

شناسایی الگوهای غالب در وقوع حوادث کار با استفاده از داده کاوی به طریق تحلیل قوانین همبستگی

چکیده

سابقه و هدف: در کنار اثرات نامطلوب وقوع حوادث کار، فرصت یادگیری برای سازمان در جهت پیشگیری از وقوع حوادث مشابه و یا کاهش اثرات حوادث بسیار حائز اهمیت است. هدف این تحقیق معرفی چگونگی استفاده از یک روش یادگیری مبتنی بر پایگاه داده حوادث در یک کارخانه تولید کننده قطعات فولادی است.

مواد و روش ها: در این تحقیق از داده کاوی به طریق تحلیل قوانین همبستگی به عنوان یک رویکرد شناسایی الگو در پایگاه داده برای شناسایی الگوهای شایع حوادث استفاده می شود. روش Apriori در محیط نرم افزار IBM SPSS Modeler 18 برای این منظور استفاده می گردد.

یافته ها: چهار قانون از الگوهای شایع در وقوع حوادث در کارخانه مورد مطالعه حاصل گردید که بیانگر وجود روابط قابل اتکا بین وقوع اتفاقاتی چون برخورد با فلزات، فرو رفتن جسم و سقوط اجسام با صدماتی از نوع بریدگی و کوفتگی می باشد. محوطه دپو موقت نیز کانونی برای وقوع حوادث شناسایی گردید.

نتیجه گیری: با کنار هم قرار دادن قوانین استخراج شده و شواهد میدانی، عدم تناسب تجهیزات حفاظت فردی در بخش پوشش، وجود رویه های موقتی در عملیات کارخانه بدون تعریف پروتوکل های ایمنی مناسب و سهل انگاری در پیروی از رویه های ایمنی جاری از مهمترین علل وقوع حوادث پر تکرار در این واحد صنعتی مورد شناسایی قرار گرفتند.

واژگان کلیدی: تجزیه و تحلیل حوادث، یادگیری، داده کاوی، تحلیل قوانین همبستگی

Identifying dominant occupational accident patterns using data mining via association rules analysis

Abstract

Background and Objective: Along with the adverse effects of workplace accidents, learning opportunities for the organization to prevent similar accidents or reduce the impact of accidents are crucial. The objective of this research is to introduce how to use a learning approach based on accident database in a steelmaking plant .

Materials and Methods: In this research, data mining through the association rules analysis is used as a pattern recognition approach in the database to identify common patterns of accidents. The Apriori method is used in the IBM SPSS Modeler 18 software environment for this purpose.

Results: Four rules were derived from common patterns in the factory accidents. They indicate the existence of important relationships between events such as metal collisions, falling objects whit injuries include cuts and sprains.

Conclusion: Putting together extracted rules and field evidence, inadequate personal protective equipment in the clothing, temporary works in factory operations without defining appropriate safety protocols, negligence in following safety procedures were the most important identified causes of current accidents this plant.

Keywords: Accident Analysis, Learning, Data Mining, Association Rules Analysis

تهیه گزارش حوادث در محیط های کاری بنابر دلایل مختلفی صورت می پذیرد که از جمله مهمترین آنها ایجاد زمینه ای برای شناسایی عوامل حادثه و مقصران احتمالی جهت طی مراحل قانونی جبران خسارت است. اما در صورت وجود سازکارهای مناسب در یک سازمان، نوعی یادگیری می تواند در سازمان اتفاق افتد که بر مبنای آن از وقوع حوادث دیگر در سازمان پیشگیری به عمل آید یا اثرات حادثه به حداقل ممکن محدود گردد. به بیان دیگر در کنار اثرات نامطلوب حوادث کاری برای سازمان ها، یک دست آورد جنبی مثبت، فرصتی است که حوادث برای یادگیری در سیستم های مدیریت ایمنی ایجاد می نماید. شیوه های یادگیری کلاسیک را می توان از یک منظر به دو دسته تقسیم کرد. نخست شامل تمام روش های تجزیه و تحلیل حوادث می شود که در آن هر پرونده حادثه بصورت مجزا تشکیل و بر مبنای آن کشف علل حادثه در لایه های مختلف صورت گرفته و در نهایت پیشنهاداتی برای جلوگیری از وقوع حوادث مشابه ارائه می گردد. دسته دوم از روش های یادگیری به منظور پیشگیری از حوادث، مطالعه گزارش های حوادث در کنار یکدیگر و نیل به کشف یک یا چند الگوی حادثه می باشد. مدل های مفهومی مختلفی از تعدادی از این الگو های حوادث نیز در گذشته ارائه شده است [۱]. با این حال این الگوهای بیان شده بصورت مدل بسیار کلی بوده و استفاده از آنها در عمل نیازمند بومی سازی مدل مورد نظر در هر سیستم یا زیر سیستم است. اما کامپیوتری شدن فرایندهای بررسی حوادث و تهیه گزارشات از یک سو و توسعه روش های خودکار یادگیری ماشینی مبتنی بر علم آمار، هوش مصنوعی و بازیابی اطلاعات از سوی دیگر موجب پیدایش یک نوع مدل سازی جدید از الگو های رفتاری حاکم بر پدیده ها از جمله حوادث شده است [۱۳]. داده کاوی (Data Mining) کلید واژه ای است که عموماً برای این شیوه جدید شناخت مورد استفاده قرار می گیرد. داده کاوی که خود فرایندی چند مرحله جهت تبدیل داده خام انباشته در پایگاه های داده به دانش مفید می باشد، از طریق رویکرد های تحلیل داده همچون خوشه بندی (Clustering)، دسته بندی (Classification)، پیش بینی (Forecasting) و تحلیل قوانین همبستگی (Association Rules Analysis) به اجرا در می آید. می توان گفت تمامی این رویکرد های تحلیل داده قادر به پاسخگویی سؤالاتی چون "عوامل موثر بر یک پدیده چه هستند؟"، "چه عامل یا عواملی بیشتر تاثیر را در بروز یک رفتار خاص سیستم دارند؟"، "سیستم تحت یک شرایط مشخص چه رفتار نشان خواهد داد؟" و "چه شرایطی باید محقق شود تا سیستم رفتار ویژه ای را انجام دهد؟".

لذا استفاده از داده کاوی در تجزیه و تحلیل حوادث می تواند پاسخ سؤالات بسیار مهم را آشکار سازد. با این حال به خاطر وابستگی عملیات داده کاوی به وجود یک پایگاه داده مناسب، لازم است ابتدا چنین پایگاه داده ای ایجاد گردد. موضوع خوشحال کننده در این بخش آن است که زمینه ایجاد چنین پایگاه داده ای در اغلب سازمان که گزارشات دقیق از حوادث

تهیه می کنند مهیا می باشد. در واقع با کنار هم گذاشتن و دسته بندی اطلاعات حوادث به شکل هوشمندانه می توان به راحتی به چنین پایگاه داده ارزشمندی دست یافت.

در سال های اخیر استفاده از داده کاوی در تجزیه و تحلیل حوادث به منظور یادگیری در سازمان مورد توجه محققین و مدیران صنایع تولیدی و خدماتی قرار گرفته است. به عنوان مثال استفاده از داده کاوی در تجزیه و تحلیل حوادث جاده ای بسیار متداول است [۲]، [۳]، [۴]، [۵] و [۱۲]. جنتی و همکاران پارامترهای تأثیرگذار بر علل رخداد حوادث شغلی صنعت ساخت و ساز با استفاده از روش تلفیقی Fuzzy logic-AHP و نظریه پارتو در استان کردستان ارزیابی کرده اند. در این مطالعه پارامترهای مؤثر در رخداد حوادث صنعت ساخت و ساز شناسایی و پرسش نامه ای بر مبنای تحلیل سلسله مراتبی فازی طراحی و توزیع شده که در ارزیابی این مطالعه، از مجموعه ی هشت فاکتور مورد نظر، شرایط نایمن، تأثیرگذارترین فاکتور شناخته شده است. همچنین، از زیرفاکتورهای این شاخص، دو پارامتر فقدان داربست و تجهیزات ایمن کار در ارتفاع و فقدان نرده های حفاظتی در پرتگاه های موجود، دارای بیشترین اهمیت بودند. همچنین توجه به پارامترهای تعیین شده توسط نظریه پارتو در هر فاکتور، می تواند به سهولت بیش از ۸۰ درصد از علل رخداد حوادث شغلی این صنعت را کاهش بدهد [۱۴]. تجزیه و تحلیل حوادث کار نیز به شیوه مشابه توسط داده کاوی صورت پذیرفته است. به عنوان مثال چنگ و همکاران در سال ۲۰۱۲ به استخراج الگو های موجود در حوادث صنعت ساختمان تایوان از طریق روش درخت تصمیم در داده کاوی پرداخته اند [۶]. به شکل مشابه سانمیکول و همکاران در سال ۲۰۱۸ به تجزیه و تحلیل حوادث کار با استفاده از تجزیه و تحلیل قوانین همبستگی در معادن اسپانیا پرداخته اند [۷]. همچنین پیش بینی حوادث نیز پس از شناسایی الگوها امکانپذیر است [۸] و [۱۱]. با بررسی تحقیقات در این زمینه می توان دریافت تمامی قابلیت های داده کاوی در حوزه تجزیه و تحلیل حوادث نیز قابل استفاده است. اما نکته قابل توجه در استفاده از داده کاوی در تجزیه و تحلیل حوادث آن است که چون نتایج حاصله بر مبنای داده های جمع آوری شده در یک حوزه خاص صنعتی و تحت شرایط محیطی، فنی و فرهنگی همان حوزه است، مطالعات بیشتر برای حصول نتایج تحقیقاتی دقیق در صنایع و جغرافیای مختلف مورد نیاز است.

با توجه به آمار نسبی بالای حوادث در کشور های در حال توسعه مانند ایران [۹] و نیز حادثه خیزی برخی صنایع گسترش یافته مانند فولاد در کشورمان، لزوم بکارگیری بیش از پیش روش های نوین تجزیه و تحلیل حوادث برای ایجاد یادگیری به منظور کاهش حوادث ضروری به نظر می رسد. از این رو، در این مقاله روش تحلیل قوانین همبستگی برای داده های جمع آوری شده از پرونده های حوادث یکی از واحد های صنعتی مشغول به فعالیت در صنعت تولید قطعات فولادی شمال ایران مورد استفاده قرار گرفته است. نتایج این تحقیق که آشکار کننده برخی عوامل پنهان در بروز حادثه است، می تواند برای سایر کارخانجات فعال در این حوزه مورد استفاده قرار بگیرد.

روش پژوهش

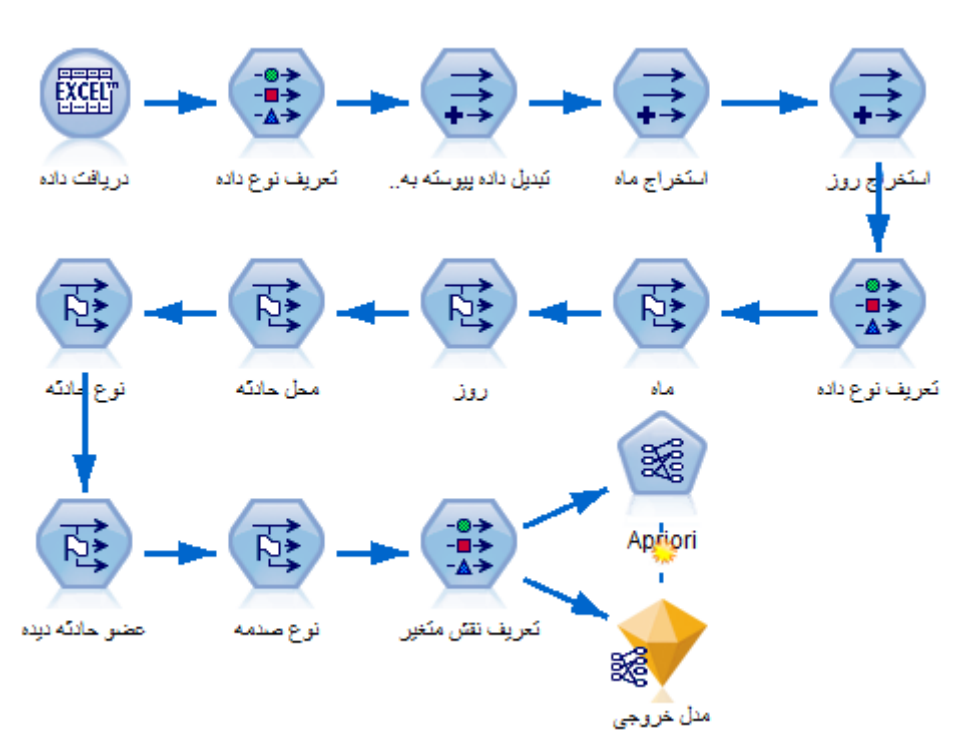
همانطور که در بخش مقدمه ذکر گردید، نخستین گام در استفاده از داده کاوی ایجاد یک پایگاه داده متناسب با هدف تحلیل می باشد. برای این منظور دسته بندی اطلاعات موجود در ۹۱۰ پرونده ی حادثه در کارخانه مورد مطالعه در طی سال های ۱۳۹۵ و ۱۳۹۶ در گام نخست صورت پذیرفت. پس از این دسته بندی یک ماتریس داده شامل مشخصه های تاریخ (معرف روز و ماه حادثه)، محل حادثه در کارخانه، نوع حادثه، عضو حادثه دیده و نوع صدمه برای هر حادثه ایجاد گردید. در جدول شماره ۱ بخشی از ماتریس داده مورد استفاده ارائه گردیده است.

جدول ۱: بخشی از ماتریس داده مورد استفاده در تحلیل مستخرج از گزارشات حوادث

ردیف	تاریخ	محل حادثه	نوع حادثه	نوع صدمه	عضو حادثه دیده
۱	۹۵۰۱۰۴	پرس ضربه ای	گیرکردن بین دستگاه	کوفتگی و ضربدیدگی	انگشت دست راست
۲	۹۵۰۱۰۴	سالن تولید	تماس با شی برنده	بریدگی	انگشت اشاره دست راست
۳	۹۵۰۱۱۷	سالن تولید-جوشکاری	برخورد با دستگاه جوش برق	شکستگی	انگشت دست راست
۴	۹۵۰۱۲۳	سالن تولید	برخورد به دستگاه پرس	کبودی و بریدگی	انگشت شصت دست راست
۵	۹۵۰۱۲۴	سالن تولید	افتادن جسم	بریدگی	روی پای چپ
۶	۹۵۰۱۳۱	سالن تولید	بخورد به دستگاه پرس	ضرب دیگی	انگشت سوم دست راست
⋮					
۹۱۲	۹۶۱۲۲۴	سالن تولید-گیوتین	فرو رفتن جسم	بریدگی	ساق پای راست

گام دوم آماده سازی داده می باشد. با توجه به اینکه از روش Apriori برای استخراج قوانین همبستگی در نرم افزار IBM SPSS Modeler استفاده می شود، متغیر های اولیه از نوع اسمی (Nominal) باید به نوع صفر-یک (Flag) تبدیل شوند. در یک متغیر از نوع اسمی ارزش ها برای هر رکورد، از مجموعه ای شامل تعدادی عضو انتخاب می شوند. لذا برای ایجاد متغیر های صفر-یک معادل لازم است ابتدا به تعداد اعضای مجموعه مزبور متغیر صفر-یک هر کدام متناظر با یک عضو آن مجموعه تعریف شود. سپس برای هر رکورد چنانچه ارزش خاصی مد نظر باشد در متغیر صفر-یک متناظر مقدار ۱ و در سایر متغیر های جدید صفر ثبت می شود. لازم به ذکر است این تبدیل برای همه متغیر ها اعمال می شود.

گام سوم مدل سازی الگو های موجود در حوادث می باشد. با توجه به اینکه در این تحقیق تمرکز بر روی نوع صدمه است، به دنبال قوانینی مانند اگر A آنگاه B هستیم که در آن B نوع صدمه و A مجموعه ای از شرایط محیطی، فنی و زمانی می باشد. برای اینکه هر قانون بدست آمده از لحاظ کیفیت قابل قبول باشد، از دو معیار پشتیبان (Support) و اطمینان (Confidence) استفاده می شود. این دو معیار مقداری بین صفر الی ۱ (یا ۰ الی ۱۰۰ درصد) اختیار می کنند و قانونی قابل اتکا است که همزمان هر دو این معیار های برای آن به ۱ (یا ۱۰۰ درصد) نزدیک باشد. پشتیبان یک قانون بیان کننده نسبت حوادثی است که مجموعه عوامل A و B در آن مشاهده شده است. اطمینان هم بیان می کند در چه نسبتی از حوادث عوامل B وجود دارد اگر عوامل A وجود داشته باشد. مساله اساسی در این گام چگونگی یافتن قوانینی است که دارای پشتیبان و اطمینان قابل قبول باشند. هر چند که تعداد کل قوانین ممکن محدود است اما با این حال تعداد زیادی را شامل می شود. بنابراین روش های ویژه ای استفاده می شود که بدون بررسی همه قوانین ممکن در با تلاش کمتری به بهترین قوانین دست یافت. از جمله این روش ها Apriori است. برای آگاهی از مبانی تئوری این روش می توانید به مرجع [۱۰] مراجعه نمایید. این روش در این تحقیق در محیط نرم افزار IBM SPSS Modeler 18 پیاده سازی گردیده است. خاطرنشان می سازد آماده سازی ویژه ای روی متغیر های ورودی به این روش در نرم افزار به گونه ای که در گام دوم ذکر گردید لازم است. در شکل ۱ فرایند مدل سازی در محیط نرم افزار ارائه شده است.



شکل ۱: فرایند تحلیل داده در نرم افزار IBM SPSS Modeler 18 جهت استخراج الگو های پنهان حوادث

یافته های تحقیق

پس از اجراء روش Apriori در محیط نرم افزار قوانین مندرج در جدول ۲ که ناظر بر الگوهای قابل اتکا از حوادث می باشند، استخراج گردید. لازم به ذکر است در استخراج قوانین، حداقل سطح اطمینان ۶۰٪ و پشتیبان ۵٪ در نظر گرفته شده است. همچنین قوانین بدیهی کنار گذاشته شده اند.

جدول ۲: قوانین قابل اتکا استخراج شده از داده های حوادث با روش Apriori

ردیف	اگر (Antecedent)	آنگاه (Consequent)	اطمینان قانون (%)	پشتیبان قانون (%)
۱	نوع حادثه برخورد با فلز	نوع صدمه بریدگی	۹۲.۲	۱۵.۵
۲	نوع حادثه فرو رفتن جسم	نوع صدمه بریدگی	۷۹.۸	۷.۶
۳	محل کارگاه محوطه و ماه خرداد	نوع صدمه بردیگی	۶۱.۲	۹.۶
۴	افتادن جسم	نوع صدمه ضرب دیدگی	۷۷.۷	۱۰.۸

بحث

بر اساس نتایج بدست آمده که در جدول شماره ۲ ارائه شده است، بریدگی در اعضاء بدن بر اثر برخورد با فلز یا فرورفتن جسم یک الگوی فراگیر در کارخانه مورد مطالعه است. لذا می توان استدلال نمود که تجهیزات ایمنی جهت پوشش دارای ضعف می باشد که نسبت به اصلاح آن باید اقدام گردد. یکی دیگر از الگوهای قابل توجه، وقوع حوادث در محوطه کارخانه است. با بررسی های میدانی مشخص شد دیو مواد در محوطه بدون رعایت مسائل ایمنی در ماه های پرکاری وجود دارد که به دلیل نبود پروتوکل های ایمنی به خاطر موقتی بودن دیو باعث ایجاد حادثه می گردد. از طرفی عدم وجود تجهیزات حفاظت شخصی نوع صدمه بریدگی شیوع دارد. ضرب دیدگی بخاطر سقوط اجسام الگوی متداول دیگری در حوادث شناسایی شده است. مجدداً با بررسی میدانی مشخص گردید سهل انگاری در مهار صحیح مواد دیو و در حین حمل، بی توجهی به محدوده های تردد و استقرار کارگران در حین کار و پیچیدگی عملیات از علل بالقوه هستند. در مورد اخیر ارتقاء فرهنگ ایمنی می تواند در جلوگیری از این نوع حوادث کارساز باشد.

در این مقاله روش تحلیل قوانین همبستگی که از رویکرد های داده کاوی در حل مسائل می باشد جهت تجزیه و تحلیل حوادث استفاده شد. تمرکز اصلی در استفاده از این روش کشف الگوهای شایع که منجر به انواع صدمه در یک کارخانه تولید کننده قطعات فولادی می شود، است. برای این منظور، در سه گام ایجاد پایگاه داده، آماده سازی داده و مدل سازی شناسایی الگوهای متداول به اجرا درآمد. تعداد ۴ قانون قابل اتکا با حداقل سطح اطمینان و پشتیبانی قانون انتخاب گردید. با کنار هم قرار دادن قوانین استخراج شده و شواهد میدانی، عدم تناسب تجهیزات حفاظت فردی در بخش پوشش، وجود رویه های موقتی در عملیات کارخانه بدون تعریف پروتوکل های ایمنی مناسب و سهل انگاری در پیروی از رویه های ایمنی جاری از مهمترین علل وقوع حوادث پر تکرار در این واحد صنعتی مورد شناسایی قرار گرفت.

References

- [1] Khanzode VV, Maiti J, Ray PK. Occupational injury and accident research: A comprehensive review. *Safety Science*. 2012; 50(5): 1355-1367.
- [2] Chang LY, Chen WC. Data mining of tree-based models to analyze freeway accident frequency. *Journal of safety research*. 2005; 36(4): 365-375.
- [3] Bayam E, Liebowitz J, Agresti W. Older drivers and accidents: A meta-analysis and data mining application on traffic accident data. *Expert Systems with Applications*. 2005; 29(3), 598-629.
- [4] Mirabadi A, Sharifian S. Application of association rules in Iranian Railways (RAI) accident data analysis. *Safety Science*. 2010; 48(10): 1427-1435.
- [5] Kumar S, Toshniwal D, Parida MA. comparative analysis of heterogeneity in road accident data using data mining techniques. *Evolving systems*. 2017; 8(2): 147-155.
- [6] Cheng, C. W., Leu, S. S., Cheng, Y. M., Wu, T. C., & Lin, C. C. (2012). Applying data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry. *Accident Analysis & Prevention*, 48, 214-222.
- [7] Sanmiquel L, Bascompta M, Rossell JM, Anticoi HF, Guash E. Analysis of occupational accidents in underground and surface mining in Spain using data-mining techniques. *International journal of environmental research and public health*. 2018; 15(3): 462.
- [8] Sarkar S, Verma A, Maiti J. Prediction of occupational incidents using proactive and reactive data: a data mining approach. In *Industrial Safety Management*. 2018; 65-79.
- [9] Hämmäläinen P, Takala J, Saarela KL. Global estimates of occupational accidents. *Safety science*. 2006; 44(2), 137-156.
- [10] Hegland M. The apriori algorithm—a tutorial. In *Mathematics and computation in imaging science and information processing*. 2007; 209-262.

- [11] Özeydin, E., Fışkın, R., Uğurlu, Ö., & Wang, J. (2022). A hybrid model for marine accident analysis based on Bayesian Network (BN) and Association Rule Mining (ARM). *Ocean Engineering*, 247, 110705.
- [12] Sarkar, S., & Maiti, J. (2020). Machine learning in occupational accident analysis: A review using science mapping approach with citation network analysis. *Safety science*, 131, 104900.
- [13] Garcia-Arroyo, J. A., & Segovia, A. O. (2020). Occupational accidents in immigrant workers in Spain: The complex role of culture. *Safety science*, 121, 507-515.
- [14] jannaty R, abedi K, jannaty L. Evaluation of effective parameters on the causes of occupational accidents in the construction industry using the combined method of Fuzzy logic-AHP and Pareto theory (Case study: Kurdistan province). *tkj* 2023; 14 (4) :14-24