بر همه نوع داده‌ی کمی و کیفی قابلیت پیاده‌سازی دارد. حوزه‌های تجارتی، مدیریتی، پزشکی، ورزشی، اقتصادسنجی و... را می‌توان به‌عنوان حوزه‌های کاربردی داده‌کاوی نام برد.[۱۳]

داده‌کاوی فرایند به خدمت گرفتن یک روش‌شناسی رایانه‌یی است که با استفاده از تکنیک‌ها و الگوریتم‌های مختلف دانش نهفته در داده‌ها را جستجو می‌کند.[۱۵] این فرایند مشارکتی میان انسان و رایانه[۱۶] در نهایت به دنبال کشف الگوها و قواعد معنادار در میان داده‌هاست. دو هدف کارکردی داده‌کاوی را «پیش‌بینی[۱۱]» و «توصیف[۱۲]» ذکر کرده‌اند. در روش‌های پیش‌بینی به برآورد یک متغیر هدف با استفاده از مقادیر چند متغیر دیگر می‌پردازیم، و در نوع توصیفی نیز بیشترین تمرکز بر یافتن الگوها و روابط قابل توصیف در میان داده‌هاست.[۱۱]

در روش‌های داده‌کاوی از الگوریتم استانداردی باید استفاده کرد که مراحل انجام یک مدل‌سازی را به‌صورت کامل و قدم به قدم برای ما تشریح کند. این الگوریتم را با نام «کریسپ[۱۳]» می‌شناسیم و برای استخراج قاعده از داده‌ها دقیقاً مطابق با شش مرحله‌ی آن (شکل ۱) پیش می‌رویم. در نوشتار حاضر نیز از این روش‌شناسی استاندارد داده‌کاوی (هر چند به‌شکل تعدیل‌شده) و با گام‌های زیر استفاده می‌کنیم:



شکل ۱. نمودار مراحل روش‌شناسی CRISP-DM.

۱. درک موقعیت کسب و کار؛

۲. درک داده‌ها؛

۳. پیش‌پردازش و آماده‌سازی داده‌ها؛

۴. مدل‌سازی؛

۵. ارزیابی و استنتاج مدل؛

۶. به‌کارگیری و توسعه.

اگرچه عواملی همچون ارتباط شغلی، کیفیت خروجی، قابلیت نمایش نتایج، زمان واکنش و جواب‌گویی، و نوع داده‌ها[۱۷] در پذیرش داده‌کاوی مؤثرند، در وهله‌ی اول باید به بعد وظیفه‌گرایی توجه داشت. توجه به این عامل ما را از حضور مقادیر خوب و مناسب شاخص‌های میانگین ارتقاء، مؤثر بودن اطلاعات در دسترس، فرایند حل و استفاده مردم و سازمان‌ها مطمئن می‌سازد.[۱۸]

۱.۳. قواعد باهم‌آیی[۱۴]

داده‌کاوی دارای شش گروه متفاوت ابزار اصلی است که قواعد باهم‌آیی یکی از آنهاست.[۱۵و۱۹] استخراج قواعد باهم‌آیی یک تکنیک غیر نظارتی[۱۵] در داده‌کاوی است. این روش در تلاش و جست‌وجو برای یافتن ویژگی‌هایی از متغیرهاست که در بیشتر مواقع و با احتمال نسبتاً بالایی با یکدیگر همراه می‌شوند. به عبارت دیگر این روش به تحلیل وابستگی‌ها و مطالعه‌ی ویژگی‌ها یا خصوصیاتی می‌پردازد که با یکدیگر همراهند. این روش به دنبال استخراج قواعدی به‌منظور کمّی‌کردن ارتباط میان دو یا چند خصوصیت است. به‌صورت رسمی‌تر قواعد باهم‌آیی پیروی کردن پارامترها از یکدیگر را به‌صورت جزء به جزء توضیح می‌دهد.[۲۰]

استفاده از قواعد باهم‌آیی برای بیان دلایل شکست (در اینجا اندازه پارامترهای خارج از محدوده‌ی بالا و پایین پذیرش) و استخراج دانش مفید از آن‌ها، در الگوهای متناوب فرایندهای صنعتی بسیار مفید خواهد بود.[۲۱]

قواعد باهم‌آیی به شکل $x \rightarrow y'$ بیان می‌شوند که هرکدام از X و Y زیرمجموعه‌یی دلخواه از کل داده‌های موجود است. این سری از قواعد در جست‌وجوی ویژگی‌های پنهانی هستند که معمولاً با یکدیگر همراه می‌شوند. قاعده‌ی $X \rightarrow Y$ همان بیان اگر X آنگاه Y را تداعی می‌کند؛ البته باید توجه داشت که ذات قواعد باهم‌آیی احتمالی است و بنابراین از خاصیت‌های رایجی مانند شرکت‌پذیری و توزیع‌پذیری و... پیروی نمی‌کند. برای اعتبارسنجی هریک از این قواعد از دو معیار پشتیبان[۱۶] و اطمینان[۱۷] استفاده می‌کنیم. معیار پشتیبان نشان‌گر درصدی از کل مجموعه‌ی تراکنش‌هاست که هر دو مجموعه‌ی X و Y در آنها موجود باشد. این معیار میزان جذابیت و مفید بودن قاعده‌ی استخراجی را به ما معرفی می‌کند. اما معیار اطمینان بیان‌گر میزان وابستگی مجموعه‌ی Y به X است؛ به عبارتی بیان می‌کند که چند درصد از تراکنش‌هایی که X را دارند Y را نیز شامل می‌شوند. برای محاسبه‌ی این معیار:

$$\text{Confidence} = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)}$$

برای استخراج قواعد با مشخص کردن کمینه‌ی هریک از این دو معیار، اکتشاف را شروع می‌کنیم.

۲.۳. الگوریتم اپریوری[۱۸]

الگوریتم اپریوری در سال ۱۹۹۳ ابداع شد[۲۰] و به‌لحاظ صرفه‌جویی بسیار در وقت و حافظه‌ی مورد نیاز برای جست‌وجوی پایگاه داده شدیداً مورد توجه قرار گرفت.

به‌عبارتی این الگوریتم به‌جای کنکاش در یافتن اقلام پرتکرار به جست‌وجو برای زیرمجموعه‌های پرتکرار در پایگاه داده می‌پردازد؛ و چون اقلام یک زیرمجموعه‌ی پرتکرار خود نیز هم‌زمان پرتکرارند پس با یک بار جست‌وجو و خواندن پایگاه داده به‌راحتی می‌توان تمام اقلام تکراری را در زیرمجموعه‌های بزرگ شناسایی کرد و بدین طریق نیازی به خواندن چندین مرتبه‌یی پایگاه داده توسط نرم‌افزار نخواهد بود. طریقه‌ی محاسبه‌ی اقلام پرتکرار (C_k) در ادامه آمده است.[۱۰] در این جا می‌بینیم که C_kهای اضافی چگونه حذف می‌شوند:

$Apriori - gen(L_{k-1})$

Join Step

Insert Into C_k

Select $p.item_1, p.item_2, ...p.item_{k-1}, q.item_{k-1}$

from $L_{k-1}p, L_{k-1}q$

where $p.item_1 = q.item_1, ..., p.item_{k-2} = q.item_{k-2}$

Prune Step

$p.item_{k-1} < q.item_{k-1}$

for all item sets c in C_k do

for all $(k-1)$ − subsets s of c do

if (s isnot in L_{k-1}) then delete c from C_k

۴. تعریف مسئله

خط رنگ‌آمیزی به‌روش الکترودیپوزیشن دارای دو مرحله است: ۱. اقدامات اولیه (PT)[۱۹] شامل مراحل شست‌وشو، چربی‌زدایی[۲۰] در دو یا چند مرحله، آبکشی[۲۱] در یک یا دو مرحله، فعال‌سازی[۲۲]، فسفاته کردن[۲۳] تری‌کاتیونیک، خنثی‌سازی و آبکشی و چکه‌گیری با آب DI است که مجموعاً سطح فلز را پاک‌سازی می‌کند؛ ۲. الکترودیپوزیشن که مربوط به رنگ‌آمیزی فلز و شامل مراحل حوضچه‌ی رنگ ED، آبکشی، شست‌وشو با محلول فوق تصفیه (UF)[۲۴]، آبکشی با آب تازه DI، مرحله‌ی بدون تخلیه و کوره‌ی الکترودیپوزیشن (ED) است.

برای کنترل کیفیت محصولات و قطعات خروجی چون نمی‌توان خود قطعات را در هر مرحله خارج و بررسی کرد، از رویکرد کنترل فرایند استفاده می‌شود و با ثابت نگه‌داشتن چندین پارامتر شیمیایی و الکتریکی (حدود ۴۰ پارامتر) در هریک از مراحل مذکور کیفیت رنگ خروجی تضمین می‌شود. پارامترهای شیمیایی مورد نظر ما در جدول ۱ آمده است، اما این نمونه‌گیری و بررسی پارامترها زمان‌بر و پرهزینه است و برای تغییر هر یک از آنها در حد قابل قبول و در محدوده‌ی پذیرش، زمانی نسبتاً طولانی مورد نیاز خواهد بود. این نوع نمونه‌گیری، کنترل و انجام اقدامات اصلاحی که روشی سنتی و ساده‌یی به نظر می‌رسد علاوه بر کند بودن هزینه زیادی را در پی خواهد داشت، چرا که برای انجام اقدامات اصلاحی زمان کم‌تری در اختیار مهندسین فنی قرار می‌گیرد و این امر باعث معیوب و بی‌کیفیت شدن قطعات بیشتری می‌شود. از این رو، رویکرد مد نظر این مقاله به کاهش تعداد این پارامترها کمک شایانی خواهد کرد. این در حالی است که از کیفیت نتایج پایانی نیز کم نخواهد شد. اگرچه مراحل آماده‌سازی و رنگ‌آمیزی

جدول ۱. محدوده‌بندی پارامترهای کنترلی.

نام پارامتر	A <		≤ B <		≤ C <		≤ D <		≤ E
	حد پایین	حد بالا	حد پایین	حد بالا	حد پایین	حد بالا	حد پایین	حد بالا	حد پایین	حد بالا
COND. DI		۵	۵	۱۰	۱۰	۱۵	۱۵	۳۰	۳۰	
COND آب انتهای مسیر		۳۰	۳۰	۶۰	۶۰	۹۰	۹۰	۱۲۰	۱۲۰	
COND رنگ	۱۲۰۰	۱۴۰۰	۱۴۰۰	۱۵۰۰	۱۵۰۰	۱۶۰۰	۱۶۰۰	۱۸۰۰	۱۸۰۰	
PH آب DI		۴	۴	۶	۶	۷	۷	۸	۸	
PH آب انتهای مسیر		۳	۳	۴	۴	۵	۵	۶	۶	
PH خنثی‌سازی		۳	۳	۴	۴	۵	۵	۶	۶	
PH رنگ		۳	۳	۵	۵	۶	۶	۷	۷	
PH فعال‌سازی		۷	۷	۸	۸	۹	۹	۱۰	۱۰	
اسید آزاد فسفاته		۰/۵	۰/۵	۱	۱	۱/۵	۱/۵			
اسیدکل خنثی‌سازی		۴	۴	۵	۵	۶	۶	۷	۷	
اسیدکل فسفاته		۲۰	۲۰	۲۲	۲۲	۲۴	۲۴	۲۶	۲۶	
دمای فعال‌سازی		۲۰	۲۰	۲۵	۲۵	۳۰	۳۰	۳۵	۳۵	
دمای کوره		۱۸۰	۱۸۰	۱۸۵	۱۸۵	۱۹۰	۱۹۰	۱۹۵	۱۹۵	
شتاب‌دهنده	۰	۰/۵	۰/۵	۱	۱	۲	۲	۳/۵	۳/۵	
قلیائی‌شدگی آزاد	۰	۱۲	۱۲	۱۳	۱۳	۱۴	۱۴	۱۵	۱۵	
لجن فسفاته		۵	۵	۱۰	۱۰	۱۵	۱۵			

در نگاه اول جدا از یکدیگر به نظر می‌رسد، چون در یک خط رنگ پیوسته و با یک نوع تنظیمات قرار دارند و قطعات در حرکت بین مراحل ویژگی‌های هر مرحله را به مراحل بعدی منتقل خواهند کرد، ویژگی‌های مرتبط با هم خواهند داشت؛ یعنی تغییر در مقدار هر متغیر باعث تغییر در متغیر دیگری نیز خواهد شد. در اینجا رویکرد ما پی بردن به این روابط و کشف دانش نهفته در میان این داده‌هاست.

۵. روش‌شناسی تحقیق

مطابق شکل ۲ ابتدا به بررسی ویژگی‌های خاص این نوع خط رنگ می‌پردازیم. سپس به چگونگی جمع‌آوری داده‌های مورد نیاز خواهیم پرداخت و پیرامون کمیت و کیفیت داده‌های نمونه‌برداری شده شرح می‌دهیم. همچنین بیان می‌کنیم که در مرحله‌ی پیش‌پردازش داده‌ها برای معتبرشدن داده‌ها چگونه سعی کرده‌ایم که داده‌های غیر معتبر و خطاهای معمول (مانند خطای نوع اول یا دوم) را تا حد ممکن کم‌اثر کنیم.

بعد از آن شرح خواهیم داد که داده‌های موجود را چگونه رده‌بندی می‌کنیم به‌طوری که هر رده دارای معنای معتبر و پرمفهومی باشد. برای ترتیبی کردن این داده‌های پیوسته از محدوده‌ی کنترلی و حد فاصل ۶ انحراف معیار هریک از پارامترها مطابق با استانداردهای شرکت‌های تأمین کننده استفاده خواهد شد. سپس قواعد را از این داده‌های ترتیبی استخراج می‌کنیم. گفتنی است که میزان جذابیت هر قاعده

شکل ۲. روند روش‌شناسی تحقیق.

برای لحاظ شدن به‌عنوان یک قاعده‌ی مفید، بسیار حائز اهمیت قرار گرفته است و برای تحقق این مهم از چند تن از کارشناسان و سرپرستان باتجربه‌ی خط رنگ مذکور بهره گرفته شده است.

۶. پیاده‌سازی روش‌شناسی در مطالعه‌ی موردی

در اینجا خط رنگ ED یک شرکت قطعه‌ساز خودرو که انواع قطعات بدنه و لوازم جانبی را برای دو شرکت ایران خودرو و سایپا رنگ‌آمیزی می‌کند مورد بررسی قرار داده‌ایم. این قطعات در این شرکت رنگ آستر می‌شوند و رنگ‌آمیزی اصلی‌شان در داخل هریک از شرکت‌های فوق انجام می‌گیرد. این شرکت نیز چنان که بیان شد با مشکل عمده‌ی نمونه‌برداری‌های زیاد مواجه است که باعث افزایش زمان و هزینه‌ی تولید و درنتیجه افزایش قطعات بی‌کیفیت خواهد شد. همچنین به‌دلیل زیاد بودن پارامترهای کنترلی و تعداد دفعات نمونه‌برداری، احتمال تأثیر خطاهای انسانی و افزایش خطای اندازه‌گیری نیز بیشتر خواهد بود و درنتیجه امکان تولید قطعات بی‌کیفیت افزایش خواهد یافت. شرکت فوق دارای گواهی‌نامه‌های ISO۹۰۰۰,۹۰۰۱ در زمینه‌ی کیفیت است و کلیه‌ی الزامات استاندارد خودروسازان فوق را رعایت می‌کند. خط رنگ مذکور در دو شیفت کاری مشغول به کار است و در حد فاصل این دو شیفت کاری نیز خاموش نمی‌شود؛ درنتیجه تنظیمات[۲۵] آن فقط در ابتدای روز و برای یک مرتبه در طی روز انجام می‌گیرد.

نمونه‌برداری پارامترها ۴ مرتبه در طی یک روز (دومرتبه در هر شیفت کاری) و در فواصل زمانی برابر و مشخص و مطابق با استانداردهای شرکت‌های تولید کننده‌ی مواد اولیه انجام گرفته است (شرکت کانسای مواد لازم برای خط ED و شرکت هنکل نیز مواد اولیه‌ی خط PT آن را تأمین می‌کند). داده‌های این پژوهش در بازه زمانی حدود شش ماه جمع‌آوری شده که بیشتر از ۱۰۰۰۰ بار نمونه‌برداری را شامل می‌شود.

برای رسیدن به نتایج مطلوب‌تر و همچنین بی‌تأثیر یا کم تأثیرکردن خطاهای احتمالی در مراحل مختلف، در بین مقادیر روزانه میانگین‌گیری شده است و فرایند اکتشاف قواعد روی این میانگین‌ها به‌عنوان نماینده‌ی هر شیفت کاری پیاده‌سازی شده است. برای پایایی بیشتر قواعد استخراج شده، پارامترهای اندازه‌گیری شده‌در روزهای خاص از داده‌ها حذف شده‌است. منظور از روزهای خاص، شرایط غیرطبیعی مانند قطعی برق، توقف خط رنگ به هر دلیل اضطراری، برقراری دیواره‌ی آتش[26] برای عدم خروج قطعات معیوب و دیگر اتفاقات غیر منتظره و غیر طبیعی است.

باید توجه داشت که این بازه‌بندی داده‌ها با توجه به محدوده‌های مورد پذیرش برای هر پارامتر انجام گرفته است. این محدوده‌های قابل قبول برای هر پارامتر به‌صورت مجزا از سوی شرکت‌های تولید مواد اولیه‌ی همین خطوط رنگ و با توجه به استانداردهای رسمی آن‌ها استخراج می‌شود. در این خط رنگ از استانداردهای کنترل کیفیت شرکت‌های هنکل آلمان و کانسای ژاپن که به‌صورت رسمی از طرف همین کمپانی‌ها اعلام شده، استفاده شده است.

از آن جایی که داده‌های نمونه‌برداری شده از خط رنگ فوق تماماً کمی بوده‌اند، مطابق جدول ۱ در مرحله‌ی پیش‌پردازش داده، به داده‌های ترتیبی[27] تبدیل شده‌اند تا بتوان از الگوریتم Apriori روی آنها استفاده کرد. یعنی هر پارامتر براساس مقادیر کمی درج شده که می‌توانسته شامل شود به یکی از حروف A, B, C, D, E نسبت داده شده است.

۷. استخراج قواعد و تحلیل آن

پس از آماده‌سازی داده‌ها و اجرای الگوریتم Apriori به نتایج درج شده در جدول ۲ دست یافته‌ایم. در این جدول فقط ۱۵ قاعده از کل قواعد درج شده است. باید توجه داشت که روی همین داده‌ها به حدود ۱۰۰ قاعده با جذابیت و اطمینان مناسب دست یافتیم که فقط تعداد محدودی از آنها از گذشته برای بازرسان کنترل کیفی شرکت مشهود بود.

قواعد درج شده در جدول ۲ بر اساس مقدار معیار اطمینان مرتب شده است. تفسیر جدول هم مانند قواعد باهم‌آیی است، اما در تفسیر حروف هر پارامتر باید به بازه‌های درج شده در جدول ۱ نیز توجه داشت. برای مثال قاعده‌ی اول را باید چنین تفسیر کرد:

زمانی که PH آب انتهای مسیر برابر C شود (یعنی بازه بین ۴ تا ۵) آنگاه با اطمینان ۹۹٫۷۳۹٪ می‌توان گفت که PH رنگ نیز برابر C (یعنی در بازه بین ۵ تا ۶) است. این رخداد (یعنی برابر بودن همزمان PH آب انتهای مسیر با C و PH رنگ با C) در ۷۰٫۳۵٪ از کل رکوردهای جاری اتفاق افتاده که جذابیت نسبتاً بالایی نیز دارد.

قاعده ششم چنین تفسیر می‌شود:

اگر در آن واحد و همزمان هر دو شرط رسانایی الکتریکی (Cond)[28] آب انتهای مسیر برابر B (یعنی در بازه ۳۰ تا ۶۰) و اسید آزاد فسفاته برابر B (یعنی در بازه بین ۰٫۵ تا ۱) محقق شود، آنگاه با اطمینان ۸۲٫۷۶۸٪ می‌توان گفت که PH آب انتهای مسیر برابر C (یعنی در بازه بین ۴ تا ۵) قرار می‌گیرد. این اتفاق (یعنی بروز همزمان رسانایی الکتریکی آب انتهای مسیر = B و اسید آزاد فسفاته = B و PH آب انتهای مسیر = C) در ۲۶٫۹۸٪ از کل رکوردهای موجود به وقوع پیوسته است. به‌صورت واضح‌تر در حدود ۲۷٪ از کل دفعات نمونه‌برداری می‌توانیم با اندازه‌گیری دو مقدار مقدم در قاعده‌ی فوق با احتمال بسیار بالایی درمورد یکی دیگر از پارامترها بدون اندازه‌گیری آن نتیجه‌گیری کنیم. در اینجا به‌دلیل کمبود فضا و زمان از ذکر و تشریح دیگر قواعد اجتناب کرده‌ایم، اما گفتنی است که با رعایت کلیه‌ی این قواعد، کاهش نمونه‌برداری‌ها بسیار مشهود است و در نهایت به کاهش هزینه‌های کیفیت و جلوگیری زودتر از عدم انطباق محصولات خواهیم رسید.

۱.۷. تحلیل فراتر قواعد و نتایج کاربردی

در این قسمت سعی می‌کنیم تا از قواعد استخراج شده نتایج کاربردی و کلی‌تری استخراج شود، به‌گونه‌یی که با اندازه‌گیری یک معیار خاص به‌راحتی بتوان با احتمال و اطمینان بالایی درمورد کلیت یک پارامتر دیگر اظهار نظر کرد. براساس قواعد به دست آمده در سطح اول، حالت خاصی از یک پارامتر را می‌توان براساس مقدار خاصی از پارامتری که به آن وابسته است، تخمین زد. در مرحله‌ی دوم تحلیل سعی بر آن است که رویکرد جدیدی ارائه دهیم تا بتوان ارتباط بین پارامترها را در سطح کلان‌تر و مستقل از مقدار خاص آن‌ها استخراج کرد، به‌طوری که بتوان مقدار یک پارامتر را براساس پارامتری که به آن وابسته است، مستقل از مقادیر خاص پارامترها تخمین زد. در اینجا بر ۴ مورد از پارامترهای اندازه‌گیری PH (شامل PH آب DI، PH آب انتهای مسیر، PH خنثی‌سازی و PH فعال‌سازی) متمرکز می‌شویم. برای مثال در جدول ۳ نتایج وابستگی PH آب DI به PH آب انتهای مسیر مورد بررسی قرار گرفته است که نتایج حاصل از این بررسی عبارت است از:

۱. با افزایش PH آب DI، PH آب انتهای مسیر نیز نسبتاً و به‌تدریج افزایش می‌یابد.

۲. نتایج به‌صورت محدوده‌یی، کاملاً معنادارند. مثلاً زمانی که PH آب DI در بازه B قرار می‌گیرد با احتمال حدود ۹۲٪ می‌توان گفت که PH آب انتهای مسیر در بازه B یا C قرار خواهد گرفت، یا به‌صورت تفکیک شده برای بازه B با احتمال ۳۵٪ و برای بازه C با احتمال ۵۷٪ اتفاق خواهد افتاد. در ضمن این قاعده در ۱۴٫۲۲٪ از کل تراکنش‌ها به وقوع پیوسته است.

۳. با استفاده از جدول ۳ می‌توان حدود ۷۵٪ از کل مقادیر PH آب انتهای مسیر را از روی مقادیر PH آب DI به دست آورد.

در جدول ۴ نیز همین رویکرد، برای PH آب انتهای مسیر به‌عنوان مقدم[29] و PH خنثی‌سازی به‌عنوان تالی[30] و نتیجه‌ی قاعده به کار گرفته شده و نتایج آن نیز درج شده است.

اما چنان که در جدول ۵ نیز ذکر شده، درمورد PH فعال‌سازی اصلاً نمی‌توان نتیجه‌ی مشخصی گرفت. زیرا به هیچ‌یک از PHهای دیگر وابسته نیست و در هر رده از مقادیر PHهای دیگر، مقادیر کاملاً مستقل به خود را با احتمال تقریباً برابر در هر محدوده دارد.

در اینجا مشخص است که با جدول‌های فوق بهینه‌سازی محسوسی در تعداد نمونه‌برداری‌ها و زمان آنها صورت خواهد گرفت. قواعد استخراج شده در این مقاله برای مدت حدود ۲ ماه روی خط رنگ مذکور آزمایش شد. برای مثال مطابق جدول‌های ۳ تا ۵ با اندازه‌گیری PH آب DI در ابتدای مسیر رنگ‌آمیزی به PH چهار پارامتر دیگر مطابق همین جدول‌ها پی بردیم. سپس اگر هریک از این چهار پارامتر خارج از محدوده‌ی بالا و پایین قابل پذیرش خود بودند، اقدام اصلاحی بلافاصله توسط کارشناسان خط رنگ انجام می‌گرفت. تمامی این اقدامات فقط در مدت زمان ۱۰-۲۰ دقیقه صورت می‌پذیرفت. اما نمونه‌برداری از هر چهار پارامتر و اندازه‌گیری مقادیر هریک به‌صورت مجزا و پس از آن انجام اقدام اصلاحی بین

جدول ۲. قواعد استخراج شده به همراه شاخص‌های اعتبارسنجی آنها.

ردیف	مقدم (شرط)	تالی (نتیجه)	شاخص (٪)	
			اطمینان	پشتیبان قانون*
۱	آب انتهای مسیر = PH = C	رنگ = PH = C	۹۹٫۷۳۹	۷۰٫۳۵
۲	خنثی‌سازی = PH = C	رنگ = PH = C	۹۹٫۵۲۶	۷۷٫۲۵۶
۳	اسید آزاد فسفاته = B	رنگ = PH = C	۹۹٫۳۳۷	۶۸٫۹۶۹
۴	اسید کل خنثی‌سازی = C and رنگ = PH = C	خنثی‌سازی = C = PH	۹۲٫۹۴۷	۳۳٫۹۷۸
۵	COND. DI = A and خنثی‌سازی = PH = C and رنگ = PH = C	آب انتهای مسیر = C = PH	۹۲٫۲۸۵	۲۸٫۶۳۷
۶	آب انتهای مسیر = COND = B and اسید آزاد فسفاته = B	آب انتهای مسیر = C = PH	۸۲٫۷۶۸	۲۶٫۹۸
۷	آب DI = PH = C	خنثی‌سازی = PH = C	۸۱٫۵۲۴	۳۹٫۴۱۱
۸	اسید آزاد فسفاته = B and خنثی‌سازی = PH = C	آب انتهای مسیر = C = PH	۷۶٫۳۵۱	۴۱٫۶۲۱
۹	فعال‌سازی = PH = C and آب انتهای مسیر = PH = C	اسید آزاد فسفاته = B	۷۵٫۴۰۵	۲۵٫۶۹۱
۱۰	آب DI = PH = C and آب انتهای مسیر = PH = C	COND. DI = A	۶۴٫۹۲۵	۲۴٫۰۳۳
۱۱	COND. DI = A and آب انتهای مسیر = PH = C and خنثی‌سازی = PH = C	آب انتهای مسیر = B = COND	۶۴٫۷۴۴	۱۸٫۶
۱۲	قلیائیت آزاد = A and آب انتهای مسیر = PH = C	دمای فعال‌سازی = C	۶۳٫۳۰۹	۱۶٫۲۰۶
۱۳	COND. DI = A and آب DI = PH = C and رنگ = PH = C	آب انتهای مسیر = B = COND	۶۲٫۸۹۸	۱۶٫۳۹
۱۴	قلیائیت آزاد = A and خنثی‌سازی = PH = C and رنگ = PH = C	دمای فعال‌سازی = C	۶۱٫۱۸۹	۱۶٫۱۱۴
۱۵	رنگ = COND = D	دمای فعال‌سازی = C	۵۲٫۰۰۹	۲۱٫۴۵۵

* rule support

جدول ۳. شاخص‌های اعتبارسنجی قاعده: PH آب انتهای مسیر → PH آب DI.

ردیف	مقدم (شرط)	تالی (نتیجه)	شاخص (٪) اطمینان	شاخص (٪) پشتیبان قانون
۱	DI آب PH = B	آب انتهای مسیر PH = B	۳۵٫۰۴	۵٫۴۲
۲	DI آب PH = B	آب انتهای مسیر PH = C	۵۶٫۹۳	۸٫۸۰
	مجموع		۹۱٫۹۷	۱۴٫۲۲
۳	DI آب PH = C	آب انتهای مسیر PH = C	۳۰٫۸۷	۱۰٫۳۸
۴	DI آب PH = C	آب انتهای مسیر PH = D	۴۵٫۶۴	۱۵٫۳۵
	مجموع		۷۶٫۵۱	۲۵٫۷۳
۵	DI آب PH = D	آب انتهای مسیر PH = C	۲۷٫۹۶	۶٫۶۶
۶	DI آب PH = D	آب انتهای مسیر PH = D	۳۳٫۶۵	۸٫۰۱
۷	DI آب PH = D	آب انتهای مسیر PH = E	۲۳٫۷۰	۵٫۶۴
	مجموع		۸۵٫۳۱	۲۰٫۳۲

جدول ۴. شاخص‌های اعتبارسنجی قاعده: PH خنثی‌سازی → PH آب انتهای مسیر.

ردیف	مقدم (شرط)	تالی (نتیجه)	شاخص (٪) اطمینان	شاخص (٪) پشتیبان قانون (٪)
۱	آب انتهای مسیر PH = B	خنثی‌سازی PH = B	۳۳٫۰۳	۴٫۰۶
۲	آب انتهای مسیر PH = B	خنثی‌سازی PH = C	۴۵٫۸۷	۵٫۶۴
	مجموع		۷۸٫۹۰	۹٫۷۱
۳	آب انتهای مسیر PH = C	خنثی‌سازی PH = B	۳۶٫۷۰	۱۱٫۰۶
۴	آب انتهای مسیر PH = C	خنثی‌سازی PH = C	۲۰٫۹۷	۶٫۳۲
	مجموع		۵۷٫۶۸	۱۷٫۳۸
۵	آب انتهای مسیر PH = D	خنثی‌سازی PH = D	۶۱٫۱۸	۱۷٫۶۱
	مجموع		۶۱٫۱۸	۱۷٫۶۱
۶	آب انتهای مسیر PH = E	خنثی‌سازی PH = D	۳۵٫۲۶	۶٫۸۸
۷	آب انتهای مسیر PH = E	خنثی‌سازی PH = E	۲۹٫۴۸	۵٫۷۶
	مجموع		۶۴٫۷۴	۱۲٫۶۴

جدول ۵. شاخص‌های اعتبارسنجی قواعد استخراجی پارامتر PH فعال‌سازی.

مقدم (شرط)	تالی (نتیجه)	شاخص (٪) اطمینان	شاخص (٪) پشتیبان قانون (٪)
DI آب PH = B	فعال‌سازی PH = B	۲۸٫۴۷	۴٫۴۰
DI آب PH = D	فعال‌سازی PH = B	۲۱٫۸۰	۵٫۱۹
DI آب PH = E	فعال‌سازی PH = B	۱۸٫۵۴	۳٫۷۲
DI آب PH = E	فعال‌سازی PH = C	۵۳٫۳۷	۱۰٫۷۲
DI آب PH = D	فعال‌سازی PH = D	۳۳٫۱۸	۷٫۹۰
DI آب PH = E	فعال‌سازی PH = D	۲۶٫۴۰	۵٫۳۰
DI آب PH = B	فعال‌سازی PH = D	۲۳٫۳۶	۳٫۶۱
آب انتهای مسیر PH = C	فعال‌سازی PH = B	۲۹٫۵۹	۸٫۹۲
آب انتهای مسیر PH = B	فعال‌سازی PH = B	۲۹٫۳۶	۳٫۶۱
آب انتهای مسیر PH = D	فعال‌سازی PH = B	۱۹٫۶۱	۵٫۶۴
آب انتهای مسیر PH = B	فعال‌سازی PH = C	۴۹٫۵۴	۶٫۰۹
آب انتهای مسیر PH = E	فعال‌سازی PH = D	۳۶٫۹۹	۷٫۲۲
آب انتهای مسیر PH = D	فعال‌سازی PH = D	۳۶٫۰۸	۱۰٫۳۸
آب انتهای مسیر PH = C	فعال‌سازی PH = D	۲۲٫۴۷	۶٫۷۷
آب انتهای مسیر PH = B	فعال‌سازی PH = D	۲۰٫۱۸	۲٫۴۸

۲۵-۳۵ دقیقه زمان نیاز داشت. این حداقل ۱۵ دقیقه زمان ذخیره شده فقط با اندازه‌گیری یک پارامتر، با در نظر گرفتن رنگ‌آمیزی چهار قطعه‌ی بزرگ (مانند درب خودرو) در هر ۱۵-۲۰ ثانیه از صرفه‌جویی حدود ۱۸۰-۲۰۰ قطعه‌ی معیوب جلوگیری به عمل می‌آورد. باید توجه داشت که این تنها هزینه‌ی تولید قطعات معیوب است و باید هزینه‌ی کاهش نمونه‌های اندازه‌گیری نشده را نیز به این صرفه‌جویی اضافه کنیم.

واضح و مشخص است که ادامه‌ی این روند برای دیگر قواعد مطرح شده در جدول ۲ این بهینه‌سازی را تا مقادیر بالاتری افزایش می‌دهد. به‌صورت خلاصه در بیشتر از ۷۰٪ موارد تنها با اندازه‌گیری سه پارامتر PH آب DI، اسیدکل خنثی‌سازی و قلیائی‌شدگی آزاد (و البته در برخی از موارد نیز اسید آزاد فسفاته) می‌توانستیم درمورد ۷ پارامتر دیگر اظهار نظر صحیح و درستی داشته باشیم. دقت این تصمیمات نیز در جدول ۲ به تفکیک هر قاعده درج شده است. مطابق نظر کارشناسان خط رنگ مذکور این صرفه‌جویی در نمونه‌برداری و زمان علاوه بر کاهش هزینه‌های نمونه‌برداری و مستقیم به میزان ۲۰-۳۰ درصد از تولید قطعات با کیفیت پایین و معیوب صرفه‌جویی به عمل می‌آورد.

۸. نتیجه‌گیری

در این مطالعه، رویکرد جدیدی با استفاده از قواعد باهم‌آیی برای کاهش هزینه‌های کیفیت ارائه شد. با استفاده از قواعد به دست آمده می‌توان بررسی پارامترهای کنترل فرایند خط رنگ مذکور را تا حدود بسیار زیادی (و با کاهش تعداد نمونه‌گیری‌های پارامترها) کاهش داد، طوری که کیفیت رنگ‌آمیزی قطعات کاهش نیابد. همچنین می‌توان با اندازه‌گیری چند پارامتر کلیدی که بر دیگر پارامترها تأثیرگذارترند، مقدار دیگر پارامترها را تخمین زد تا به‌راحتی و با کم‌ترین مقدار هزینه بتوان در بهترین زمان

ممکن اقدامات اصلاحی مفید و با اثربخشی بالا انجام داد تا کیفیت محصولات خروجی همچنان در حد مطلوب قرار بگیرد. این رویکرد در سرعت بخشیدن به فرایند کنترل کیفیت و همچنین کاهش هزینه‌های آن ضمن افزایش عملکرد و کارایی آن‌ها نقش به‌سزایی ایفا می‌کند. هدف کلی از قواعد کاربردی ارائه شده نیز این است که به‌جای اندازه‌گیری همه پارامترها و حصول اطمینان ۱۰۰٪ از محدوده‌ی آن پارامتر بتوان با اندازه‌گیری یکی از آن پارامترها با اطمینان بالا (که البته کم‌تر از ۱۰۰٪ خواهد بود) درمورد مقدار چندین پارامتر دیگر قضاوت کرد. با این کار با از دست دادن مقدار ناچیزی از درصد اطمینان نتایج، به مقدار بسیار زیادی در زمان و هزینه‌ی کنترل کیفیت صرفه‌جویی می‌کنیم. گفتنی است چون در این حالت اقدامات اصلاحی سریع‌تری روی خط تولید اجرا می‌کنیم از بروز اشکالات احتمالی در محصول نیز سریع‌تر جلوگیری کرده‌ایم. روشن است که این اقدام به‌موقع می‌تواند از هزینه‌های اقلام نامنطبق جلوگیری کند و از این رو باعث افزایش سودآوری محصول شود.

استفاده از این نوع الگوریتم‌های کشف دانش به نتایجی منجر خواهد شد که حتی برای دست‌اندرکاران همان صنعت نیز جالب و بدیع بوده و می‌تواند به‌صورت کاربردی و عملی توسط آنها مورد استفاده قرارگیرد. الگوی بیان شده به‌طرز چشم‌گیری قابل گسترش به انواع دیگر خطوط کنترلی است.

برای پژوهش‌های آینده پیشنهاد می‌شود نقش پارامترهای نهایی مختصر شده، بر پارامترهای رنگ خروجی بررسی و ارزیابی شود تا بتوان تنها با ارزیابی و اندازه‌گیری چند پارامتر در خطوط رنگ، مانع از افت کیفیت نهایی رنگ قطعات شد.

پانوشت‌ها

1. electro deposition
2. data mining
3. statistical process control
4. decision tree
5. cost of quality (COQ)
6. electro deposition painting line
7. knowledge discovery and data mining
8. learning machine
9. data visualization
10. knowledge acquisition
11. predict
12. description
13. CRISP-DM
14. association rules (در متون فارسی نام این روش تحلیل وابستگی، تحلیل سبد بازار و یا قوانین انجمنی نیز ذکر گردیده است.)
15. undirected
16. support
17. confidence
18. apriori
19. pre-treat
20. degreasing
21. rinsing
22. activation
23. phosphating
24. ultra filter
25. setup
26. fire wall
27. ordinal
28. conductivity
29. antecedent
30. consequent

منابع (References)

1. Kim, S.B., Jitpitaklert, W., Park, S.K. and Hwang, S.J. "Data mining model-based control charts for multivariate and autocorrelated processes", *Journal of Expert Systems with Applications*, **39**, pp. 2073-2081 (2012).
2. Alzghoul, A. and Löfstrand, M. "Increasing availability of industrial systems through data stream mining", *Journal of Computers & Industrial Engineering*, **60**(2), pp. 195-205 (2011).
3. Köksal, G., Batmaz, I. and Testik, M.C. "A review of data mining applications for quality improvement in manufacturing industry", *Journal of Expert Systems with Applications*, **38**(10), pp. 13448-13467 (2011).
4. Kamsu-Foguem, B., Rigal, F. and Mauget, F. "Mining association rules for the quality improvement of the production process", *Journal of Expert Systems with Applications*, **40**, pp. 1034-1045 (2013).
5. Jabrouni, H., Kamsu-Foguem, B., Geneste, L. and Vaysse, C. "Continuous improvement through knowledge-guided analysis in experience feedback", *Journal of Engineering Applications of Artificial Intelligence*, **24**(8), pp. 1419-1431 (2011).
6. Abbasi, B. and niaki, S.T.A. "Monitoring high-yields processes with defects count in nonconforming items by artificial neural network", *Journal of Applied Mathematics and Computation*, **188**(1), pp. 262-270 (2007).
7. Zhang, C.-H., Di, L. and An, Z. "Welding quality monitoring and management system based on data mining technology", *International Conference on Machine Learning and Cybernetics*, **1**(2-5), pp. 13-17 (2003).
8. Alaeddini, A., Ghazanfari, M. and Amin Nayeri, M. "A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts", *Journal of Information Science*, **179**(11), pp. 1769-1784 (2009).
9. Evans, J.R. and Lindsay, W.M., *The Management and Control of Quality*, (5th ed., Cincinnati, OH: South-Western (2001).
10. Ghazanfari, M., Alizadeh, S. and Teimourpour, B., *Data Mining & Knowledge Discovery*, Tehran, Iran University of Science and Technology (2008)
11. Tan, P.N., Steinbach, M. and Kumar, V., *Introduction to Data Mining*, Boston, Pearson (2006).

12. Ahmadvand, A.M. and Akhondzade, E. "The usage framework of data mining techniques in modeling crimes", *Journal of Police Human Development*, **30**, pp. 11-23 (2010).

13. Han, J. and Kamber, M., *Data Mining: Concepts and Techniques*, Second Edition, Morgan Kaufman Publisher (2006),

14. Liao, S-H., Chu, P-H. and Hsiao, P-Y. "Data mining techniques and applications– A decade review from 2000 to 2011", *Journal of Expert Systems with Applications*, **39**(12), pp. 11303-11311 (2012).

15. Kantardzic, M., *Data Mining: Concepts, Models, Methods, and Algorithms*, 2nd Revised. Wiley-Blackwell (JOHN WILEY & SONS, INC. , Publication). ISBN-10:0470890452. (2011).

16. Berson, A., Smith, S. and Thearling, K., *Building Data Mining Applications for CRM*, McGraw Hill (2000).

17. Huang, T.C.-K., Liu, C.-C. and Chang, D.-C. "An empirical investigation of factors influencing the adoption of data mining tools", *International Journal of Information Management*, **32**(3), pp. 257–270 (2012).

18. Detlor, B. "Information management", *International Journal of Information Management*, **30**(2), pp. 103-108 (2010).

19. Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y. and Sun, X. "The application of data mining techniques in financial fraud detection: A classification framework and/ an academic review of literature", *Journal of Decision Support Systems*, **50**(3), pp. 559-569 (2011).

20. Agrawal, R., Imielinski, T. and Swami, A. "Mining associations between sets of items in large databases", *In ACM SIGMOD int'l Conf. on Management of Data*, Washington D.C., pp. 207-216 (May 1993).

21. Martínez-de-Pisón, F.J., Sanz, A., Martínez-de-Pisón, E., Jiménez, E. and Conti, D. "Mining association rules from time series to explain failures in a hot-dip galvanizing steel line", *Journal of Computers & Industrial Engineering*, **63**(1), pp. 22-36 (2012).