بررسی پارامترهای تأثیرگذار در ایجاد عیب چسبندگی ورق‌های فولادی تحت نورد سرد با کشف دانش از داده‌های فرآیندی

محمدتقی رضوان

استادیار گروه مهندسی صنایع، دانشکده مهندسی، دانشگاه کاشان، کاشان، ایران[1]

مصطفی ابویی اردکان

استادیار گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران

علی زینل همدانی

استاد دانشکده مهندسی صنایع، دانشگاه صنعتی اصفهان، اصفهان، ایران

علی اکبر باطنی

کارشناس شرکت فولاد مبارکه، اصفهان، ایران

دوفصلنامه علمی- پژوهشی

مدیریت اطلاعات

دوره ۳، شماره۱ – شماره پیاپی۴، بهار و تابستان۱۳۹۶

چکیده: بهبود کیفیت تولید با بهینه‌سازی شرایط تولید، یک هدف مستمر برای همه شرکت‌های صنعتی-تولیدی است. کاهش ضایعات و محصولات معیوبی که در اثر عوامل کنترلی در فرآیندهای تولید ایجاد می‌شوند، همواره از اهداف مدیران شرکت است. فرآیندهای تولید شرکت‌های فولادسازی نیز از این موضوع مستثنا نیستند. یکی از عیوب سطحی ورق‌های فولادی، عیب چسبندگی است که نوعی جوش‌خوردگی لایه‌ها بوده که در خط تولید در حین باز شدن کلاف آنیل شده و درصورتی‌که نیروی موردنیاز جهت باز کردن کلاف بیش از استحکام تسلیم باشد، اتفاق می‌افتد. در صورت بروز این عیب در ورق‌های فولادی، هزینه‌های سنگینی به شرکت‌های فولادسازی وارد می‌شود. در راستای کاهش عیب چسبندگی می‌توان داده‌های عملیاتی فرآیندی را تجزیه‌وتحلیل و پردازش کرد؛ به‌گونه‌ای که انعکاس‌دهنده مشخصه‌های تأثیرگذار بر عیب باشد. برای هدایت این کار می‌توان تکنیک‌های داده‌کاوی را مورد بهره‌برداری قرار داد؛ چراکه این تکنیک‌ها قادرند دانش را بازیابی کرده و به استخراج قواعد عملیاتی از یک مجموعه داده بپردازند. داده‌های عملیاتی از مجموعه داده‌های غیرمتمرکز یک شرکت فولادی گرفته شد و بر اساس متدولوژی استاندارد CRISP-DM مورد تجزیه‌وتحلیل قرار گرفت. مجموعه داده تحت بررسی یک مجموعه داده نامتوازن بوده که با استفاده از ابزارهای شبکه عصبی، رگرسیون لجستیک، ماشین‌های بردار پشتیبان و درخت تصمیم C5.0 مورد تجزیه‌وتحلیل قرار گرفت که بالاترین دقت، مربوط به درخت تصمیم C5.0 بوده است. بر این اساس، از درخت تصمیم، قوانینی استخراج شد که با تجارب کارشناسان خبره نیز چک و کنترل گردید. نتایج به‌دست‌آمده نشان داد که بهره‌گیری از داده‌کاوی برای تجزیه‌وتحلیل پارامترهای مؤثر بر عیوب می‌تواند منجر به بهبود کیفیت گردد؛ زیرا از این رویکرد می‌توان برای تنظیم پارامترهای عملیاتی فرآیندهای تولیدی استفاده نمود.

کلیدواژه‌ها: بازیابی دانش، داده‌کاوی، درخت تصمیم، شبکه عصبی، عیب چسبندگی، ماشین‌های بردار پشتیبان.

۱. نویسنده مسئول: rezvan@kashanu.ac.ir

مقدمه

در صنعت، روزانه حجم وسیعی از داده و اطلاعات در مورد پارامترهای مختلف فرآیندی تولید، جمع‌آوری و ثبت می‌شود. البته تولید، جمع‌آوری و ثبت داده، گاه دستی و گاه در قالب بانک اطلاعاتی، صورت می‌گیرد. از این داده‌های ثبت‌شده می‌توان به‌طور بالقوه‌ای برای پیش‌بینی سیر تکاملی متغیرهای جالب یا روندهای محیط بیرونی بهره گرفت. اما آن‌چنان‌که بایدوشاید از این پتانسیل بهره‌برداری کاملی نشده است. درواقع، دو مشکل اساسی در این راه وجود دارد که عبارت‌اند از: اطلاعات در سیستم‌های آرشیو مختلف، پراکنده هستند. به‌عبارت‌دیگر، این سیستم‌ها به یکدیگر متصل نبوده و ساختاری ناکارآمد از داده و اطلاعات در سازمان ایجاد نموده است که این پایگاه‌های داده غیرمتمرکز، توجه چندانی را جلب نمی‌نماید. همچنین آگاهی‌های لازم نسبت به ابزارهای داده‌کاوی و پتانسیل آن‌ها برای تشریح اطلاعات و تحلیل داده وجود ندارد.

با رشد فزاینده فناوری‌های جمع‌آوری و ثبت و نگهداری داده‌ها، استخراج دانش مفید پنهان در داده‌ها و عمل کردن بر اساس آن دانش، از اهمیت قابل‌توجهی برخوردار شده است. فرآیند کلی به‌کارگیری متدولوژی‌های کامپیوتر محور که دربردارنده روش‌هایی برای کشف دانش از داده است را داده‌کاوی می‌گویند. مجموعه تکنیک‌ها و روش‌های داده‌کاوی می‌تواند مدیران و تصمیم‌گیرندگان را در زمینه‌های مختلف یاری رساند. ازجمله آن‌ها، در کیفیت تولید می‌توان با بهره‌گیری از داده‌ها، پارامترهای مؤثر در ایجاد عیوب و زمینه‌های کاهش کیفیت را شناسایی کرده و حتی با تنظیم شرایط مناسب در راستای کاهش عیوب و افزایش کیفیت اقدامات اساسی را صورت داد یا در زمینه ایمنی و سلامت محیط کار، با بهره‌گیری از داده‌های ثبت و نگهداری شده می‌توان عوامل ریسک و خطر را شناسایی و اهمیت آن‌ها را مشخص نموده و در راستای کاهش ریسک حرکت کرد.

نورد سرد یکی از فرآیندهای کلیدی برخی از شرکت‌های فولادی در زمینه فلزکاری است که در آن، ضخامت ورق فولادی با عبور از بین دو غلتک کاهش می‌یابد. شرایط عملیاتی مختلف در این فرآیند می‌تواند بر کیفیت نهایی تولید ورق‌های فولادی تأثیر بگذارد. بنابراین شناخت شرایط تولید و عملیاتی که منجر به کاهش ضایعات و کالاهای معیوب و در سطح بالاتری افزایش بهره‌وری می‌شود، از اهمیت بالایی برخوردار است؛ به‌ویژه اینکه نیاز به انجام آزمایشی جدید و تحمیل هزینه جدیدی به آن شرکت نباشد. بدین ترتیب که بتوان با بهره‌گیری از داده‌های جمع‌آوری‌شده دانش ضمنی و پنهان در داده‌ها را استخراج و مورد بهره‌برداری قرار داد.

در نورد سرد، عیوب سطحی مختلفی در انواع مختلف ظاهر می‌شوند که عیب چسبندگی یکی از شناخته‌شده‌ترین عیوب سطحی است. صنایع خودرو و هوافضا در فرآیند تولید خود از ورق‌های فولادی تولیدشده با فرآیند نورد سرد استفاده می‌کنند. این صنایع، ورق‌هایی را که دارای عیوب سطحی هستند را رد می‌کنند و بنابراین این عیوب می‌تواند منجر به تحمیل زیان‌های بزرگی برای شرکت‌های فولادی شود. لذا این شرکت‌ها سعی می‌کنند کیفیت تولیدات نهایی خود را بهبود بخشند و به دنبال آن، بهره‌وری فرآیند را افزایش دهند. این تلاش‌ها می‌تواند در عمل مؤثر واقع شود و مقدار این عیب را کاهش دهد.

عیب چسبندگی یک تغییر پلاستیک است که به‌وسیله جوش فشاری سطوح بدون پوشش به وجود می‌آید. عیب چسبندگی در خط تمپر[1] در حین باز شدن کلاف آنیل‌شده و در هنگامی‌که نیروی موردنیاز جهت باز کردن کلاف، بیش از استحکام تسلیم باشد؛ اتفاق می‌افتد. به‌عبارت‌دیگر، چسبندگی به‌وسیله جوش فشاری سطوح بدون پوشش به وجود می‌آیند. تاکنون به‌طورقطع اثبات نشده است که این فرآیند جوش ناشی از جوش نفوذ[2] بین لایه‌ها، فرآیند سینترینگ[3] یا مکانیسم‌های دیگر چسبندگی است. چسبندگی عمومی به دو پارامتر استحکام تسلیم ورق‌آنیل شده و تنش اعمالی جهت بازنمودن کلاف بستگی دارد. درنتیجه، اگر جوش بین لایه‌های ورق با نیرویی کمتر از استحکام تسلیم ورق باز شود، موجب عیب چسبندگی نخواهد شد (Wendt et al. 2007).

اشکال مختلفی همچون ماهوری[4]، موضعی[5]، لبه‌ای[6]، عمومی[7] از این عیب قابل‌مشاهده است. چسبندگی ماهوری، در یک باند کاملاً محدود از عرض ورق اتفاق می‌افتد و علت اصلی آن، نیز ناهمگونی پروفیل ضخامت در عرض ورق است که فشار بالای شعاعی در بین لایه‌ها را باعث می‌شود. چسبندگی موضعی، در نقاط محلی خاص رخ می‌دهد و علت اصلی آن فشار بالای پینس‌هاست که موجب پیچش ناخواسته ورق در یک کویل[9] می‌شود. چسبندگی لبه‌ای، همان‌طوری که از نامش مشخص است، بر روی لبه‌ها اتفاق می‌افتد. این عیب به دلیل حرکت نسبی لایه‌های کلاف و کنوکتور[10] یا وجود اختلاف زیاد بین نرخ پخت و خنک کاری ایجاد می‌شود. آسیب‌دیدگی ناشی از جابجایی کویل به دلیل ناصافی سطوح درگیر پینس است. این عیب با استفاده از کنوکتورهای با منگنز بالا که سختی بالایی دارند، قابل کاهش است. چسبندگی عمومی، خطوط منحنی شکل نامنظمی هستند که داسی شکل بوده و پهنای ۵۰ تا ۴۰۰ میلی‌متر داشته و عمدتاً روی ناحیه بزرگی از وسط عرض ورق قرار دارند. آن‌ها همچنین به خطوط لودرز[11] باند نیز شباهت دارند (Wendt et al. 2007).

چسبندگی به دو پارامتر استحکام تسلیم ورق آنیل شده و تنش اعمالی جهت بازنمودن کلاف بستگی دارد که عوامل مختلفی می‌توانند بر روی این دو پارامتر اثرگذارند و ایجادکننده این عیب باشند. این عوامل علاوه بر فرآیند آنیلینگ[12] می‌تواند مربوط به فرآیندهای بالادست و پایین‌دست آن باشد. شناسایی عوامل مهم و مؤثر در ایجاد عیب چسبندگی، تعیین و تعریف دامنه مجاز برای این عوامل و کنترل آن‌ها می‌تواند منجر به کاهش این عیب گردد. ازآنجایی‌که این عوامل، اثرات متقابلی بر روی یکدیگر داشته

1 Temper Mill
2 Diffusion
3 Sintering
4 Ridge Sticker
5 Spot Sticker
6 Edge Sticker
7 General Sticker
8 Coil Tongs
9 Coil
10 Convector
11 Luders
12 Annealing

نمی‌توان به‌راحتی با کنترل یک عامل و ثابت نگه‌داشتن سایر عوامل، فرآیند تولید را تحت کنترل نگه داشت؛ بنابراین استفاده از روش‌های چندمتغیره بجای تک متغیره که فاقد فرضیات برای مدل‌سازی باشد می‌تواند کمک‌کننده باشد. داده‌کاوی یکی از زمینه‌هایی است که از آن می‌توان برای تحلیل‌های چندمتغیره استفاده نمود.

ازآنجاکه هر کلاف ورق دو حالت وجود یا عدم وجود عیب را خواهد داشت می‌توان گفت مسئله تحت بررسی، یک مسئله دسته‌بندی دودسته‌ای بوده که می‌توان برای تحلیل آن، از فرآیند داده‌کاوی یکپارچه استفاده نمود و به دنبال آن، دامنه‌ای مجاز برای عوامل ایجاد یا تشدیدکننده عیب تعیین و تعریف کرد. لذا استخراج قواعد مناسب دسته‌بندی برای تحلیل مسئله جهت تعیین دامنه مجاز برای عوامل شناسایی‌شده می‌تواند حائز اهمیت باشد.

هدف از این مقاله، استخراج دانش از داده‌های پایگاه‌های داده موجود در یک شرکت فولادی بزرگ بر روی عیب چسبندگی است تا بتوان با تنظیم پارامترهای فرآیندی تولید کلاف‌های ورق تحت فرآیند نورد سرد، این عیب را کاهش داد. برای انجام این کار، یک سیستم یکپارچه برای استخراج مجموعه‌ای از قواعد دسته‌بندی ارائه می‌شود تا کاربر نهایی بتواند از این قواعد استفاده کرده و اقداماتی را انجام داده تا این عیب در ورق‌های فولادی کاهش یابد. کیفیت قواعد استخراج‌شده بر اساس معیارهای ارزیابی شامل قابل‌درک بودن آن‌ها و قابل‌توجه بودن و همچنین دقت پیش‌بینی موردسنجش و ارزیابی قرار می‌گیرد.

ساختار مقاله بدین شرح است: پیشینه تحقیق و ادبیات موضوع مرتبط با توسعه مدل‌های دسته‌بندی برای تشخیص و پیش‌بینی عیب در فرآیندهای مختلف تولیدی و نیز انتخاب شرایط مطلوب عملکردی در بخش دوم مرور می‌شود. در بخش سوم، روش کار برای استفاده از متدولوژی استاندارد CRISP-DM برای این مسئله دسته‌بندی و استخراج قواعد مناسب معرفی می‌گردد. در این بخش جزئیات مسئله شامل ویژگی‌های مجموعه داده مورد تحلیل نیز تبیین می‌شود. بخش چهارم، نتایج به‌دست‌آمده از اجرای مدل‌های دسته‌بندی مختلف روی مجموعه داده تحت بررسی ارائه می‌گردد. نتیجه‌گیری کلی و پیشنهادها برای تحقیقات آتی، بخش پنجم را تشکیل می‌دهد.

پیشینه پژوهش

در سال‌های اخیر، به‌کارگیری داده‌کاوی در بسیاری از زمینه‌های تولید همانند کنترل کیفیت، به‌عنوان یک رویکرد جدیدی مطرح‌شده است. کارهای معمول داده‌کاوی عبارت‌اند از: دسته‌بندی[1]، خوشه‌بندی[2]، پیش‌بینی[3]، تحلیل انحراف[4]، قواعد التزام (وابستگی)[5]. بسیاری از مسائل تولیدی را می‌توان در قالب یک مسئله دسته‌بندی مطرح کرد (Choudhary et al. 2009). واضح است که عیوب، هزینه‌های تولید را

1 Classification
2 Clustering
3 Prediction
4 Deviation Analysis
5 Association Rules

افزایش می‌دهد و باید در جهت کاهش عیوب اقدامات اساسی صورت گیرد. به‌منظور کاهش عیوب باید عوامل مختلف ایجادکننده یا تشدیدکننده عیب را شناسایی و آن‌ها را کنترل نمود. استخراج الگوهای و قواعدی برای دستیابی به شرایط مناسب فرآیندهای تولید، همیشه موردتوجه محققان و کارشناسان صنعتی بوده است. در ادامه، برخی از تحقیقات در این زمینه را مرور می‌شود.

کوسیاک و کوراسک[1] یک روش داده‌کاوی مبتنی بر نظریه مجموعه سخت را برای شناسایی علت عیب لحیم‌کاری در یک تخته مدار چاپی استفاده کرد. قواعد استخراج‌شده از مجموعه داده مورد تحلیل، یک نشانه پایداری از اینکه محل لحیم بسیار باریک است فراهم می‌سازد و بدین ترتیب، بررسی‌های بیشتر برای علت عیوب توپ لحیم‌کاری را اثربخش‌تر می‌سازد (Kusiak and Kurasek 2001). یانگ[2] و همکاران، سیستم شبکه عصبی بهینه‌شده با الگوریتم ژنتیک را برای انتخاب مواد بهینه کامپوزیت و شرایط عملیاتی آن‌ها معرفی کرد (Yang et al. 2003). تسنگ[3] و همکاران، نظریه مجموعه سخت را برای استنتاج قواعد و شناسایی مهم‌ترین مشخصه‌ها برای برطرف کردن مشکلات کنترل کیفیت تخته‌های مدار چاپی (PCB[4]) توسعه دادند (Tseng et al. 2004). چن[5] و همکاران، قواعد التزام یا وابستگی را برای تشخیص عیب در تولیدات نیمه‌هادی کشف کردند. آن‌ها وابستگی بین ماشین‌های مختلف را نتیجه گرفته و اینکه با ترکیب آن‌ها با عیوب، ماشین ایجادکننده عیب را تشخیص دهند (Chen et al. 2005). وانگ[6] و همکاران، یک رویکرد مبتنی بر داده‌کاوی را با استفاده از خوشه‌بندی سلسله مراتبی با روش بخش‌بندی K-means پیشنهاد دادند که منتج شده به‌وسیله الگوریتم EM گوسی برای تشخیص و دسته‌بندی خودکار عیب بود (Wang et al. 2006). تسای[7] و همکاران، یک روش استنتاج مبتنی بر نمونه بکار گرفتند که از رویکردهای استنتاج و شاخص گذاری هوشمند برای توسعه یک سیستم پیش‌بینی عیب برای تخته مدار چاپی جدید (PCBs) استفاده کردند (Tsai et al. 2006). یزدچی[8] و همکاران، شبکه عصبی و سیستم استنتاج فازی را به‌عنوان یک ابزار دسته‌بندی برای تشخیص عیب و دسته‌بندی فولاد نورد سرد استفاده کردند. آن‌ها تصویر یک تصویر به‌دست‌آمده را با استفاده از یک روش کاهنده خاص افزایش دادند و موقعیت عیب را با استفاده از آنتروپی و مورفولوژی محلی یافتند (Yazdchi et al. 2008).

دنگ[9] و همکاران، یک مدل ترکیبی از شبکه عصبی مصنوعی و الگوریتم ژنتیک پیشنهاد دادند (Deng et al. 2009). ژانگ[10] و همکاران، ترکیبی از استنتاج مبتنی بر نمونه و نظریه مجموعه سخت

1. Kusiak &Kurasek
2. Yang
3. Tseng
4. Printed-Circuit Board (PCB)
5. Chen
6. Wang
7. Tsai
8. Yazdchi
9. Deng
10. Zhang

برای بهینه‌سازی پارامترهای فرآیند آسیاب کردن میل بادامک استفاده نمودند (Zhang et al. 2013). والاوانیس و کسموپولس[۱] برای تشخیص و دسته‌بندی عیب در رادیوگرافی جوش، سه ابزار دسته‌بندی متفاوت شامل شبکه عیبی، ماشین بردار پشتیبان و نزدیک‌ترین K همسایه را بکار گرفتند (Valavanis and Kosmopoulos 2010). صدیقی و افشاری[۲] شبکه عصبی و الگوریتم ژنتیک را برای توصیف روابط غیرخطی گوناگون بین پارامترهای سنگ‌زنی توصیف کردند (Sedighi and Afshari 2010).

بااین‌وجود، مطالعه منتشرشده برای فرآیند نورد سرد برای کنترل عیب چسبندگی محدود به تحقیقات زیر است: پارالایکز[۳] و همکاران، روش‌های نیمه تجربی را برای پیش‌بینی کرنش در عملیات شکل‌گیری کلاف توسعه دادند. این داده‌ها، با داده‌های یک واحد صنعتی تأیید و همبستگی خوبی را نشان دادند. این مدل‌ها، یک بینش برای انتقال‌های مختلف و تأثیر آن‌ها بر روی کرنش‌ها ارائه کردند (Paralikas et al. 2009). زاریت و دیس[۴] با استفاده از روش FCANN[۵]، دانش را از شبکه‌های عصبی استخراج و ارائه کردند. این رویکرد، برای درک رابطه میان پارامترهای فرآیند از طریق قواعد ضمنی در شرایط مختلف عملیاتی بر روی منحنی بار از فرآیند نورد سرد، قابل‌اجرا است (Za´rate and Dias 2009). آگروال و شیوپوری[۶] یک روش جدید برای پیش‌بینی عیوب سطحی در تولیدات میله‌ای ارائه کردند و سپس یک سیستم خودکار به نام تشخیص‌هایی روی نورد غلتکی توسعه دادند که نتایج آن با بازرسی دستی بسیار نزدیک بود (Agarwal and Shivpuri 2013). پرز[۷] و همکاران، از یک رویکرد کاهش ابعاد برای تحلیل بصری یک فرآیند نورد سرد استفاده کردند. روش پیشنهادی آن‌ها، روابط بین رفتارهای پویا و خطای پراکندگی در فرآیند نورد را پیدا می‌نماید (Pérez et al. 2013).

در این مقاله، مسئله عیب چسبندگی در کلاف‌های فولادی نورد سرد را به‌عنوان یک مسئله دسته‌بندی تعریف می‌شود تا با استخراج قواعد دسته‌بندی شرایط عملیاتی مطلوب برای فرآیند تعیین شود.

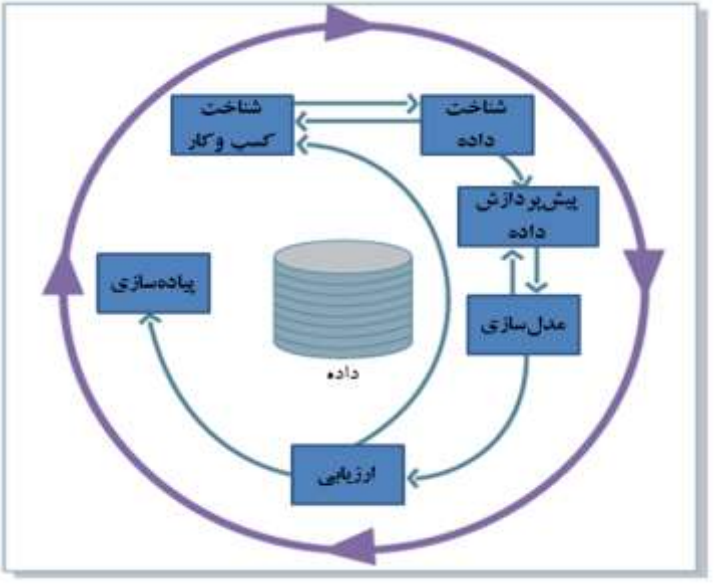
روش‌شناسی پژوهش

در این بخش، ضمن مرور متداول‌ترین فرآیند داده‌کاوی یعنی فرآیند استاندارد صنعتی داده‌کاوی[۸] (CRISP-DM) به‌منظور هدایت اصولی تحلیل‌های داده‌کاوی، ابعاد مسئله تحت بررسی از دیدگاه این فرآیند تبیین می‌گردد. این فرآیند در نسخه نرم‌افزار کلمنتاین[۹] گنجانده‌شده است.

---

1. Valavanis & Kosmopoulos
2. Sedighi & Afshari
3. Paralikas
4. Za´rate & Dias
5. FCANN (Formal Concept Artificial Neural Network)
6. Agarwal & Shivpuri
7. Pérez
8. Cross Industry Standard Process for Data Mining (CRISP-DM)
9. Clementine

این فرآیند را می‌توان در قالب شکل ۱ نمایش داد که شامل مراحل شش‌گانه‌ای است که در ذیل بدان اشاره می‌گردد.

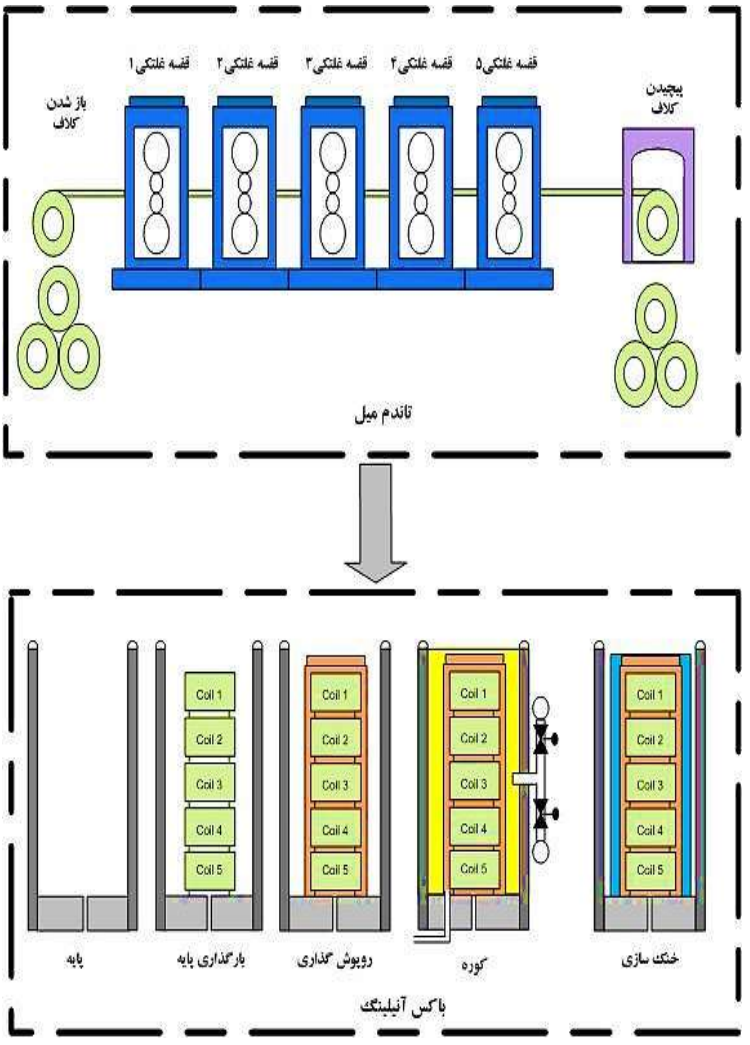


شکل ۱. فرآیند استاندارد صنعتی داده‌کاوی

الف) شناخت کسب‌وکار: شناخت کسب‌وکار شامل تعیین اهداف تجاری، ارزیابی موقعیت‌ها، شکل‌دهی اهداف داده‌کاوی و توسعه دادن برنامه پروژه است.

بخشی از محصولات نورد گرم فرایندهای تولید محصولات فولادی، برای تولید محصولات باکیفیت بالاتر به نورد سرد جهت انجام و تکمیل عملیات ارسال‌شده تا درنهایت محصول مورد مصرف در اتومبیل‌سازی و ازاین‌دست را تولید نماید. درواقع، فرآیند نورد سرد برای تولید محصولات نورد با سطح پرداخت عالی و تلورانس ضخامت مشخصی استفاده می‌شود. در نورد سرد، تغییر شکل با نیروی اعمالی از سوی غلتک‌ها بر روی ورق بازشده، صورت می‌گیرد. فرآیند نورد سرد شامل دو زیر فرآیند است که تحت عنوان تاندم‌میل[1] و باز پخت شناخته می‌شوند. نمودار چیدمان فرآیند نورد سرد در شکل ۲ نمایش داده‌شده است.

1Tandem Mill

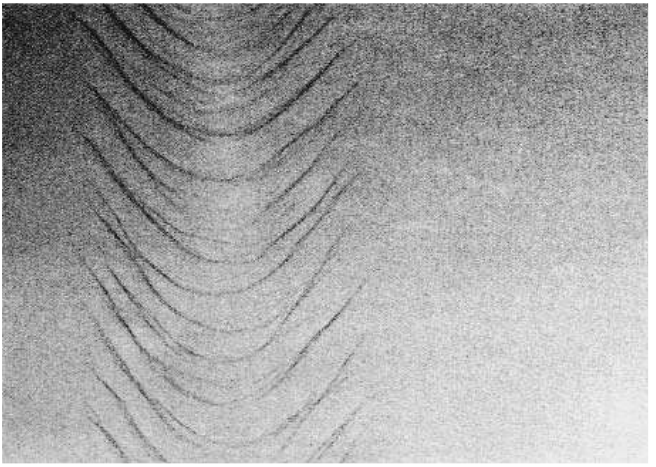


شکل ۲. نموداری از فرآیند نورد سرد

در تاندم‌میل، ورق‌های فولادی از قفسه‌های غلتکی عبور می‌کنند، به‌طوری‌که ورق‌های باضخامت مطلوب تولید شود. در شروع خط، هر کلاف به‌وسیله بازکننده کلاف بازمی‌گردد و سپس از ۵ قفسه غلتکی عبور می‌کند. هر قفسه، یک جفت غلتک با ابعاد دو فوت در قطر دارد که توسط موتورهای الکتریکی قدرتمندی به حرکت درمی‌آیند و با تنگنا قرار دادن فولاد، ضخامت را کاهش می‌دهند. در پایان خط، ورق به‌صورت کلاف پیچیده می‌شود. بعد از تاندم میل، ورق بایستی باز پخت شود. این مرحله یعنی باز پخت جعبه‌ای، فرآیند گرمایش، تابش متعادل و خنک‌سازی به‌منظور به دست آوردن خواص ویژه مواد است. در این مرحله، پنج کلاف ورق را در یک کوره مسدود روی یک‌پایه گذاشته به‌طوری‌که گاز H2 و یا HNX در

کوره دمیده می‌شود تا در مقابل هوا از آن محافظت شود. روی این کوره یک کلاهک گذاشته می‌شود تا فرآیند باز پخت انجام گردد. در این مرحله، باز پخت آرام و تبلور مجدد صورت می‌گیرد. درواقع، باز پخت آرام رسوب‌های سانتی‌متری و حالات آتشفشانی را که استحکام فولاد را کاهش می‌دهد، کمینه می‌نماید. تبلور مجدد باز پخت، فرم‌های بلورین را در حالت شکل‌دهی قبل از سرد شدن، تجدید ساخت می‌نماید. درنهایت، مواد به‌دست‌آمده سطح تمیز و همواری به‌منظور صنایع اتومبیل‌سازی دارند.

در برخی از موارد بعد از مرحله باز پخت، جدایش زیرسطحی از جوش لایه‌ای رخ می‌دهد که تحت عنوان عیب چسبندگی شناخته می‌شود که تجزیه‌وتحلیل این عیب و پارامترهای مؤثر بر آن و تنظیم این پارامترها برای کاهش عیب از اهداف این مقاله به شمار می‌رود. تصویری از عیب چسبندگی را می‌توان در شکل ۳ مشاهده نمود.



شکل ۳. یک نمونه از عیب چسبندگی

درگذشته علت اصلی عیب چسبندگی، کوره‌های باز پخت، فرض می‌شد؛ چون این عیب در پایان مرحله باز پخت مشاهده می‌شود؛ اما باید گفت که عوامل مختلف دیگری می‌توانند ایجادکننده عیب چسبندگی باشند. شرکت‌هایی در دنیا وجود دارند که بیش از ۱۰ درصد محصولات آن‌ها دارای این عیب است. با انجام اقدامات اصلاحی در خطوط نورد می‌توان این عیب را تا یک درصد کاهش داد.

بنابراین باوجوداینکه عیب چسبندگی در مرحله باز پخت مشاهده می‌شوند، اما عوامل مؤثر بر شکل‌دهی این عیب را می‌توان علاوه بر باز پخت، در مراحل قبل و بعد از باز پخت هم جستجو کرد. در مقیاس بالای تولید، حدود پنج درصد عیب چسبندگی، زیان‌های هنگفتی را به شرکت‌های تولیدکننده فولاد تحمیل می‌نماید؛ بنابراین سعی و تلاش برای کاهش این عیب همواره وجود دارد. بررسی عوامل ایجادکننده این عیب به‌تنهایی و یا حتی زوجی عوامل، نتایج چندان مناسبی را ایجاد نکرده است؛ زیرا عوامل ایجادکننده عیب چسبندگی، به پارامترهای مختلف فرآیندی مربوط می‌شود که بعضی از آن‌ها اثر

معکوس بر روی یکدیگر دارند؛ بنابراین به‌کارگیری رویکرد یکپارچه‌ای که قادر باشد، هم‌زمان عوامل بالقوه مختلف شکل‌دهی عیب چسبندگی را بررسی کند، مفید است.

این عوامل را می‌توان تحت عنوان مشخصه‌های شرطی یک مسئله دسته‌بندی تعریف کرد که یک دسته کلاف‌های بدون عیب و دسته دیگر کلاف‌های دارای عیب چسبندگی باشند. مطالعه این مجموعه داده باهدف پاسخ‌گویی به این سؤال صورت می‌گیرد که چگونه می‌توان عیب چسبندگی را بر روی ورق‌های نورد سرد کاهش داد. برای پاسخ‌دهی به این سؤال، باید بررسی نمود چه مشخصه‌هایی با چه مقادیری در ایجاد این عیب مؤثر هستند. به‌عبارت‌دیگر اهمیت مشخصه‌ها را در ایجاد عیب می‌بایست بررسی نمود. ایجاد یک مدل دسته‌بندی که فقط دقت بالایی داشته و بتواند یک نمونه را پیش‌بینی نماید برای تحلیل مسئله تحت بررسی کفایت نمی‌کند و لازم است علاوه بر تعیین اهمیت مشخصه‌ها، قواعد دسته‌بندی استخراج گردد تا بر اساس آن بررسی نمود که آیا می‌توان برای مشخصه‌های فرآیندی محدوده‌ای تنظیم نمود که عیب به حداقل مقدار ممکن برسد یا نه.

ب) شناخت داده: با شکل‌گیری اهداف کسب‌وکار و برنامه پروژه، مرحله بعدی، شناخت داده با توجه به نیازمندی‌های اطلاعاتی است. این مرحله، شامل جمع‌آوری اولیه اطلاعات، توصیف، شناسایی و تأیید کیفیت اطلاعات است. شناسایی اطلاعات مانند مشاهده آماره‌های خلاصه‌شده که مشاهده بصری متغیرهای قطعی و قیاسی را در بردارد، می‌تواند نقطه انتهایی این مرحله باشد. در طول این مرحله، مدل‌هایی همچون تحلیل خوشه‌ای را به‌منظور تعیین الگوهای داده می‌توان بکار برد.

برای اینکه بتوان عیب چسبندگی را مورد تجزیه‌وتحلیل قرار داد، نیاز است تا پارامترهایی که می‌توانند نقشی در ایجاد عیب داشته باشند تعیین کرد. بدین منظور، فهرستی از پارامترهای مختلف تعیین گردید که از یک کلاف به کلاف دیگر تغییر می‌کرد. این پارامترها، بر اساس کار تحقیقاتی توسط وندت و همکاران در سال ۲۰۰۷ و همچنین با نظر کارشناسان آن شرکت فولادی تعیین گردید. درواقع، این پارامترها به‌عنوان مشخصه‌های شرطی و «وجود یا عدم وجود عیب چسبندگی» به‌عنوان مشخصه تصمیم مسئله دسته‌بندی در نظر گرفته شد. ویژگی‌های این مشخصه‌ها در جدول ۱، قابل‌مشاهده است. بر اساس این جدول، هفت مشخصه شرطی عددی[1] و چهارمشخصه شرطی رسته‌ای[2] هستند.

جدول ۱. فهرست مشخصه‌های مجموعه داده عیب چسبندگی

ردیف	نام مشخصه	نام لاتین	نوع مشخصه	واحد	دامنه
۱	کشش کلاف‌پیچ تاندم	Tension	رسته‌ای	-	صفر برای $۲۵N/mm^2$ یا یک برای $۲۲N/mm^2$

1Numerical Feature
2Categorical Feature

ردیف	نام مشخصه	نام لاتین	نوع مشخصه	واحد	دامنه
۲	گریدهای فولاد	CRMQLY	رسته‌ای	-	F1R, F3R, F5R, G4R, J1R, X1R, X1RY, X2ER, X2R, X2MI, X3R, X3MI, X5R, X7R, X8R, X4ER, X4MI
۳	موقعیت ورق در کوره	PositionInCharge	رسته‌ای	-	B, C_1, C_2, C_3, T
۴	گاز مورداستفاده در کوره	Gas	رسته‌ای	-	HNX, H_2
۵	زبری	Roughness	عددی	cm^{-1}	[۱۸ ،۹۹]
۶	تمیزی	Cleaning	عددی	mg/m^2	[۳۰ ،۸۹]
۷	عرض ورق	Width	عددی	mm	[۶۱۵ ،۱۶۱۰]
۸	ضخامت ورق	Thickness	عددی	mm	[۰/۲۲ ،۳]
۹	قطر بیرونی کلاف ورق	CoilDiameter	عددی	mm	[۹۴۴ ،۲۰۹۲]
۱۰	وزن شارژ	ActualWeightOfCharge	عددی	ton	[۱۲۲۳۵۰،۱۹۲۱۰]
۱۱	زمان حرارت دهی عملی	ActualHeatingTime	عددی	hour	[۱۲/۱۳۳، ۶۲/۱۳۳]
۱۲	مشخصه عیب	Non-Defect/Defect	رسته‌ای	-	۰ یا ۱

مشخصه‌های دیگری نیز وجود داشتند که اهمیت زیادی برای مطالعه داشتند، اما به دلیل ثابت نگه‌داشته شدن آن‌ها برای همه ورق‌های تولیدی، کنار گذاشته شدند که ازاین‌دست می‌توان به‌سرعت باز شدن ورق، دمای حرارت دهی[1] و زمان آغاز خنک کاری اشاره نمود. مشخصه‌هایی همچون شیفت و تیم کاری نیز باوجود اطلاعات موجود حذف شدند؛ زیرا بر اساس نظر کارشناسان شرکت، عملکرد تیم‌های مختلف در شیفت‌های متفاوت یکسان است. همچنین مقادیر برخی از مشخصه‌ها همچون پروفیل نورد گرم به دلیل ضعف در سیستم اطلاعاتی برای ورق‌های مختلف موجود نبودند که این مشخصه‌ها هم از مجموعه مشخصه‌ها کنار گذاشته شدند.

ج) پیش‌پردازش داده: بعد از شناسایی منابع اطلاعاتی موجود، نیاز است که منابع، انتخاب و پالایش شوند و به فرم دلخواه و قالب‌دار تبدیل گردند. در این مرحله، پالایش و تغییر شکل داده‌ها برای مدل

1Soaking Temperature

کردن آن‌ها صورت می‌گیرد. همچنین شناسایی داده‌ها، با دقت بیشتری انجام‌شده و مدل‌های بیشتری استفاده‌شده تا مجدداً فرصت مشاهده الگوهای شکل‌گرفته بر پایه شناخت کسب‌وکار فراهم گردد.

داده‌های موردبررسی برای کلاف‌های مختلف نورد سرد، در یک پایگاه داده متمرکز قرار نداشتند، به‌طوری‌که یک پایگاه داده اکثر مشخصه‌ها به‌جز مشخصه‌های سختی و تمیزی را دارا است؛ اما مقادیر مشخصه‌های سختی و تمیزی برای برخی از کلاف‌ها در یک پایگاه داده دیگر ذخیره می‌شود و علت این امر، آن است که به دلیل اندازه‌گیری دستی و نمونه‌ای این دو مشخصه، هنوز امکان یکپارچه‌سازی در این شرکت فولادی فراهم نشده است؛ بنابراین یکپارچه‌سازی این دو پایگاه برای این مطالعه با استفاده از مشخصه کلیدی شماره اصلی کلاف انجام می‌شود.

آنچه باید بدان توجه کرد اینکه مقدار مشخصه‌های سختی و تمیزی برای برخی از کلاف‌ها اندازه‌گیری شده و بنابراین برای کلاف‌های زیادی این مقادیر وجود ندارد. لذا پالایش مجموعه داده ایجادشده برای انجام مطالعه ضروری است. بر همین اساس، کلاف‌هایی که مقادیر این دو مشخصه برای آن‌ها وجود نداشت، کنار گذاشته شد. همچنین کلاف‌هایی که بیش از نیمی از مشخصه‌های باقیمانده آن‌ها بدون مقدار بودند، نیز از تجزیه‌وتحلیل کنار گذاشته شد.

داده‌ها در یک دوره نه‌ماهه جمع‌آوری گردید و از تعداد ۱۶۲۰۷۹ کلاف تولیدشده در واحد نورد سرد، فقط ۱۹۵۲ کلاف در مجموعه داده وجود داشت که با مطالب گفته‌شده، تطابق داشت. بر اساس آنچه گفته شد، مجموعه داده تحت مطالعه دارای ۱۲ مشخصه (۱۱ مشخصه شرطی و یک مشخصه دسته) و ۱۹۵۲ رکورد است که تعداد رکوردهای دارای عیب آن ۱۴۳ و مابقی کلاف‌های بدون عیب هستند.

برای آشنایی بیشتر با مجموعه داده برخی از آماره‌های شناخته‌شده برای مشخصه‌های عددی در جدول ۲ مرور می‌گردد. بر اساس اطلاعات این جدول می‌توان گفت فقط برخی از مقادیر رسته‌ای مرتبط با گریدهای فولاد است که در آن عیب دیده نمی‌شود و موقعیت C_4 نیز در بین موقعیت دارای ورق‌های بدون عیب است. هرچند تعداد ورق‌های بدون عیب در این موقعیت، پایین است که این به دلیل ضعف سیستم اطلاعاتی بوده که اطلاعات کلاف‌های حاوی این موقعیت جزو مقادیر گمشده بوده و از مجموعه داده کنار گذاشته شدند.

د) مدل‌سازی: برای تحلیل اولیه، نرم‌افزارهای داده‌کاوی با ترسیم داده‌ها و بررسی ارتباطات آن‌ها مفید است. هنگامی‌که شناخت بیشتری در مورد داده‌ها حاصل شد، می‌توان مدل‌های جزئی بیشتری را متناسب با نوع داده بکار برد. همچنین در این مرحله، تقسیم اطلاعات به مجموعه‌های آموزشی و آزمایشی برای مدل‌سازی موردتوجه قرارگرفته و اعتبار مدل ساخته‌شده مورد ارزیابی قرار می‌گیرد.

جدول ۲. آماره‌های مشخصه‌های عددی مجموعه داده تحت بررسی

ردیف	نام مشخصه عددی	میانگین	کمینه	بیشینه	دامنه	انحراف استاندارد	میانه	مد
۱	زبری	۵۰/۵۵۵	۱۸	۹۹	۸۱	۱۵/۸۵۵	۴۸	۴۱
۲	تمیزی	۵۹/۳۹۴	۳۰	۸۹	۵۹	۶/۷۵۰	۶۰	۶۰
۳	عرض ورق	۱۰۹۴/۹۱۵	۶۱۵	۱۶۱۰	۹۹۵	۱۷۴/۴۲۹	۱۰۱۵	۱۰۰۰
۴	ضخامت ورق	۰/۹۹۴	۰/۲۲	۳/۰	۲/۷۸	۰/۳۶۹	۰/۸	۰/۷
۵	قطر بیرونی کلاف ورق	۱۶۹۵/۸۷۶	۹۴۴	۲۰۹۲	۱۰۹۸	۹۶/۵۹۴	۱۶۸۸	۱۶۹۴
۶	وزن شارژ	۶۴۴۷۰/۷۶۹	۱۹۲۱۰	۱۲۲۳۵۰	۱۰۳۱۴۰	۱۳۳۵۸/۳۰۵	۶۰۸۲۵	۵۳۹۲۰
۷	زمان حرارت دهی عملی	۲۲/۳۶۸	۱۲/۱۳۳	۶۲/۱۳۳	۵۰	۵/۱۲	۲۱/۸۰	۲۱/۱۰

در این قسمت، مدل‌های دسته‌بندی شناخته‌شده همچون درخت تصمیم C5.0، رگرسیون لجستیک، شبکه‌های عصبی و ماشین بردار پشتیبان بر روی مجموعه داده عیب چسبندگی مورد ارزیابی قرار می‌گیرند. برای این ارزیابی، مجموعه داده به‌طور تصادفی به دو مجموعه آموزش و آزمون با نسبت‌های ۸۰ و ۲۰ تقسیم‌شده که هر آزمایش پنج بار تکرار شده است. اجرای مدل‌های دسته‌بندی شناخته‌شده با استفاده از نرم‌افزار SPSS Clementine صورت می‌پذیرد. مدل‌های دسته‌بندی مورداستفاده در ادامه به‌طور اجمالی معرفی می‌شوند.

درخت تصمیم[1] C5.0: هر مدل درخت تصمیم، از مجموعه‌ای از قوانین، برای تقسیم یک جمعیت ناهمگن و وسیع به گروه‌های کوچک‌تر و همگن‌تر بر اساس یک متغیر هدف خاص تشکیل‌شده است. در هر تقسیم‌بندی متوالی، اعضای مجموعه‌های حاصل بیش‌ازپیش به یکدیگر شباهت پیدا می‌کنند. الگوریتم C5.0، جدیدترین نسخه الگوریتم‌های درخت تصمیم است. الگوریتم C5.0، ابتدا یک درخت بزرگ‌تر از

1. Decision Tree

حد لازم را ایجاد می‌کند و سپس آن را برای ایجاد یک مدل باثبات‌تر، هرس می‌نماید. این الگوریتم، برای انتخاب از میان درخت‌های فرعی نامزد از یک مجموعه اعتبار استفاده نکرده بلکه از همان داده‌های استفاده‌شده برای تصمیم‌گیری درباره چگونگی هرس درخت، استفاده می‌کند. این الگوریتم، درختان را با بررسی نرخ خطای آن‌ها در هر گره و فرض اینکه نرخ واقعی خطا درواقع بدتر از این است، هرس می‌نماید.

شبکه‌های عصبی مصنوعی[۸]: یک شبکه عصبی مصنوعی، ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته‌شده و مانند مغز به پردازش اطلاعات می‌پردازد. شبکه‌های عصبی، باقابلیت قابل‌توجه آن‌ها در استنتاج معانی از داده‌های پیچیده یا مبهم، می‌تواند برای استخراج الگوها و شناسایی روش‌هایی که آگاهی از آن‌ها برای انسان و دیگر روش‌های کامپیوتری بسیار پیچیده و دشوار است؛ به کار گرفته شود. از شبکه‌های عصبی برای پیش‌بینی، دسته‌بندی و خوشه‌بندی استفاده می‌کنند (Olsen and Delen 2008).

رگرسیون لجستیک[۱]: رگرسیون لجستیک یک مدل آماری رگرسیون برای متغیرهای وابسته رسته‌ای است. این مدل را می‌توان به‌عنوان مدل خطی تعمیم‌یافته‌ای که از تابع لوجیت به‌عنوان تابع پیوند استفاده می‌کند و خطایش از توزیع چندجمله‌ای پیروی می‌کند، به‌حساب آورد. در این مدل، مقادیر نسبت احتمالات برای هرکدام از متغیرهای توصیفی که رابطه معنی‌دار آماری با احتمال موفقیت ناشی از متغیر وابسته داشته باشند، محاسبه می‌شود (Olsen and Delen 2008).

ماشین‌های بردار پشتیبان[۲]: ماشین‌های بردار پشتیبان، مجموعه‌ای از روش‌های یادگیری نظارتی[۳] هستند که به‌منظور رگرسیون و دسته‌بندی مورداستفاده قرار می‌گیرند. به عبارت ساده‌تر، در مجموعه‌ای از داده‌های یادگیری که هر نمونه به یکی از دودسته تعلق دارد، یادگیری ماشین بردار پشتیبان می‌تواند پیش‌بینی کند که یک داده جدید به کدام دسته تعلق دارد. در سال ۱۹۶۳ ماشین‌های بردار پشتیبان به‌منظور دسته‌بندی دودویی داده‌ها معرفی شدند. در این مسئله هدف جدا کردن دودسته به کمک تابعی است که از نمونه‌ها استنتاج می‌شود. این جداکننده باید به‌گونه‌ای باشد که کارایی لازم را به‌منظور پیش‌بینی مناسب داده‌های جدید داشته باشد. در ماشین‌های بردار پشتیبان، پیش‌بینی به معنای دسته‌بندی نظارتی است که شامل دو گام می‌شود: در اولین گام، ماشین بردار پشتیبان با تعدادی از داده‌های موجود در یک مجموعه داده، به‌عنوان یک ابزار دسته‌بندی آموزش داده می‌شود و در دومین گام، از این ابزار دسته‌بندی آموزش‌یافته، به‌منظور دسته‌بندی بقیه داده‌ها استفاده می‌گردد (Olsen and Delen 2008).

هـ) ارزیابی: با توجه به اهداف تجاری شکل‌گرفته در مرحله اول یعنی شناخت کسب‌وکار، نتایج مدل ارزیابی می‌شود. این ارزیابی منجر به شناسایی نیازهای دیگر می‌گردد. در این مرحله گاهی لازم است به‌طور مکرر به مراحل قبلی فرآیند CRISP-DM، بازگشت داشت. به دست آوردن شناخت کسب‌وکار در

1 Logistic Regression
2 Support Vector Machines
3 Supervised Learning

داده‌کاوی، یک فرآیند تکراری است تا اینکه نتایج روش‌های مختلف ترسیمی، آماری و ابزارهای هوش مصنوعی به کاربر، روابط جدیدی را ارائه و شناخت عمیق‌تری از عملیات سازمان فراهم نماید.

در مسائل دسته‌بندی، منبع اولیه ارزیابی‌های عملکرد، ماتریس انطباق یا تطابق[1]است که ماتریس دسته‌بندی نامیده می‌شود. جدول ۳ یک ماتریس انطباق را برای یک مسئله دسته‌بندی با دودسته صفر و یک را نشان می‌دهد.

جدول ۳. یک ماتریس انطباق ساده

		دسته درست	
		مثبت	منفی
دسته پیش‌بینی‌شده	مثبت	شمارش مثبت درست (TP)	شمارش مثبت غلط (FP)
	منفی	شمارش منفی غلط (FN)	شمارش منفی درست (TN)

اعداد در امتداد قطر، از بالاترین در سمت چپ تا پایین‌ترین در سمت راست، تصمیمات صحیح اتخاذشده را نشان می‌دهد و اعداد خارج از این قطر، خطاها را بازنمایی می‌کند. نرخ مثبت درست[2]یک مدل دسته‌بندی، با تقسیم مثبت‌های درست دسته‌بندی‌شده (تعداد مثبت‌های درست) بر کل تعداد مثبت‌ها برآورده می‌گردد. دقت[3]یک مدل دسته‌بندی با تقسیم تعداد مثبت‌ها و منفی‌های درست دسته‌بندی‌شده بر تعداد کل نمونه‌ها به دست می‌آید (Olsen and Delen 2008).

دقت معمول‌ترین معیار ارزیابی مدل‌های دسته‌بندی به شمار می‌رود که بر اساس نحوه تقسیم‌بندی مجموعه داده و تکرار فرآیند یادگیری، عموماً به‌صورت متوسط بیان می‌شود.

فرمول ۱.

$$\text{True Positive Rate} = \frac{TP}{TP + FN}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

و) پیاده‌سازی: داده‌کاوی می‌تواند در بازبینی فرضیات گذشته و یا برای کشف دانش یعنی شناسایی روابط مفید و غیرمنتظره مورداستفاده قرار گیرد. می‌توان از طریق دانش کسب‌شده در مراحل نخستین فرآیند CRISP-DM، مدل‌های با کمترین عیب را به دست آورد که ممکن است برای اهداف زیادی در عملیات کسب‌وکار شامل پیش‌بینی و یا تعیین موقعیت‌های کلیدی استفاده گردد. این مدل‌ها باید برای

1Coincidence Matrix
2True Positive Rate
3Accuracy

تغییر شرایط عملیاتی، آگاهی‌دهنده باشند، چون ممکن است که آنچه امروز درست است، سال آینده درست نباشد. اگر تغییرات مهمی اتفاق بیفتد، مدل‌سازی باید دوباره صورت پذیرد. همچنین عاقلانه به نظر می‌رسد که نتایج پروژه داده‌کاوی را به‌عنوان اسنادی نگهداری کرد تا به‌عنوان شواهدی برای مطالعات آینده موجود باشند. در حقیقت این نتایج، دانش پنهانی است که از داده‌های بانک‌های اطلاعاتی با بهره‌گیری از داده‌کاوی استخراج‌شده و می‌تواند کمک‌کننده در تصمیم‌گیری باشد و در چرخه مدیریت دانش، در تولید و استفاده مدنظر قرار گیرد.

برای صحه‌گذاری بر قواعد استخراج‌شده از مجموعه داده اصلی و همچنین بررسی قابلیت پیاده‌سازی آن‌ها، نتایج برای کارشناسان آن شرکت فولاد ارائه گردید و بر اساس صحبت‌های انجام‌شده، مشخص گردید که برخی از قواعد در سیستم اعمال‌شده تا توسط خبره خارجی به آن‌ها ارائه‌شده است.

یافته‌های پژوهش

نتایج به‌دست‌آمده از اجرای مدل‌های دسته‌بندی موردبررسی شامل درخت تصمیم C5.0، رگرسیون لجستیک، شبکه‌های عصبی و ماشین بردار پشتیبان، در جدول ۴ برای مجموعه داده اصلی نشان داده‌شده است. این نتایج شامل متوسط و انحراف معیار برای دقت کلی، متوسط و انحراف معیار دقت برای دسته‌های تعریف‌شده یعنی عیب‌دار (دسته یک) و بدون عیب (دسته صفر) هستند.

علت نمایش دقت برای دودسته، این است که به دلیل نامتوازن بودن مجموعه داده اصلی اکثر مدل‌ها توانایی جزئی در تشخیص دسته اقلیت دارند. این موضوع را ستون پنجم جدول ۳، تأیید می‌نمایید. با توجه به اینکه دسته اقلیت حدود هشت درصد مجموعه داده اصلی را شکل می‌دهد، برخی از مدل‌ها همچون شبکه عصبی فقط داده را در دسته بدون عیب، دسته‌بندی می‌نماید و به همین دلیل متوسط و انحراف معیار دقت برای دسته یک، صفر است. این نتیجه به‌هیچ‌وجه نمی‌تواند مطلوب باشد؛ زیرا عملاً دسته‌بندی انجام نمی‌شود و بر اساس اینکه فراوانی دسته اکثریت، بالاتر است، همه رکوردها، صفر پیش‌بینی می‌گردد.

جدول ۴. نتایج مدل‌های دسته‌بندی بر روی مجموعه داده اصلی

ردیف	نام مدل دسته‌بندی	متوسط (انحراف معیار) دقت دسته صفر	متوسط (انحراف معیار) دقت دسته یک	متوسط (انحراف معیار) دقت
۱	رگرسیون لجستیک	۹۸/۹۴ (۰/۳۷)	۱۵/۹۷ (۶/۲)	۹۲/۶۹ (۱/۶۹)
۲	شبکه‌های عصبی	۱۰۰/۰ (۰/۰)	۰/۰ (۰/۰)	۹۲/۶۴ (۰/۶۴)
۳	ماشین بردار پشتیبان	۹۸/۶۵ (۰/۶۵)	۱۶/۷۳ (۴/۷۳)	۹۲/۶۳ (۱/۰۸)
۴	درخت C5.0	۹۸/۶۰ (۱/۳۷)	۲۶/۵۶ (۱۹/۵۴)	۹۳/۲۴ (۲/۱۸)

در میان مدل‌های مورداستفاده برای مجموعه داده اصلی، درخت تصمیم C5.0 ضمن ایجاد بالاترین متوسط دقت کلی، بیشترین متوسط دقت را برای دسته اقلیت داشته که این موضوع نشان‌دهنده عملکرد مناسب این مدل است. همچنین با این مدل می‌توان قواعد دسته‌بندی را استخراج کرد که از یک‌سو کاربران درک خوبی نسبت به آن‌ها دارند و از سوی دیگر می‌توان با به‌کارگیری این قواعد، پارامترهای از خط تولید که در دسترس مهندسین تولید هستند را در جهت کاهش عیب تنظیم نمود. به‌عبارت‌دیگر، ممکن است برخی از مشخصه‌های مجموعه داده اعم از گرید فولاد، عرض و ضخامت ورق عموماً جزو مشخصه‌های تعریفی توسط مشتری باشند و در تنظیم آن‌ها مهندسین تولید نقشی ندارند، اما می‌توان پارامترهایی همچون زبری، تمیزی و گاز مورداستفاده در کوره را بر اساس مشخصه‌های تعریف‌شده توسط مشتری طوری تنظیم نمود که عیب چسبندگی کاهش یابد. درخت تصمیم C5.0 به دلیل عملکرد مناسب ازنظر متوسط دقت کلی و متوسط دقت برای دسته اقلیت در مجموعه داده اصلی را می‌توان برای تحلیل و تصمیم‌گیری مورداستفاده قرار داد. این نتایج شامل تعیین اهمیت مشخصه‌ها و مجموعه قواعد تولیدشده با استفاده از درخت تصمیم C5.0 است.

مجموعه قواعد استخراج‌شده از مجموعه داده اصلی با استفاده از درخت C5.0 در جدول ۵ گنجانده‌شده است. بر اساس این جدول باید گفت که ۲۱ قاعده استخراج‌شده است که از این میان، نه قاعده مربوط به دسته یک یعنی ورق‌های عیب‌دار و مابقی مربوط به دسته صفر یعنی ورق‌های بدون عیب است. در مقابل هر قاعده یک عبارت پرانتزی وجود دارد که شامل دو بخش است که بخش اول تعداد رکوردهای پوشش داده‌شده و بخش دوم دقت قاعده را نمایش می‌دهد. دو قاعده با شماره‌های ۱۱ و ۱۵ وجود دارد که هیچ رکوردی را پوشش نمی‌دهند که یکی از آن‌ها برای پیش‌بینی دسته یک و دیگری برای پیش‌بینی دسته صفر است و با توجه به این مجموعه داده، عملاً ارزشی ندارند؛ ولی به دلیل ساختار درخت، ایجادشده‌اند؛ بنابراین برای این دو قاعده، پرانتز مقابل آن به‌صورت (۰) است. دقت کلی برای مجموعه داده اصلی بر اساس این قواعد، ۹۴/۱۶ بوده که برای دسته اقلیت این مقدار به ۲۷/۹۷ می‌رسد. تعداد قواعد با طول شش، بیشترین فراوانی را با هفت قاعده دارد و تعداد قواعد با طول یک و دو، هرکدام یک است. متوسط طول قواعد این مجموعه، ۴/۷۲ است.

برای تعمیم نتایج به‌دست‌آمده بر اساس مجموعه قواعد، می‌بایست اطلاعات رکوردهایی از کلاف‌هایی که مشخصه‌های آن‌ها به‌طور کامل پرشده‌اند را به مجموعه داده اصلی اضافه نمود. همچنین قبل از اضافه نمودن رکوردها به مجموعه داده می‌توان دقت مجموعه قواعد استخراج‌شده را مورد ارزیابی قرار داد. بر اساس مجموعه قواعد به‌دست‌آمده از مجموعه داده اصلی می‌توان موارد زیر را بیان کرد که با استفاده از آن‌ها در راستای کاهش عیب چسبندگی، گام برداشت. هرچند برای عملیاتی ساختن این موارد، باید محدودیت‌های فنی را در نظر گرفت.

الف) بر اساس مجموعه قواعد و با توجه به داده‌های پالایش‌شده تغییر کشش از ۲۵N/mm2 به ۲۲N/mm2 تأثیری در ایجاد یا عدم ایجاد عیب چسبندگی نداشته است، هرچند، داده‌های

جمع‌آوری‌شده به‌گونه‌ای است که ورق‌های با مقدار دوم کشش، بسیار بیشتر از مقدار اول کشش بوده است.

ب) با توجه به‌قاعده یک، گاز HNX در مقایسه با گاز H2 ورق‌های دارای عیب بسیار کمی تولید می‌نماید؛ به‌طوری‌که این قاعده تعداد زیادی از رکوردها را با دقت بالا در دسته صفر (بدون عیب) قرار می‌دهد. اگر مشکلی در تأمین گاز HNX وجود نداشته باشد، پیشنهاد می‌شود از این گاز در کوره‌ها استفاده گردد.

جدول ۵. مجموعه قواعد استخراج‌شده از مجموعه داده اصلی به‌وسیله ابزار C5.0

شماره	قاعده
Rule#1	If Gas = HNX then non-defect (816; 0.974)
Rule#2	If Gas = H2 and Roughness > 40 then non-defect (517; 0.93)
Rule#3	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F1R" "F5R" "X1R" "X2ER" "X2MI" "X3MI" "X5R" "X7R"] then non-defect (90; 0.878)
Rule#4	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness <= 0.750 and PositionInCharge in ["B"] and ActualHeatingTime <= 31.730 then defect (3; 1.0)
Rule#5	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness <= 0.750 and PositionInCharge in ["B"] and ActualHeatingTime > 31.730 then non-defect (4; 0.75)
Rule#6	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness <= 0.750 and PositionInCharge in ["C1" "C2" "C3"] then non-defect (7; 0.857)
Rule#7	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness <= 0.750 and PositionInCharge in ["T"] and ActualHeatingTime <= 31.170 then non-defect (4; 0.75)
Rule#8	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness <= 0.750 and PositionInCharge in ["T"] and ActualHeatingTime > 31.170 then defect (4; 1.0)
Rule#9	If Gas = H2 and Roughness <= 40 and CRMQLY in ["F3R"] and Thickness > 0.750 then non-defect (3; 1.0)
Rule#10	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X2R"] then defect (7; 0.714)
Rule#11	If Gas = H2 and Roughness <= 40 and CRMQLY in ["G4R" "J1R" "X1RY" "X4ER" "X4MI"] then non-defect (0)
Rule#12	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["B" "C1" "C3"] then defect (16; 0.625)
Rule#13	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width

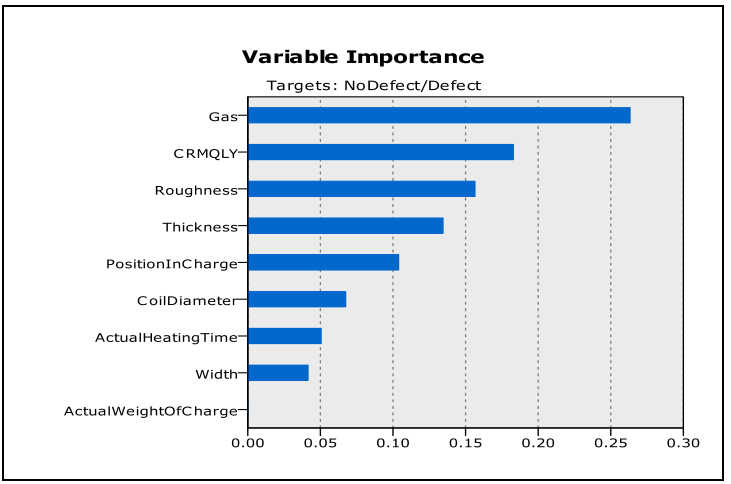
شماره	قاعده
	<= 1,250 and PositionInCharge in ["C2"] and ActualHeatingTime <= 24.220 then non-defect (3; 1.0)
Rule#14	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["C2"] and ActualHeatingTime > 24.220 then defect (3; 1.0)
Rule#15	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["C4"] then defect (0)
Rule#16	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["T"] and Thickness <= 0.670 then defect (3; 1.0)
Rule#17	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["T"] and Thickness > 0.670 and ActualWeightOfCharge <= 66,965 then defect (3; 1.0)
Rule#18	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width <= 1,250 and PositionInCharge in ["T"] and Thickness > 0.670 and ActualWeightOfCharge > 66,965 then non-defect (11; 0.818)
Rule#19	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X3R"] and Width > 1,250 then non-defect (17; 1.0)
Rule#20	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X8R"] and CoilDiameter <= 1,630 then defect (4; 1.0)
Rule#21	If Gas = H2 and Roughness <= 40 and CRMQLY in ["X8R"] and CoilDiameter > 1,630 then non-defect (11; 0.818)

ج) درصورتی‌که گاز H2 مورداستفاده قرار می‌گیرد، می‌توان بر اساس قواعد استخراج‌شده طوری عمل نمود که وجود عیب در ورق‌ها کاهش یابد.

- ✓ با توجه به‌قاعده دو، اگر زبری ورق‌ها بالاتر از ۴۰ باشد با احتمال بالایی، عیب چسبندگی وجود ندارد. پوشش تعداد ۵۱۷ کلاف با دقت ۹۳ درصد در دسته بدون عیب، کارایی این قاعده را نمایان می‌سازد.
- ✓ با توجه به‌قاعده ۱۲، اگر گرید فولاد X3R بوده و عرض آن کمتر از ۱۲۵۰ باشد باید زبری کمتر از ۴۰ داشته و همچنین دریکی از موقعیت‌های «B»، «C1» و «C3» باشد، با احتمال خوبی، این عیب دیده نمی‌شود. این قاعده، ۱۶ کلاف را با دقت ۶۲/۵ درصد در دسته بدون عیب قرار می‌دهد. هرچند دقت این قاعده خیلی مناسب نیست.
- ✓ با توجه به‌قاعده ۱۹، اگر گرید فولاد X3R است و عرض آن بیشتر از ۱۲۵۰ باشد باید زبری کمتر از ۴۰ باشد تا با احتمال خوبی، عیب دیده نمی‌شود. این قاعده، ۱۷ کلاف را با دقت یک در دسته بدون عیب قرار می‌دهد.

- ✓ با توجه به‌قاعده ۲۱، درصورتی‌که گرید فولاد X8R است اگر زبری کمتر از ۴۰ و قطر خارجی کلاف بیشتر از ۱۶۳۰ باشد، در این صورت با احتمال خوبی عیب دیده نمی‌شود. پوشش ۱۱ کلاف با دقت ۸۱/۸ درصد می‌توان گواهی بر این عبارت باشد.

می‌توان اهمیت مشخصه‌های مجموعه داده عیب چسبندگی بر اساس ابزارهای مختلف دسته‌بندی و ازجمله درخت تصمیم C5.0 تعیین کرد. علت تمرکز بر درخت تصمیم C5.0، عملکرد بهتر آن نسبت به سایر ابزارهاست. اهمیت مشخصه‌ها بر اساس درخت C5.0 در شکل ۴ نمایش داده‌شده است. بر اساس این شکل می‌توان گفت که مشخصه‌های گاز مورداستفاده، گرید فولاد، زبری و ضخامت دارای بیشترین اهمیت هستند. همچنین به نظر می‌رسد که وزن شارژ و تمیزی دو مشخصه‌ای هستند که نقش ناچیزی در ایجاد عیب دارند. همچنین باید گفت بر اساس سایر ابزارها که عملکرد منطقی دارند، مشخصه‌های وزن شارژ و تمیزی، مشخصه‌های کم‌اهمیتی قلمداد شده و سایر ابزارها مشخصه‌هایی که درخت تصمیم C5.0 مهم تلقی کرده را با تغییر ناچیزی از مقدار اهمیت، مهم تشخیص داده‌اند.



شکل ۴. اهمیت مشخصه‌های مجموعه داده اصلی با به‌کارگیری درخت تصمیم C5.0

با توجه به پایین بودن متوسط دقت پیش‌بینی در دسته یک، می‌توان گفت که این احتمال وجود دارد که مشخصه‌های دیگری در ایجاد عیب مؤثر هستند که با ثبت اطلاعات آن‌ها می‌توان به بررسی میزان تأثیر آن‌ها پرداخت. یکی از این مشخصه‌ها پروفیل نورد گرم است که در حال حاضر داده‌ای برای آن در سیستم اطلاعاتی شرکت ثبت نمی‌شود. علاوه بر این، تعداد رکوردهای مجموعه داده اصلی با جمع‌آوری اطلاعات مرتبط با رکوردهای دیگر باید افزایش یابد تا امکان افزایش متوسط دقت پیش‌بینی برای دسته یک و درنهایت متوسط دقت کلی ابزار دسته‌بندی، فراهم شود.

برای صحه‌گذاری بر قواعد استخراج‌شده از مجموعه داده اصلی و همچنین بررسی قابلیت پیاده‌سازی آن‌ها، نتایج برای کارشناسان آن شرکت فولاد ارائه گردید و بر اساس صحبت‌های انجام‌شده، مشخص گردید که قاعده یک به‌صورت یک سیاست مدیریتی برای ورق‌های بازرگانی اعمال‌شده است؛ بدین معنی

که شرکت این‌گونه ورق‌ها را در کوره‌های با گاز HNX شارژ نموده است که این کار سبب کاهش قابل‌توجه عیب بر روی این ورق‌ها شده است. همچنین ورق‌هایی که عملیات باز پخت آن‌ها در کوره‌های با گاز H2 انجام می‌شود، می‌بایست دارای زبری بالاتر از ۴۰ باشند که این دانش توسط یک خبره خارجی به آن واحد ارائه‌شده است.

نتیجه‌گیری و پیشنهادها

امروزه یکی از عوامل کلیدی متمایزکننده سازمان‌های هوشمند از یکدیگر شناخت دقیق دانش و اطلاعات و همچنین بازیابی و بهره‌گیری از آن است. خلق، تأیید، ارائه، توزیع و به‌کارگیری دانش و اطلاعات می‌تواند منجر به بهبود بهره‌وری و خدمت‌رسانی بهتر به ذی‌نفعان سازمان و هماهنگی بیشتر با تغییرات باشد. مدیریت دانش، ابزار مهمی است که سازمان‌ها با استفاده از آن می‌توانند مدیریت کارآمدتری بر اطلاعات و در سطحی بالاتر و به‌صورت ساختاریافته بر دانش سازمانی داشته باشند. چرخه مدیریت دانش شامل کشف و تولید، ارزیابی، اشاعه و تسهیم دانش و همچنین به‌کارگیری دانش است. داده‌های جمع‌آوری و نگهداری شده در بانک‌های اطلاعاتی شرکت‌ها می‌تواند یکی از منابع کشف و تولید دانش باشد. درواقع می‌توان با داده‌کاوی دانش پنهان موجود در داده‌ها را کشف نمود و مورد بهره‌برداری قرار داد. برخی از سازمان برای پیاده‌سازی مدیریت دانش اقدام به سرمایه‌گذاری گسترده‌ای در زمینه فناوری اطلاعات و ارتباطات کرده‌اند؛ این در حالی است که در پاره‌ای از موارد، از پایگاه‌های داده خود غافل شده و استفاده‌ای از دانش ضمنی داده‌های خود نکرده‌اند. داده‌هایی که به‌تدریج و به مقاصدی همچون تحت کنترل بودن فرآیندهای کاری جمع‌آوری‌شده‌اند و اکنون می‌توان از آن‌ها برای مقاصد دیگر نیز بهره برد.

در این مقاله، یک کاربرد مهم از داده‌کاوی در زمینه بهبود کیفیت موردبحث و بررسی قرار گرفت. بدین ترتیب که مسئله عیب چسبندگی ورق‌های فولادی تحت فرآیند نورد سرد، به‌صورت یک مسئله دسته‌بندی تعریف گردید و اکثر داده‌های مرتبط با آن از پایگاه‌های داده شرکت گرفته شد. هرچند برخی از آن‌ها نیز به‌صورت دستی جمع‌آوری گردید که درنهایت با مشخصه کلیدی یعنی شماره اصلی کلاف ورق، تطابقی صورت گرفت تا در یک بازه نه‌ماهه مجموعه داده‌ای یکپارچه جهت انجام تحلیل‌های لازم تهیه شود. پالایش این مجموعه داده با حذف رکوردهایی که تعداد زیادی از مشخصه‌های آن‌ها خالی بود صورت گرفت و درنهایت یک مجموعه داده جهت انجام تحلیل‌های لازم ایجاد گردید.

درنهایت، این مجموعه داده با مدل‌های مختلفی دسته‌بندی همچون رگرسیون لجستیک، شبکه‌های عصبی، ماشین بردار پشتیبان و درخت تصمیم C5.0 موردبررسی قرار گرفت. مقایسه مدل‌های مختلف بر روی مجموعه داده‌ها نشان داد که درخت تصمیم C5.0 عملکرد بهتری نسبت به سایرین دارد. این ابزار ضمن داشتن متوسط دقت بالا، استخراج‌کننده قواعدی است که می‌توان از آن‌ها برای تنظیم خط تولید به‌عنوان راهبرد استفاده نمود. این قواعد استخراج‌شده را می‌توان برای تنظیم پارامترهای قابل‌کنترل برای فرآیند تولید مورداستفاده قرار داد.

جهت استفاده بیشتر از نتایج، به مجموعه داده اصلی می‌بایست مشخصه‌هایی همچون پروفیل نورد گرم، زمان سرد شدن و سرعت باز شدن ورق‌ها اضافه گردد، به‌شرط آنکه اطلاعات آن‌ها جمع‌آوری شود و همچنین مجموعه داده اصلی شامل تعداد رکوردهای بیشتری جهت تعمیم نتایج باشد.

فهرست منابع

Agarwal, K., and·Shivpuri, R. 2013. "On line prediction of surface defects in hot bar rolling based on Bayesian hierarchical modeling". *Journal of Intelligent Manufacturing* 26(4): 785-800.

Chen, W. C., Tseng, S. S., and Wang, C. Y. 2005. "A Novel manufacturing defect detection method using association rule mining techniques". *Expert System with Applications* 29: 807-815.

Choudhary, A. K., Tiwari, M. K., and Harding, J. A. 2009. "Data Mining in Manufacturing: A Review Based on the Kind of Knowledge". *Journal of Intelligent Manufacturing* 20(5): 501-521.

Deng, Z. H., Zhang, X. H., Liu, W., and Cao, H. 2009. "A hybrid model using genetic algorithm and neural network for process parameters optimization in NC camshaft grinding". *International Journal of Advanced Manufacturing Technology* 45(9-10): 859–866.

Kusiak, A. and Kurasek, C., 2001." Data Mining of Printed Circuit Board defects". *IEEE Transactions on Robotics and Automation* 17(2):191-196.

Olsen, D.L.and Delen, D., 2008. "Advanced data mining techniques", Springer.

Paralikas, J., Salonitis, K., and Chryssolouris, G. 2009. "Optimization of the roll forming process parameters—a semi empirical approach". *International Journal of Advanced Manufacturing Technology* 47(9–12): 1041–1052.

Pérez, D., García-Fernández, F.J., Díaz, I., Cuadrado A.A., Ordonez, D.G., Díez, A.B., and Domínguez, M. 2013. "Visual analysis of a cold rolling process using a dimensionality reduction approach". *Engineering Applications of Artificial Intelligence* 26: 1865–1871.

Sedighi, M., & Afshari, D. 2010. "Creep feed grinding optimization by an integrated GA-NN system". *Journal of Intelligent Manufacturing* 21(6): 657–663.

Tsai, C. Y., Chiu, C. C., and Chen, J. S. 2006. "A Case based reasoning system for PCB defect prediction". *Expert Systems with Applications* 28: 813-822.

Tseng, T. L., Jothishanker, M. C., and Wu, T. 2004. Quality Control Problem in Printed Circuit Board Manufacturing–An Extended Rough Set Theory Approach", *Journal of Manufacturing System* 23(1): 56-72.

Valavanis, I., and Kosmopoulos, D., 2010. "Multiclass defect detection and classification in weld radiographic images using geometric and texture features". *Expert Systems with Applications* 37(12): 7606-7614.

Wang, C. H., Kuo, W., and Bensmail, H., 2006. "Detection and classification of defects patterns on semiconductor wafers". *IIE Transactions* 38: 1059-1068.

Wendt, P., Frech, W., and Leifgen, U. 2007. "Cold rolling defect, "stickers" and countermeasures". *Heat processing* 5(2): 127-135.

Yang, S. Y., Tansel, I. N., and Kropas-Hughes, C. V. 2003. "Selection of optimal material and operating conditions in composite manufacturing. Part I: Computational tool". *International Journal of Machine Tools & Manufacture* 43(2): 169–173.

Yazdchi, M. R., Golibagh Mahyari, A., and Nazeri, A. 2008. "Detection and Classification of Surface Defects of Cold Rolling Mill Steel Using Morphology and Neural Network". *International Conference on Computational Intelligence for Modelling Control & Automation, IEEE*.

Za´rate, L.E., and Dias, S.M. 2009. "Qualitative behavior rules for the cold rolling process extracted from trained ANN via the FCANN method". *Engineering Applications of Artificial Intelligence* 22: 718–731.

Zhang, X.H., Deng, Z.H., Liu, W., and Cao, H. 2013. "Combining rough set and case based reasoning for process conditions selection in camshaft grinding". *Journal of Intelligent Manufacturing* 24: 211–224.

Investigation of the effective parameters causing stickers in cold rolling mill steel by knowledge discovery from process data

Mohammad Taghi Rezvan
Assistant Professor, Department of Industrial Engineering, Faculty of Engineering, University of Kashan, Kashan, Iran[1]

Mostafa Abouei Ardakan
Assistant Professor, Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran

Ali Zeinal Hamadani
Professor, Department of Industrial Engineering, Isfahan University of Technology, Isfahan, Iran

Ali Akbar Bateni
Expert of Mobarakeh Steel Company, Isfahan, Iran

Abstract: Improving the quality of products by optimizing production procedures is a continuous goal for all industrial-manufacturing companies. Reducing waste and defective products through controlling factors in production processes is always a goal of managers' steel companies. One of the surface defects appears in steel coils is sticker issue, which is a kind of layers welding that occurs in the production line during the opening of the annealed coil when the force required to open the coil is greater than the Surrender Strength. To reduce the stickers' issue, operational data of the production process of cold rolling analyzed and presented in a fashion that reflects the effective features create this defect. For this purpose, data mining techniques can be exploited, because these techniques can retrieve knowledge and extract operational rules from a dataset. The dataset under examine was analyzed using neural network, logistic regression, support vector machines, and decision trees. Experimental results reveal the highest accuracy is related to decision tree C5.0. Accordingly, from this decision tree, rules were extracted which were checked and controlled by the experts. The results showed that the use of data mining to analyze the parameters affecting sticker defects can lead to improved quality, because this approach can be used to adjust the operational parameters of the manufacturing processes.

Keywords: Data mining Sticking defect, Decision tree, Knowledge discovery, Logistic regression, Neural network, Support vector machine.

1Corresponding Author: rezvan@kashanu.ac.ir