



AN ASSOCIATION RULE MINING MODEL FOR THE ASSESSMENT OF THE CORRELATIONS BETWEEN THE ATTRIBUTES OF SEVERE ACCIDENTS

Bilal Umut AYHAN^{1,2}, Neşet Berkay DOĞAN¹, Onur Behzat TOKDEMİR^{1*}

¹Department of Civil Engineering, Middle East Technical University, Ankara, Turkey

²Department of Civil Engineering, Istanbul Bilgi University, Istanbul, Turkey

Received 03 September 2019; accepted 03 December 2019

Abstract. Identifying the correlations between the attributes of severe accidents could be vital to preventing them. If such relationships were known dynamically, it would be possible to take preventative actions against accidents. The paper aims to develop an analytical model that is adaptable for each type of data to create preventative measures that will be suitable for any computational systems. The present model collectively shows the relationships between the attributes in a coherent manner to avoid severe accidents. In this respect, Association Rule Mining (ARM) is used as the technique to identify the correlations between the attributes. The research adopts a positivist approach to adhere to the factual knowledge concerning nine different accident types through case studies and quantitative measurements in an objective nature. ARM was exemplified with nine different types of construction accidents to validate the adaptability of the proposed model. The results show that each accident type has different characteristics with varying combinations of the attribute, and analytical model accomplished to accommodate variation through the dataset. Ultimately, professionals can identify the cause-effect relationships effectively and set up preventative measures to break the link between the accident causing factors.

Keywords: accident analysis, Association Rule Mining, data mining, network analysis.

Introduction

The construction industry has evolved into incredibly complex structures due to significant developments in technology. This evolution brings along with it some serious issues; in other words, complexity has created an increased level of problems in occupational health and safety (OHS). Throughout history, attempts have been made to take precautions to prevent death or serious injuries. In this respect, OHS is becoming a significant concern that requires detailed investigations.

In 2015, 3.2 million non-fatal accidents, each of which resulted in the loss of at least four calendar days, and 3,786 fatal accidents were recorded in European Union countries (EU-28) in total (Eurostat, 2015). In addition, 2.8 million non-fatal accidents leading to 882,730 lost workdays and 5,147 fatalities occurred in the US in 2017 (U.S Bureau of Labor Statistics, 2017). According to Eurostat (2015) and the U.S Bureau of Labor Statistics (2017), 21% and 20.7% respectively of fatal occupational accidents occurred in the construction industry. These statistics show that the rate of accidents is so high that it cannot be ignored.

Construction projects have different properties, such as scope, geographical conditions, and risk factors (Project Management Institute [PMI], 2008). Such variations make each project unique and too complicated for the creation of systematic approaches to solving the problems of industry. As mentioned above, one of the biggest challenges in construction is decreasing or even avoiding harm to people or the environment. Although there is no easy way to prevent safety failures, accidents are the work events that can be avoided by implementing the following procedures adequately. These are (1) identifying the contributory factors and preparing viable risk assessments accordingly, (2) taking the necessary precautions to control the outcome of risk assessment, (3) creating and applying strict safety procedures, and creating a sustainable environment by employing the lessons learned the system. For decades, researchers have also put enormous effort into accomplishing these.

The first step in preventing accidents is to identify the attributes leading to occupational accidents accu-

*Corresponding author. E-mail: tokdemir@metu.edu.tr

rately. Various methods have been applied to determine which attribute or attribute groups lead to accidents. For instance, Mohammadi et al. (2018) applied qualitative content analysis to categorize underlying reasons for construction accidents after reviewing papers, which concentrated on exploring the safety factors, and they developed a hierarchical framework for measuring how vital these attributes were. There are several studies previously conducted in order to investigate the correlation of contributory factors of occupational accidents in construction sites via Association Rule Mining (ARM) (e.g., Liao & Perng, 2008; Cheng et al., 2010; Verma et al., 2014; Shin et al., 2018; Guo et al., 2019; Liao et al., 2019). Some other studies established a predictive model and introduced a simple preventative measure that cannot be adaptable for changes in data (Ayhan & Tokdemir, 2019a, 2019b, 2020). The lacking point of the current studies is that they require an individual examination to capture prospective risks. That means the adaptation process of the models is not quite well. Besides, the existing models are suffering from the use of data, obtained subjectively. The OHS professionals put their opinions in the recording process, so the reality tends to diverge from it (Tixier et al., 2017). Moreover, these studies did not question the accidents in terms of their types individually or their similarities collectively.

In contrast to the literature, the proposed research contributed to the literature by examining explicitly the most commonly observed accidents at the construction site to fulfill the gap in the literature. The study aimed to introduce an adaptable and analytical model to extract correlations between the attributes, inducing the construction incidents. It promoted an attribute-based prevention system that can automatically respond to the problem no matter the type of it is.

An arm-based analytical model is developed to show the connection between the attributes by adhering to the factual knowledge related to nine different accident types through case studies. The purpose of using the association rule mining is that the arm is a popular data mining algorithm, which produced significant relationships along with the data structure. The working logic relies on the frequent itemset mining, which can successfully cooperate with the incident data because the dataset encapsulates a high number of attributes with variate observation rates. Additional analysis is carried out regardless of the type of accident to gain a comprehensive insight about high-severe accidents and validate the adaptability of the proposed model with different datasets. Ultimately, the study achieved to reveal meaningful relationships between the attributes leading to these different types of accidents. With the help of this study, OHS professionals can focus on each accident type more specifically and minimize the hazardous consequences by taking the necessary precautions regarding the significant correlations identified. Moreover, obtaining meaningful conclusions from the different accident types demonstrated the compliance of the model to any computational systems as well. Besides, the results and the proposed methodology will be compared with the accident prevention systems given in the previous studies.

1. Occupational health and safety challenges in the literature and practice

The seriousness of accidents' consequences has attracted researchers' attention for decades. They have put a great deal of effort into discovering the characteristics of accidents by identifying the attributes. Comprehending the underlying correlations among the trigger attributes of an accident will provide a tremendous opportunity to prevent work-related safety failures common to construction sites (Winge et al., 2019).

Researchers have examined the safety issue in the construction industry under several popular headings such as safety risk, safety performance and management, studies on specific accident types, and development of analytical models. Although their focus is to prevent accidents, the methodology of the studies and approaches to the problem tend to vary in each research. The studies have developed numerous analytical or expert models concerning the outcome of safety problems, but the success of the proposed or constituted model depends on deducing the correlations between the attributes. The following section shows the importance of understanding the inherent structure of the accident cases regardless of the topic or focus point of the study.

1.1. Safety risks

One of the most common topics that researchers have concentrated on is safety risks based on construction projects. Camino López et al. (2008) studied the accidents that occurred in Spain. They investigated the relationships between the affecting attributes and explored how these attributes influence the severity level of accidents. In another study, the relationships between the type of work were associated with the accident types, and correlations between them were investigated in detail (Kim et al., 2012). Aminbakhsh et al. (2013) employed an analytical hierarchy process to prioritize the safety risk elements with the help of OHS experts. Another safety risk assessment model was proposed to analyze different construction site layouts with various safety risk levels (Ning et al., 2018). Studies were conducted to investigate the similarities between the safety and risk perceptions of the stakeholders of construction projects and those of OHS professionals (Zhang et al., 2015; Zhao et al., 2016; Liao & Chiang, 2016). As the research discussed above, understanding the correlations between the attributes is critical to preventing accidents proactively since the OHS professionals should learn how to overcome the safety risks and how to manage risk assessment in advance.

1.2. Safety performance and safety management

Safety performance and safety management in the construction site are other leading topics of construction safety. Similar to the safety risk process, interpreting the correlations between the attributes is becoming crucial to satisfying better safety performance and safety assess-

ment. Choi et al. (2019) proposed an approach to determine the efficacy of the wearable sensor, which measures the physiological responses of workers. The study showed that there is a remarkable difference between workers' responses during low and high-risk activities (Choi et al., 2019). In another study, the applicability of Unmanned Aerial Systems (UAV) for safety inspection was discussed. The research was based on an analysis of the photographs taken by UAV to capture the hazardous working conditions, which were then used to implement preventative measures (Melo et al., 2017).

Performing safety management systems (SMS) is a critical element for satisfying the safety environment at construction sites. Adequate SMS requires a comprehensive investigation of the attributes that contribute to accidents. Yiu et al. (2018) examined the efficiency of implementing the SMS in the Hong Kong construction industry, and a similar study was performed for developing countries in the research of Kheni et al. (2010). Root cause analysis is commonly used to explore the core reasons behind a problem. Bayesian networks (Gerassis et al., 2016) and Social Network Analysis (Eteifa & El-adaway, 2017) were used to determine the root causes of fatal accidents. On the other hand, Shao et al. (2019) discovered that fatal accidents are more frequent in summer, on Mondays and during the time intervals of 10–11 and 15–16, and half of these accidents are falls from heights. Moreover, it was revealed that due to their biological and occupational dissimilarities, women uniquely handle occupational hazards in the workplace (Cruz Rios et al., 2017).

Lessons learned from the results of accident investigations promote remarkable advancement in safety performance. In this respect, safety training begins to play an essential role in accident prevention. The effectiveness of safety training was questioned in several studies (e.g. Bařađa et al., 2018; Loosemore & Malouf, 2019). Providing safety training is the most efficient way to transfer theoretical knowledge about safety to the employees and create awareness of OHS. Evanoff et al. (2016) designed a training program for inexperienced construction workers to enhance their knowledge about fall prevention. Furthermore, some strategies were proposed to improve the safety performance of ethnic minority workers (Chan et al., 2016). Olson et al. (2016) developed easy-to-use toolbox materials for supervisors. As well as their effectiveness, difficulties in achieving global implementation of safety practices were examined (Gao et al., 2018). Safety leadership is a significantly influential factor for the proper implementation of SMS. An investigation of how safety performance can be promoted with effective safety leadership in railway projects (Stiles et al., 2018), and Grill and Nielsen (2019) provided detailed instructions to enable safety practitioners to influence safety performance positively.

1.3. Studies for specific accident types

Researchers have started to concentrate on the most frequent and severe accidents encountered in construction

sites to cut down the rate of fatalities and critical injuries. Falling from a height is the most frequent accident type that is observed in the construction industry (U.S Bureau of Labor Statistics, 2017). Hence, many studies have been conducted to investigate and to reduce the number of falling accidents. Winge and Albrechtsen (2018) evaluated the types of accidents to determine the most frequent ones and to analyze barrier failures. That study primarily concentrated on the proximal attributes, but there exist latent interactions and correlations, which require a deeper level of investigation. In another study, an attempt was made to identify the reasons behind fatal fall accidents in the US to suggest feasible solutions (Dong et al., 2013). The accident data source used in the study did not provide information regarding some critical factors of accidents such as working conditions and environments and safety climate factors. For this reason, the results may not reflect all aspects of the accident. Fall-related equipment, which includes mast climbing work platforms (Wimer et al., 2017) and scaffolds (Rubio-Romero et al., 2012), were also evaluated in terms of their stability. Fall accidents of drywall installers were also studied to identify their characteristics (Schoenfisch et al., 2014).

Loss of balance is associated with the postural stability of workers; therefore, factors related to balance loss were identified (DiDomenico et al., 2010), and Inertial Measurement Unit sensors were used for postural stability analysis of construction workers (Jebelli et al., 2016). In addition to these analyses, a long term study was proposed for increasing the effectiveness of the training of foremen in order to decrease fall accidents (Kaskutas et al., 2013).

There are some studies that focused on other specific accident types in addition to fall accidents. Accidents due to electricity (Suárez-Cebador et al., 2014), and earthmoving equipment-related incidents were also investigated (Kazan & Usmen, 2018). Suárez-Cebador et al. (2014) investigated electricity accidents in terms of the 13 different attributes whose significance levels were lower than 0.05. However, the study was not concerned with human elements or workplace factors, which lead to the occurrence of accidents at construction sites. Tokdemir and Ayhan (2019) investigated foreign body damage and developed a hybrid model of ANN-AHP to predict the severity level of accidents. As well as the prediction process, the most frequently observed attributes in this accident type were examined to help professionals take the necessary precautions to prevent accidents. Although many researchers addressed the issue of construction accidents separately, there has been no study that encapsulated the investigation of more than one accident type with its common and distinct points. The reason for this is that accidents may carry certain characteristics that can be universal or distinct and may need significant analysis to be revealed. The proposed approach aimed to disclose correlations between the attributes of various accident types, so the distinct and similar points of the most common accidents will be presented in the remaining parts of the paper.

1.4. Analytical models for the construction safety

Artificial intelligence (AI) is based on the idea of creating super-machines that mimic the human brain and process complicated information. Patel and Jha (2014) proposed Artificial Neural Network (ANN) based models for the estimation of safety climate efficiency (Patel & Jha, 2014) and the evaluation of the safety level of employees' behavior (Patel & Jha, 2016). Fuzzy neural network technology and methods of multivariate statistical analysis were used together for the determination of environmental safety levels in Russia (Glinskiy et al., 2016). Convolutional neural networks (CNN) and long short-term memory networks were selected to detect safe and unsafe behavior on the construction site (Ding et al., 2018). Other research focused on checking the certification of workers via deep learning, which can recognize each worker's face and check if he has valid certification (Fang et al., 2018). The Fuzzy Delphi Method and the Decision-Making Trial and Evaluation Laboratory were used together to discover the factors affecting occupational accidents and their relationships (Bavafa et al., 2018). Mohandes and Zhang (2019) developed a comprehensive hybrid fuzzy-based risk assessment model to eliminate the shortcomings of existing approaches to safety management systems. Ayhan and Tokdemir (2019a) also proposed a model, which consists of ANN and Fuzzy Logic, for predicting the consequences of accidents. The model can process the collected data from real accidents and enables OHS professionals to take the required precautions to prevent work-related injuries during both preconstruction and construction stages. Ayhan and Tokdemir (2019b) improved their previous study by implementing the clustering technique to obtain better accuracy in prediction. As the last step, they introduced an enlightening system as a preventative measure with regards to the frequency of attributes observed in fatal events. Although results gave promising outcomes for incident prevention, more detailed investigation depending on different analytical models should be necessary to acquire more reliable results.

Attribute-based safety assessment strategies depending on Analytical models are also one of the popular topics among the researchers. The literature has remarkable studies for searching the correlations between the attributes of the construction incidents. The record-keeping mechanism in the construction project are mostly subjective as the personal opinions of experts may induce reality. Esmaili et al. (2015a) proposed an attribute-based risk assessment system using the reliable national database of 1812 injury reports of struck-by incidents. They aimed to overcome the problem of the requirement of individual evaluation and investigation in existing studies. The methodology introduced addressed the construction accidents of struck-by accidents only. Although the study was limited due to assessing only a specific type of accident, it promoted to enhance proactive safety management strategies.

Esmaili et al. (2015b) advanced their study, which encapsulated the development of predictive models based on

generalized linear regression analysis. The specialists can forecast the risk level of the project proactively. In spite of all these contributions, the study may remain weak in some points. The proposed system cannot be implemented before the start of construction. Besides, it dealt with only struck-by accidents only.

Mistikoglu et al. (2015) conducted a decision tree analysis, which is a supervised data mining method. The research investigated a fall from height accidents taken from the US Occupational Safety and Health Administration (OSHA). The results depicted the associations between the attributes. The most prominent factors to increase the fall accidents were fall distance, fatality/injury cause, safety training, and construction operation prompting fall.

The limitation of existing studies can be described as the fact that the models depend on data, recorded subjectively, as mentioned before. To address this limitation, Tixier et al. (2017) proposed a framework, including natural language processing (NLP), to standardize the fundamental attributes of incidents. The objective was to attain a new safety knowledge from mining multidimensional data. A small amount of relevant observation was obtained from the bulk of the data.

The proposed system revealed promising results in achieving the standard way of extracting valuable insights from injury reports by applying unsupervised approaches. Researchers focused on only three different accident types, including struck by against objects, falling on the same level, caught between the objects. It may lack in giving insights about the remaining accident types.

ARM is another useful technique for indicating the relationships between the attributes, and it is an approach that is applied in the construction industry. Lin and Fan (2018) have taken into consideration the inspection grades of 990 public construction projects for the determination of the relationships between defect types and inspection type, and the genetic algorithm and ARM were combined for extracting the defect patterns (Cheng et al., 2015). Occupational accidents and fatality reports that occurred in Taiwan between 2000 and 2007 (Cheng et al., 2010) and in Korea between 2005 and 2010 (Shin et al., 2018) were also investigated. Verma et al. (2014) used ARM for the investigation and identification of the operating factors resulting in incidents in a steel company in India between March 2010 and July 2013. Guo et al. (2019) created a fusion model by combining a pre-developed structured learning algorithm and weighted association rules to generate a hazard association network. Liao et al. (2019) combined human dynamics analysis with ARM to find the time-statistical laws for the time distribution of workers' unsafe behavior in a metro construction site in Wuhan, China. ARM was used for the investigation of factors that led to traffic accidents (Geurts et al., 2012; Yao et al., 2018; Xu et al., 2018).

Furthermore, work zone crashes (Weng et al., 2016), vehicle-pedestrian crashes in Louisiana (Das et al., 2018), and crashes occurring during rainy weather (Das & Sun, 2014) were explicitly focused. Other researchers tried to

discover meaningful relationships, patterns, and trends for railway accidents in China (Chen et al., 2017) and in Iran (Mirabadi & Sharifian, 2010). Likewise, Zhang and Liu (2011) established a primary database for marine traffic accidents and employed ARM to determine dependency on the factors behind these accidents.

This research has focused explicitly on the frequently recognized accidents in construction sites to fill the missing points in the literature described in detail above. The aim is to prove how severe accidents can be prevented by uncovering the correlations between the attributes.

2. Methodology of research

The focus of the study is to examine the correlations between the attributes coherently to avoid severe accidents. The factors involved in each of nine different accident types are investigated, and an analytical model is proposed to reveal the characteristics of each accident type by controlling the various combinations of attributes. This research had three stages, as shown in Figure 1. First, the data preparation process was carried out. The accident records were taken from construction companies that had

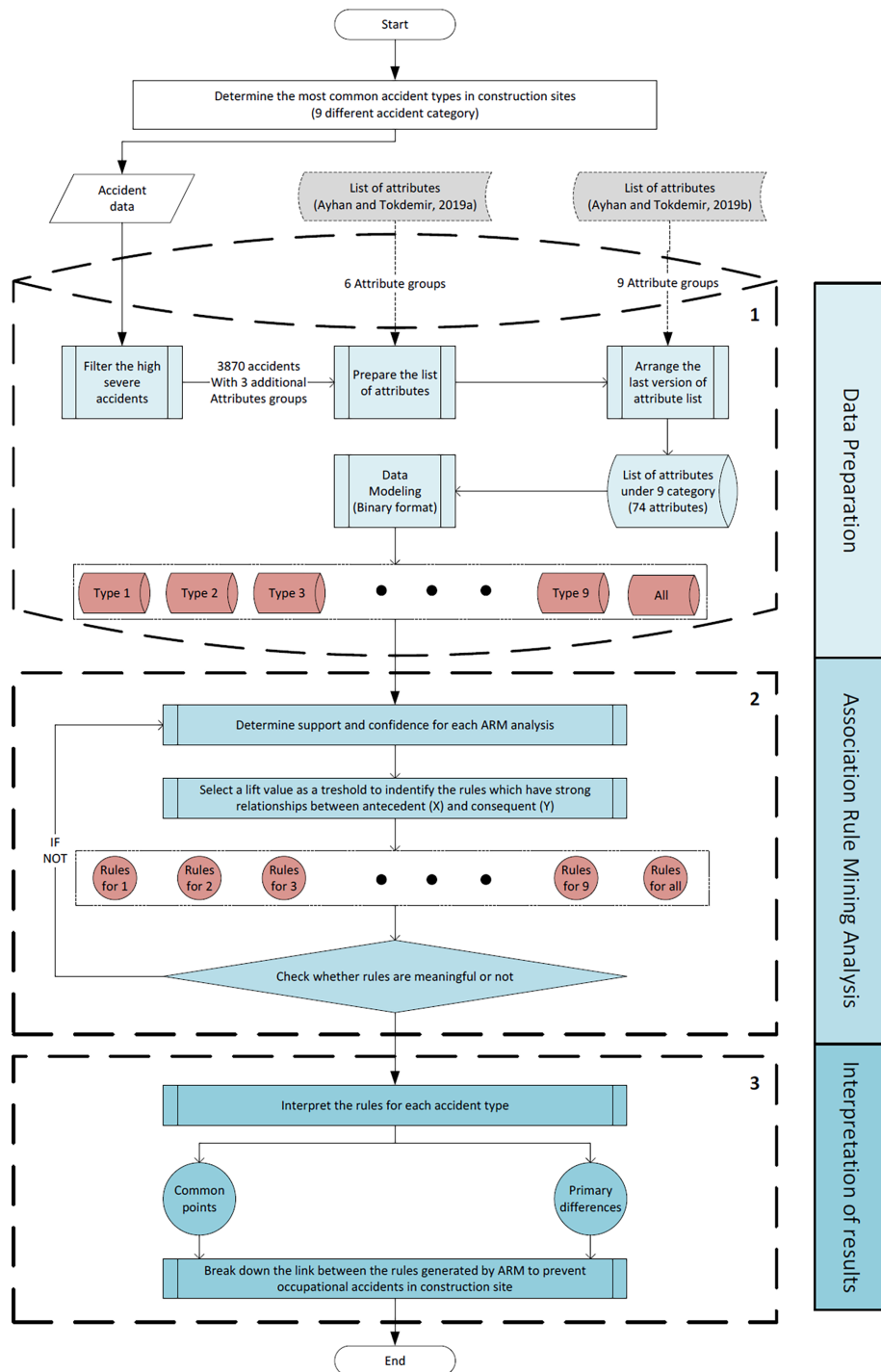


Figure 1. Research plan and process

agreed to share accident information confidentially. Then, the study developed analytical modeling with the dataset utilizing ARM analysis. In the final stage, the correlations with strong bonds were interpreted to enable the necessary precautions for each type of severe accident to be recommended.

2.1. Data preparation

The study started by defining the most common accident types encountered in construction sites, and each of these accidents is demonstrated in Figure 2. The authors contacted the leading construction companies, which have several construction sites in the Europe and Asia regions. The aggregation of the incident record regarding the countries is as follows. The most significant portion of the accident records belongs to Russia and Turkey, as 36% and 31% of the whole dataset, respectively. Accidents in Poland occupied 9%, and the portion belonging to Turkmenistan is equal to 11%. The remaining countries inside the dataset are Iraq, Saudi Arabia, and Kazakhstan, and their portions are 7%, 3%, and 3% percentage, accordingly. The companies agreed to share their high-severe accident documents, including both past and current records on a confidential basis; therefore, any personal or organizational information has not been involved in this study. After collecting the accident data, the records were filtered to ensure that all accidents belonged to the high-severe class involving first aid, medical intervention, workday loss, or fatality categories. After filtering one more time, 3,870 accident records remained as the data to be entered into an ARM.

The following step in the data preparation was to organize the list of attributes to describe the accidents. The statistical analysis requires a meaningful expression with ordinal or nominal variables so that the computational model can make the necessary calculation to extract findings from the dataset. There is significant research available in the literature to explain the accident cases rationally. For this reason, the attribute categories were determined based on Ayhan and Tokdemir (2019a) at first. Then, attributes were rearranged with the help of OHS experts and the study of Ayhan and Tokdemir (2019b) to form a more coherent list and three more categories (occupation, age, and experience duration of people involved in accidents) were integrated into this study.

Table 1 describes the attributes taken from the mentioned study. The reason for coding them in binary format is that an ARM-based analytical model can easily be adapted to work with binary expressions. Besides, more than one attribute can be observed in one case simultaneously, so this method allows them to be expressed categorically. Thus, if one of the attributes specified in Table 1 exists in the accident case, the assigned value will become 1. There are also continuous and categorical attributes that need to be converted into a binary expression. Table 2 indicates the groups of attributes with their characteristics before ARM analysis. These should also be converted into the binary format by coding as 1 and 0. This was achieved by dividing the attributes into sub-groups, as shown in Table 2 to accommodate the option defined. To illustrate, when it comes to age, four sub-attributes were created to reflect four age groups.

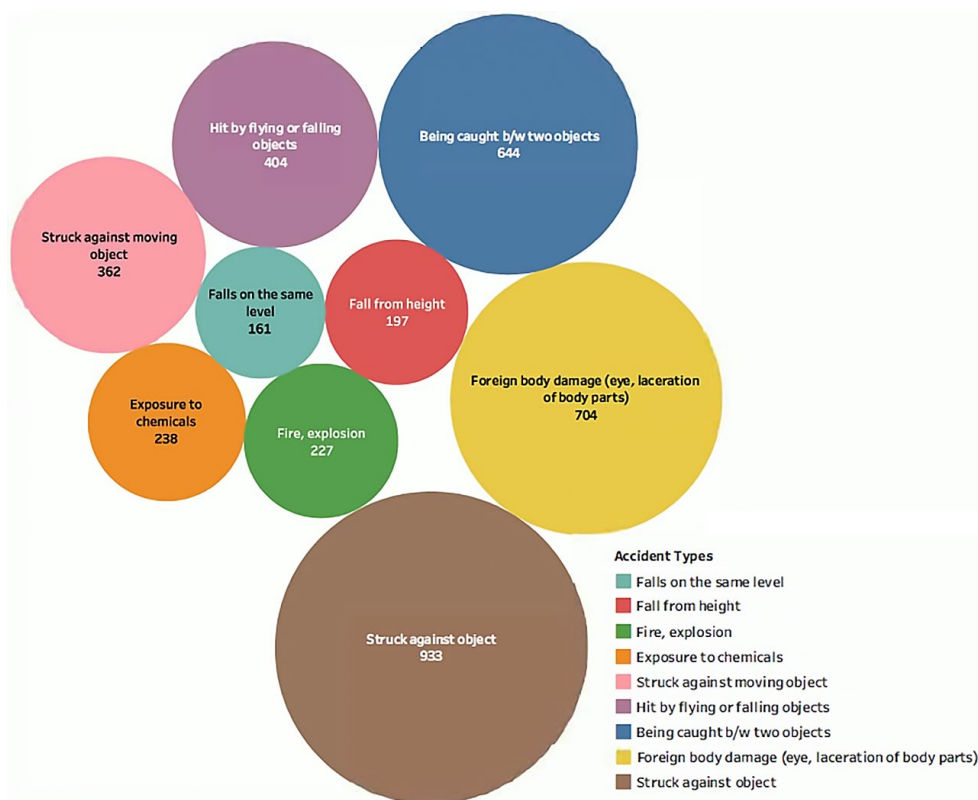


Figure 2. Type of accidents and their observation frequencies

Table 1. The master list of attributes used in this paper (based on the study of Ayhan and Tokdemir (2019a, 2019b))

Attribute Groups	ID	Expression
Activity	AC-1	Finishing works
	AC-2	Installation works
	AC-3	The usage of hand-tool / equipment
	AC-4	Daily activities
	AC-5	The usage of vehicle/construction equipment
	AC-6	equipment
	AC-7	Mobilization on/off-site
	AC-8	Lifting operations
	AC-9	Re-bar/formwork installation
	AC-10	Repair/maintenance works
	AC-11	Welding/hot works
	AC-12	Testing works
	AC-13	Geotechnical works
	AC-14	Landscaping works
	AC-15	Working at height
	AC-16	Field measurement works
	AC-17	Excavation works
	AC-18	Concreting
	AC-19	Material drop
	AC-20	Working with chemicals MEP works
Workplace factors	W-1	Problems with the method of statement
	W-2	The lack of management system
	W-3	Inadequate communication (general)
	W-4	Inadequate incident analysis systems
	W-5	Inadequate maintenance / repairment mechanism
	W-6	The lack of OHS training
	W-7	Insufficient control, tracking
	W-8	Incorrect recruitment procedures
	W-9	Incorrect protection measures
Hazardous cases	HA-1	NCRs in working environment
	HA-2	Mechanic Hazards / NCRs
	HA-3	NCR in usage of hand-tool / equipment / construction equipment
	HA-4	NCRs in safety protection measures
	HA-5	Chemical hazards
	HA-6	Natural hazards
	HA-7	Radiation exposure
	HA-8	Fire, explosion
Risky Behaviors	RB-1	Inability to perceive external risks
	RB-2	Tending to use a shortcut
	RB-3	Violation of safe work policy
	RB-4	Incorrect physical movement
	RB-5	Incorrect usage of equipment, hand-tool
	RB-6	Incorrect/absence of safe work policy
	RB-7	Working on energized equipment
Human Factors	HU-1	Insufficient skill and perception
	HU-2	Physical disability
	HU-3	Psychological disability
	HU-4	Problems related to education level
	HU-5	Problems resulting from the unbalanced workload
	HU-6	Problems resulting from incorrect management system

After performing these operations, the total number of attribute categories increased to nine, and these categories are composed of 74 attributes in total. The details of the attribute categories and attributes are shown in Table 1. Therefore, nine different datasets, as well as the all, were ready to be analyzed in ARM, as demonstrated in Figure 1.

2.2. Analytical model development via Association Rule Mining (ARM)

Association Rule Mining (ARM) is a data mining technique that was introduced by Agrawal et al. (1993) to investigate the purchase tendencies of customers. It is commonly used for the determination of hidden patterns and relationships between the variables in datasets and the dependencies among these variables. In this study, the authors applied the Apriori Algorithm, which is one of the most widely used algorithms for ARM. This algorithm consists of two phases. In the first phase, the analysis captures the most frequently observed items or attributes from the dataset, and then the analysis is advanced by generating logical associations rules between the attributes. Figure 3 represents the step of the ARM algorithm, which was adopted in the study. The algorithm steps were demonstrated based on Rapidminer Studio 9.2.0 software. In this study, both Rapidminer Studio and RStudio (2019) were utilized.

The rules are defined and presented in the form of “ $X \rightarrow Y$ ”, where X is antecedent and Y is consequent. In other words, the rules sustain the characteristics of if-then statements where if and then clauses stand for the antecedent and consequent, respectively.

While generating association rules, support, confidence, and lift are the thresholds used to identify the most robust rules, as shown in Figure 3. Support indicates how common the item is in the dataset, and confidence (also known as a conditional probability) shows how frequent the generated if-then associations are true. Lastly, the lift states how strong the relationship is. A higher lift value reflects a stronger relationship.

Support, confidence, and lift are calculated using the Eqns (1), (2), and (3).

$$\text{Support of } X \rightarrow Y = \frac{F(X, Y)}{N}; \tag{1}$$

$$\text{Confidence of } X \rightarrow Y = \frac{F(X, Y)}{F(X)}; \tag{2}$$

$$\text{Lift of } X \rightarrow Y = \frac{F(X, Y)}{F(X) * F(Y)}, \tag{3}$$

where N is the total number of items in the dataset, $F(X, Y)$ is the frequency of items and includes X and Y at the same time, $F(X)$ and $F(Y)$ are the frequency of X and Y , respectively.

Table 2. The list of additional attributes

Attribute Groups	Before ARM	ID	Expression	Converting Binary Expression	
	Type of data			Existence	Absence
Occupation	Categoric	O-1	Engineer	1	0
		O-2	Mechanical assembly crew	1	0
		O-3	Finishing work crew	1	0
		O-4	Rough work crew	1	0
		O-5	Repairman	1	0
		O-6	Construction equipment operator	1	0
		O-7	Administrative affairs	1	0
Time	Continous	TI-1	00:00–03:00	1	0
		TI-2	03:00–06:00	1	0
		TI-3	06:00–09:00	1	0
		TI-4	09:00–12:00	1	0
		TI-5	12:00–15:00	1	0
		TI-6	15:00–18:00	1	0
		TI-7	18:00–21:00	1	0
		TI-8	21:00–00:00	1	0
Experience	Continous	Exp-1	< 1 month	1	0
		Exp-2	1–3 month	1	0
		Exp-3	3–6 month	1	0
		Exp-4	6–12 month	1	0
		Exp-5	12–24 month	1	0
Age	Continous	AG-1	18–25	1	0
		AG-2	25–35	1	0
		AG-3	35–45	1	0
		AG-4	45–65	1	0

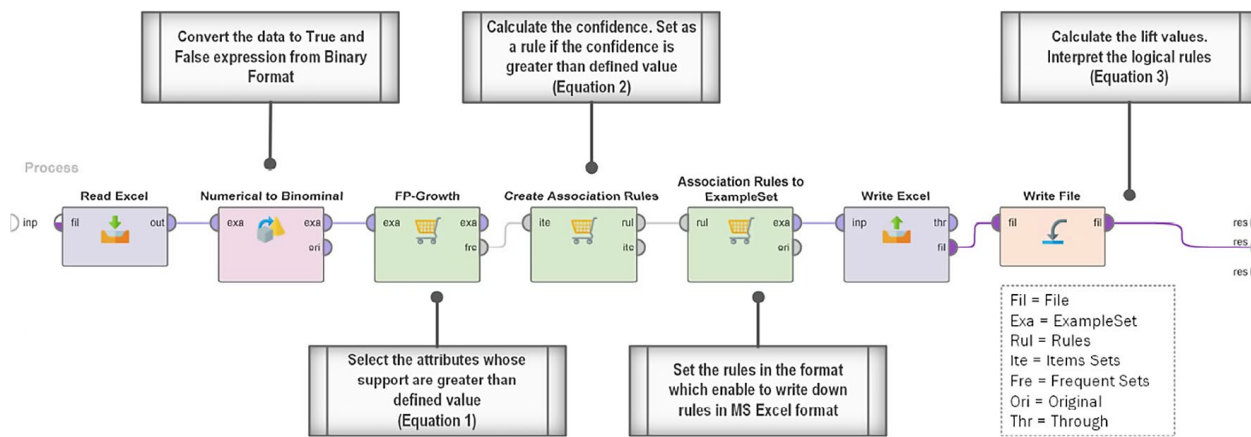


Figure 3. Representation of ARM analysis

As mentioned before, RapidMiner Studio 9.2.0 and RStudio are used for ARM analysis, and for all accident types, support, confidence, and lift are set as 0.15, 0.6 and 1.1, respectively. Gephi 0.9.2 (n.d.) software was used to obtain the visuals.

2.3. Interpretation of results

A vital part of the study was interpreting the ARM analysis for preventing the occupational accidents occurring in the construction sites. The outcomes of ARM are comprised of two elements, antecedent and consequent, as mentioned above. The rules between the attributes indicate a strong correlation between them while observing the construction accidents accordingly. If the rules constituted between the attributes are broken by taking necessary preventative

actions, the dangerous consequences of accidents can be avoided (Weng et al., 2016). Therefore, the results for all nine accident types were sorted according to their lift, which indicates how strong the correlation between antecedent and consequent is. ARM was also implemented for all accident types together to detect the variation in rules for each accident type.

3. Discussion of findings

3.1. The steps of analysis for the data including all accident types

The relationships between the attributes of accidents were evaluated with ARM by finding the correlation between them. Data of accidents was composed of nine most com-

monly observed accident types, as shown in Figure 2. ARM was applied to these nine accident types, as well as to the complete accident data, to observe the deviation in relationships between attributes. At first, implementing the ARM to the complete dataset took place to capture the common points of different accident types. Then, the analyses were done for nine accident groups to obtain a more detailed explanation for high severe construction accidents individually. Before the start of the analysis, the authors determined the threshold criteria for values of confidence, support, and lift as 0.6, 0.15, and 1.1, respectively.

The analysis began with understanding the accidents. To do so, the frequency of attributes, the accident items, was calculated. Figure 4 shows the item frequency of the attributes which had been observed in more than 15% of the 3870 accident cases. It was demonstrated that RB-1, HA-1, and RB-3 were the most profoundly observed attributes of all. Even though the figure was not capable of indicating a statistical conclusion, it gave a clue as to which attribute may require a significant level of intention to measure preventative actions. Moreover, it was expected that the rules produced by ARM would be highly associated with the frequently observed attributes.

The ARM analysis generated 202 rules for all the datasets when the support was 0.15, and confidence was 0.6. These results showed that 160 of 202 rules obtained from different accident types were unique for only one accident type at most. In other words, 160 rules did not appear in more than one accident type. The results of the analysis demonstrated the multicausal nature of accidents as also shown in complex ballasting operations (e.g. Gerassis et al., 2019). Hence, each should be investigated individually. Moreover, the results of the analysis with the whole dataset were compared with the results of individual accident analysis to examine whether deviation may exist through the accident type or not. According to the results, none of the 12 rules were involved in all of the nine accident types, whereas at least two rules were involved in at most two accident types. Therefore, it was confirmed

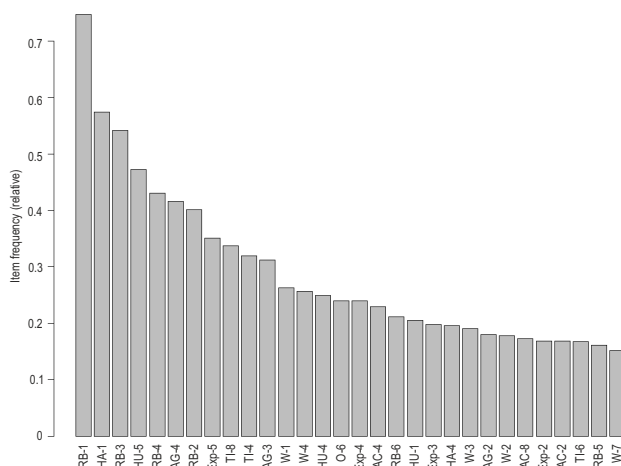


Figure 4. Attributes whose frequency was higher than 0.15

that the evaluation of accident types separately was more beneficial, since taking preventative actions based on the results of the analysis of all the data may not be sufficient to avoid an accident effectively.

Figure 5 explains how the rules were formed. The nodes without labels represented the rules. The edges which expressed the connections between the attributes were classified using different colors according to the accident type. As the accidents hold common points, they also have distinct characteristic features, as indicated in Figure 5. The nodes with labels expressed the attributes, and they are marked in terms of size and color. The bigger size and darker color indicated the higher support value and higher observation rate among the accident types. As was specified above, the rules were highly correlated with the attribute whose frequency was more significant than 0.5.

The next step was to interpret the lift and support criteria with confidence. The scatter plot of lift and support against confidence was visualized in Figure 6. The points on the graph were marked using different shapes to indicate the accident type. The color scale represents the confidence value. According to the figure, the majority of the rules occurred when the support had values between 0.15–0.20, whereas the lift values varied between 1.1–1.3. Figure 6 demonstrates that the highest confidence was obtained within these ranges. As the selected lift value was sufficient for obtaining rules with high confidence, the ARM results were filtered by applying the rule where lift values should be equal or greater than 1.1. This means that the rules whose lift was higher than 1.1 were taken into account in interpreting the interrelationships between the antecedents and consequents.

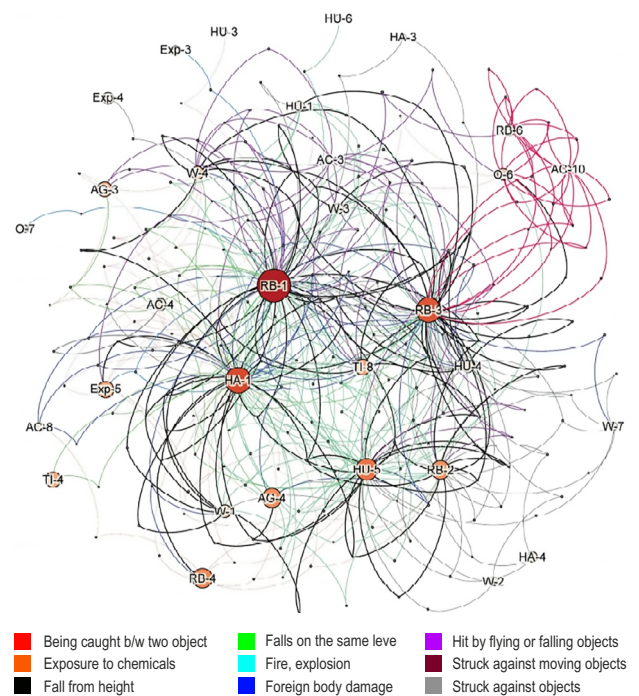


Figure 5. Visualization of a network for 202 rules (All accidents)

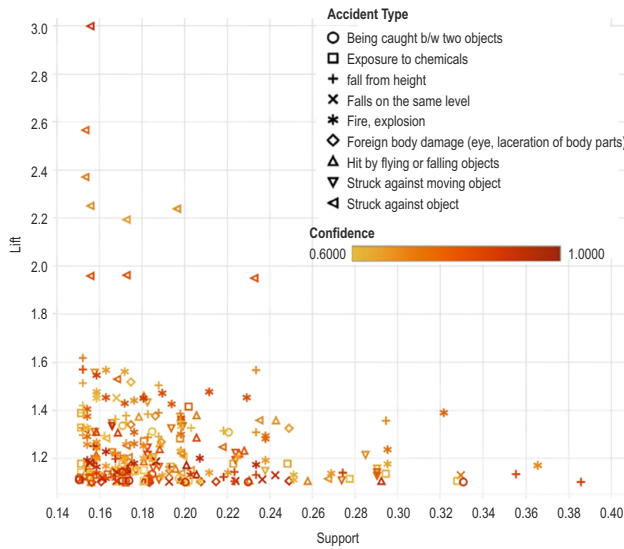


Figure 6. Scatter plot of 202 rules

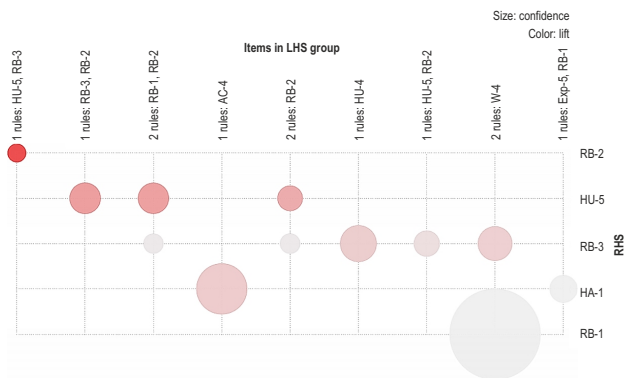


Figure 7. Balloon plot graph of 202 rules

After eliminating the rules with lower lift values, just 12 rules remained for all accident types. Figure 7 visualizes the filtered rules, considering the defined criteria. The rules were shown with balloon plots where the balloon’s size represents the frequency. The antecedents and consequents were labeled as LHS and RHS so that interpreting the relationships between the rules became easier. Figure 8 and Table 3 are given to clarify further. The network visualization of filtering rules was drafted in Figure 8 so that the relationships between the attributes can be expressed more effectively. Moreover, Table 3 gives the textual expression of logical rules for better understanding.

According to these results, a general conclusion about all accident types can be made as follows. RB-1, which is “Inability to perceive external risks”, was the most common attribute for the accident data and all types of accidents, and the accidents occurred most commonly in the time intervals of 9:00–12:00 and 21:00–00:00. Inadequate accident analysis systems may generate the inability to perceive and evaluate the external risks, and problems with training that result in a violation of safe work policies. The reasons behind these are that the root causes of the problems have not been deeply enough investigated, and the lesson learned from these accidents has not been

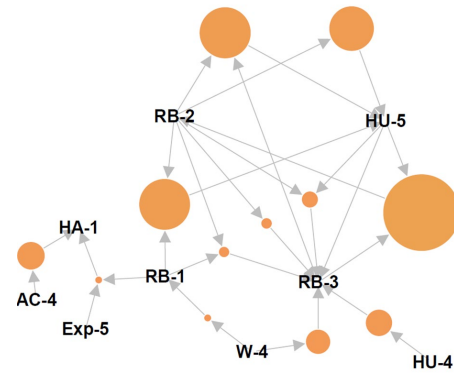


Figure 8. Network visualization of 12 rules filtered by lift defined values

Table 3. Textual expression of rules obtained by ARM

#	Antecedent	Consequent	Support	Confidence	Lift
1	RB-3, HU-5	→ RB-2	0.1566	0.6036	1.5060
2	RB-3, RB-2	→ HU-5	0.1566	0.6419	1.3583
3	RB-1, RB-2	→ HU-5	0.1742	0.6413	1.3569
4	RB-2	→ HU-5	0.2501	0.6241	1.3206
5	AC-4	→ HA-1	0.1612	0.7019	1.2220
6	HU-4	→ RB-3	0.1638	0.6584	1.2150
7	W-4	→ RB-3	0.1674	0.6519	1.2031
8	HU-5, RB-2	→ RB-3	0.1566	0.6260	1.1553
9	RB-2	→ RB-3	0.2439	0.6086	1.1232
10	RB-1, RB-2	→ RB-3	0.1651	0.6080	1.1220
11	W-4	→ RB-1	0.2116	0.8239	1.1022
12	RB-1, Exp-5	→ HA-1	0.1654	0.6312	1.1018

transferred to the workers properly. This violation creates the tendency to use shortcuts when the workload on workers is unbalanced, and as a result of this process, accidents with severe consequences have occurred.

3.2. The results of individual analysis for each accident type

The ARM was implemented for all accident types to capture more information about the interrelationships of the attributes. The potential relationships between the attributes for all accidents were disclosed previously, and this revealed the connections which should be broken to prevent accidents. Later on, the same procedure was conducted for each accident, and the results obtained are as follows. In Table 4, a summary of the rules which were generated by ARM and classified according to the type of accident was given. The analyses explored the characteristic features of each accident comprehensively. The vital difference between them all and individual analyses was the confidence values, which represent the likelihood of the occurrence for the described conditions. Separate

Table 4. Rules generated from the individual ARM analysis for each accident type

Accident Type	Antecedent	Consequent	Confidence	Support	Lift
Being caught b/w two objects	AC-4, Exp-5	RB-1, HA-1	0,743	0,171	1,145
	AC-4, HU-5	RB-2	0,601	0,185	1,312
	HU-3	RB-1	0,907	0,151	1,110
Exposure to chemicals	AC-8	RB-4	0,610	0,151	1,143
	AG-3	HA-1	0,648	0,193	1,168
	RB-1, TI-8	RB-3	0,647	0,185	1,203
	RB-4, HU-5	HA-1	0,705	0,181	1,271
	W-3	RB-3	0,762	0,202	1,417
Fall from height	Exp-4	RB-1	0,833	0,152	1,109
	HA-1, TI-8	RB-1	0,939	0,157	1,250
	HA-3	RB-3	0,805	0,168	1,109
	HA-4, RB-2	RB-3	0,868	0,168	1,196
	HU-1	RB-1	0,846	0,168	1,126
	HU-4	RB-1, RB-3	0,600	0,274	1,137
	RB-1, W-2	HU-5	0,789	0,152	1,571
	RB-3, W-7	RB-2	0,698	0,152	1,617
Falls on the same level	HA-1, W-4	RB-1	0,886	0,193	1,105
	RB-1, TI-4	HA-1	0,933	0,174	1,147
	TI-8, AG-4	RB-1	0,897	0,161	1,119
Fire, explosion	HU-5, AG-4	HA-1	0,837	0,181	1,450
	TI-8	HU-5	0,654	0,225	1,207
		RB-3	0,603	0,207	1,103
	W-1	RB-1, HU-5	0,649	0,163	1,568
Foreign body damage (eye, laceration of body parts)	Exp-3	RB-1	0,838	0,183	1,101
Hit by flying or falling objects	AC-8	RB-1, HA-1	0,636	0,168	1,157
	RB-4, W-1	RB-1	0,912	0,153	1,148
Struck against moving object	AC-3, HU-5	RB-3	0,674	0,166	1,226
	AC-3, W-4	RB-1	0,815	0,182	1,214
	RB-1, TI-8	RB-3	0,629	0,155	1,145
	RB-6	RB-1	0,815	0,182	1,214
Struck against object	HU-1	RB-3	0,723	0,198	1,221
	RB-6	AC-10	0,661	0,196	2,241
		RB-3	0,682	0,203	1,153
	W-3	RB-1	0,775	0,169	1,145

ARM analysis resulted in a higher confidence level compared with the rules in Table 3. Almost all accident types contain rules with confidence levels higher than 0.80. This means that the most likely triggers of each accident can be deduced more precisely by applying ARM separately.

For the accidents categorized as being caught between two objects, the inability to perceive external risks were generally observed to be the consequence of association rules. Construction equipment operators and workers with a physiological disability were not able to perceive the risks on the worksite. Moreover, this accident type occurred while workers with 12 to 24-months experience were working on routine activities in construction sites if the working environment was not suitable, and they were unable to evaluate the external risks. Using shortcuts were commonly involved in routine activities, and the accidents happened when this situation was combined with excessive workload.

Accidents of exposure to chemicals mostly occurred in the time interval between 21:00 and 00:00. Safe work policies have been violated in this time interval when the external risks have not been adequately evaluated. The victims of such accidents were mostly employees whose

age varied from 35 to 45, and they had exposure to these chemicals when the environment was not proper for work. Moreover, incorrect physical movement occurred during the re-bar/formwork installation. These types of movements also led to an improper working environment when they were combined with unbalanced work. Furthermore, safe work policies were violated if there was a lack of communication in the working site.

External risks could not be determined and evaluated correctly in falling from height cases if the following attributes were separately involved: inadequate education level, insufficient skill and perception, and the workers with experience of 6 to 12 months. Besides, safe work policies were violated when there was nonconformity of the hand-tools or construction equipment usage. A breach of these policies was also observed if shortcuts were preferred with improper personal protective equipment (PPE). Moreover, shortcuts were used in cases where control or tracking was insufficient, and safe work policies were violated, and also in the cases that there was an unbalanced workload and problems related to educational level. Additionally, lack of supervision and the inability to perceive external risks led to an unbalanced workload.

External risks related to falls on the same level accident type cannot be evaluated due to improper work environment and insufficient accident analysis system. Moreover, the inability to perceive these risks was most likely to indicate an improper working area during the 9:00–12:00 working period. Thus, it can be deduced that the external risks cannot be understood well in unsuitable work environments and vice versa.

In the accident due to fire or explosion category, violation of safe work policies and unbalanced workload were observed between 21:00 and 00:00. Victims of this type of accident were generally in the 45–65 age group and also worked under uneven working loads. Moreover, problems related to procedure in construction applications resulted in the inability to evaluate external risks and an unbalanced workload.

Foreign object damage was commonly due to the inability of workers with only 3 to 6 months of experience to perceive and examine the external risks. Moreover, in hit by flying or falling objects, incorrect physical movement and problems in construction procedure caused the inability to identify the external risks. During the installation of rebar or formwork, the inability to perceive external risks commonly led to accidents when it was accompanied by an improper working environment.

Hazardous consequences were also revealed due to being struck by a moving object. Incorrect safety policies or the lack of safe work policies resulted in the inability to evaluate external risks. This result was also obtained from the inadequate accident analysis system for the jobs that required the use of hand tools or other equipment. Moreover, safe work policies were violated if hand tools were used or the victims had an unbalanced workload.

Being struck by an object was the accident type most often observed, and the results showed that it generally occurred during hot jobs, for which violation or absence of safe work policies is extremely hazardous. Insufficient skills, inability to perceive risks and inadequate safety measures usually ended up in a violation of work safety policies. Similar to other accident types, lack of communication also contributed to the inability to perceive external risks.

The results show that depending on the accident type, the combination of the antecedents and consequents was changing dramatically. Even though at first sight, the results seem familiar, considering so many types of accidents it is crucial to capture dynamically changing combinations to prevent severe accidents.

3.3. Contributions to body of knowledge

The proposed study introduces a new approach explicitly for the safety prevention process. The preventative measures defined in previous studies (Ayhan & Tokdemir, 2019a, 2019b, 2020) enlighten the future work. Figure 9 stands for visualizing the development of procedures for preventative measures in the most recent studies. The studies which promoted a proactive prevention proce-

dures were investigated, and divergence with the proposed model in terms of development took place in detail. The prevention section completes the prediction process in each study, and it was improved in every work. Initially, the prevention measures were established just concerning the severity level of construction incidents. The construction cases were classified into three groups, and preventative measures formed according to severity (Ayhan & Tokdemir, 2020). The second study supplemented a risk matrix which enables to rate risk factor of the working team as well as the company. After this time, personal experience about construction safety was also taken into consideration while setting preventative measures (Ayhan & Tokdemir, 2019c). Later, the preventative measures move ahead of this by introducing a fatal accident analysis. The fatal work events were examined without performing any analytical model to understand which attributes are influential in a fatality. As a result, attribute-based naïve results guide the preventative measures proactively (Ayhan & Tokdemir, 2019b).

However, accident analysis requires a more comprehensive investigation to reveal the root causes. The analytical models like ARM-based inferential systems have significant contribution to categorize and control the correlations between the attributes to discover the hidden relationships – the proposed model concerns every aspect of accidents, including severity level, personal information, and attributes. The Apriori algorithm overcomes problems in bulk data and mines a meaningful conclusion from it. Therefore, the ARM supported the deep level of understanding and allowed to set more effective preventative measures since the hidden relationships between the attributes were disclosed. Besides, the present model is capable of adapting the changes in data since it accomplished to obtain different characteristics regarding the accident type. That means the model can work dynamically with any system and can be part of a universal system that was started to be studied in previous studies (Ayhan & Tokdemir, 2019a, 2019b, 2020).

In addition, researchers have investigated accident attributes and the correlations between them via ARM (e.g. Liao & Perng, 2008; Cheng et al., 2010; Verma et al., 2014; Shin et al., 2018; Guo et al., 2019; Liao et al., 2019), but the types of accidents were not examined in detail. In other words, these studies did not focus on the factors relevant to the various accident types. Therefore, this study covered the most common construction accidents given in Figure 2. These accidents were analyzed individually to reveal the relationships between their attributes.

Ultimately, ARM, a data mining technique, was employed in order to investigate correlations between contributory factors of occupational accidents in construction sites collectively. 3870 construction accident cases, including first aid, medical intervention, lost workday cases, and fatalities were collected via the OHS professionals of different construction companies. ARM analysis determined the different attributes for each type of construction accident.

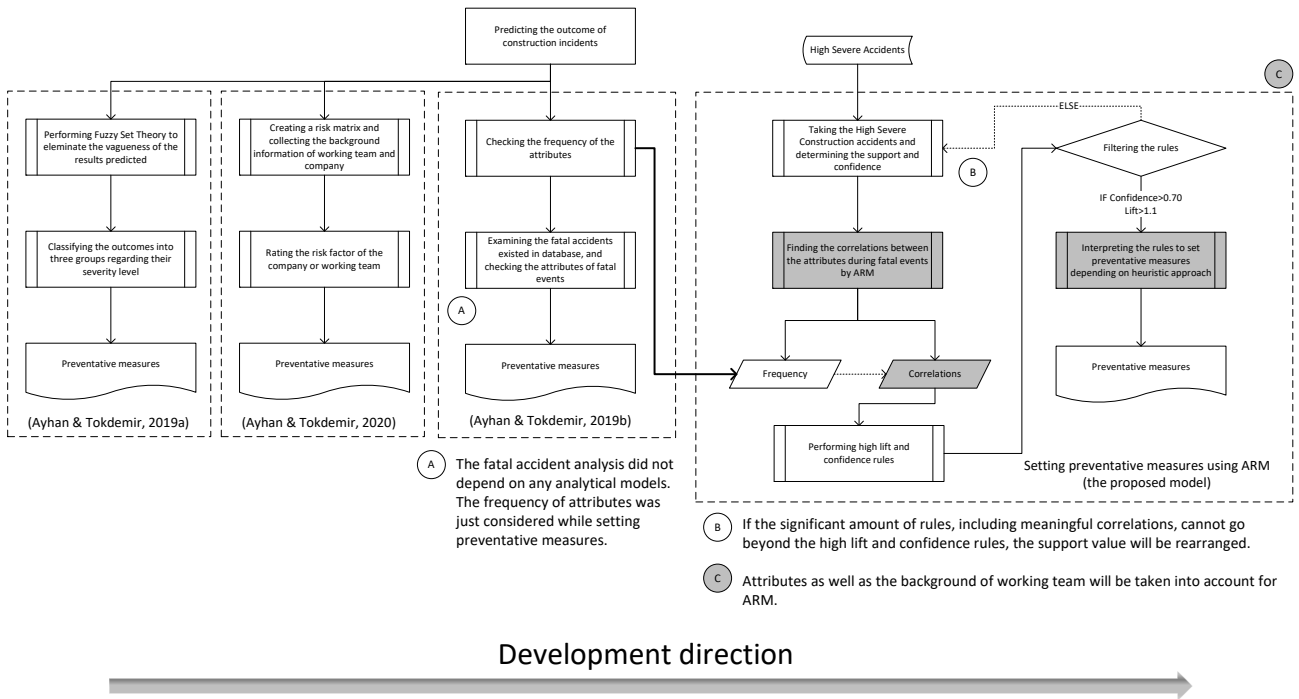


Figure 9. Contribution of the proposed model

Since the characteristics may vary depending on the accident type, different measures must be implemented to preventing casualties. This observation was also supported by analyzing the whole dataset (including every type of accident), and there was a distinct variation from the results of the individual analyses.

Conclusions

Occupational accidents are still a serious concern of the construction industry since many people working in construction sites suffer from the hazardous consequences of the accidents. Preventing accidents is becoming crucial, and the first milestone is to understand and evaluate the factors that lead to occupational accidents. The leading factors must be identified, and preventative measures must be taken in advance to avoid construction incidents. There is extensive research about construction safety in the literature, which has made significant contributions to safety. For example, Heinrich (1959) produced a domino theory which stated that the controllable unsafe acts lead to a dangerous level of incidents. Reason (1990) introduced a Swiss Case Model (SCM). SCM depicted that when the series of underlying conditions align together, it creates a path between the hazards to the accident. Although SCM is undoubtedly the most common accident causation model and Heinrich domino theory is being still widely used in industry, these models still need improvement while describing the complex interactions of components, i.e., attributes of incidents.

The accident analysis requires an in-depth level of investigation to find the root causes. Categorizing or controlling the frequency of attributes is not sufficient to extract the reason and take preventative measures. For

example, determining that the primary cause of accidents is the state of not being able to evaluate existing risk is predictable. If the investigation remains at this level, the real cause will not be discovered. For this reason, ARM can play a significant role. The ARM results also supported the deep level of understanding since assessing the construction accidents separately augmented the confidence values obtained from analyses as can be seen in the information presented in Tables 3 and 4.

Therefore, the proposed model introduced an adaptable analytical process that investigates the cause and effect relationships in occupational accidents in a more comprehensive way. OHS professionals can identify the cause-effect relationships instantly regardless of changes in data, and set up preventative actions to break the links between the accident elements, which were obtained by ARM for each common accident types. Ultimately, the present study could help OHS professionals to mitigate safety risks by showing them the relationships between attributes using statistics such as confidence, and these can serve as a guide or tool for specific accident prevention.

This study has certain limitations as follows. First, there was no guideline in the literature for the selection of threshold values for ARM analysis. The lower the selected values were; the higher the number of rules was generated. Thresholds can be re-arranged to obtain more or fewer rules according to the feedback from OHS professionals. As a future study, a fuzzy decision-making tool could be designed utilizing the results of the ARM analysis. Besides, the model was tested with only construction data and provided a good understanding of preventative studies, so it also needed to be tested for different industries, which would reveal a different combination of attributes to help professionals to prevent accidents.

References

- Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data (SIGMOD'93)* (pp. 207–216). <https://doi.org/10.1145/170035.170072>
- Aminbakhsh, S., Gunduz, M., & Sonmez, R. (2013). Safety risk assessment using analytic hierarchy process (AHP) during planning and budgeting of construction projects. *Journal of Safety Research*, 46, 99–105. <https://doi.org/10.1016/j.jsr.2013.05.003>
- Ayhan, B. U., & Tokdemir, O. B. (2019a). Predicting the outcome of construction incidents. *Safety Science*, 113, 91–104. <https://doi.org/10.1016/j.ssci.2018.11.001>
- Ayhan, B. U., & Tokdemir, O. B. (2019b). Safety assessment in megaprojects using artificial intelligence. *Safety Science*, 118, 273–287. <https://doi.org/10.1016/j.ssci.2019.05.027>
- Ayhan, B. U., & Tokdemir, O. B. (2020). Accident analysis for construction safety using latent class clustering and artificial neural network. *Journal of Construction Engineering and Management*, 146(3), 04019114. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001762](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001762)
- Başağa, H. B., Temel, B. A., Atasoy, M., & Yıldırım, İ. (2018). A study on the effectiveness of occupational health and safety trainings of construction workers in Turkey. *Safety Science*, 110, 344–354. <https://doi.org/10.1016/j.ssci.2018.09.002>
- Bavafa, A., Mahdiyar, A., & Marsono, A. K. (2018). Identifying and assessing the critical factors for effective implementation of safety programs in construction projects. *Safety Science*, 106, 47–56. <https://doi.org/10.1016/j.ssci.2018.02.025>
- Camino López, M. A., Ritzel, D. O., Fontaneda, I., & González Alcántara, O. J. (2008). Construction industry accidents in Spain. *Journal of Safety Research*, 39(5), 497–507. <https://doi.org/10.1016/j.jsr.2008.07.006>
- Chan, A. P. C., Javed, A. A., Lyu, S., Hon, C. K. H., & Wong, F. K. W. (2016). Strategies for improving safety and health of ethnic minority construction workers. *Journal of Construction Engineering and Management*, 142(9). [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001148](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001148)
- Chen, D., Xu, C., & Ni, S. (2017). Data mining on Chinese train accidents to derive associated rules. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 231(2), 239–252. <https://doi.org/10.1177/0954409715624724>
- Cheng, C. W., Lin, C. C., & Leu, S. S. (2010). Use of association rules to explore cause-effect relationships in occupational accidents in the Taiwan construction industry. *Safety Science*, 48(4), 436–444. <https://doi.org/10.1016/j.ssci.2009.12.005>
- Cheng, Y., Yu, W. Der, & Li, Q. (2015). GA-based multi-level association rule mining approach for defect analysis in the construction industry. *Automation in Construction*, 51, 78–91. <https://doi.org/10.1016/j.autcon.2014.12.016>
- Choi, B., Jebelli, H., & Lee, S. H. (2019). Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk. *Safety Science*, 115, 110–120. <https://doi.org/10.1016/j.ssci.2019.01.022>
- Cruz Rios, F., Chong, W. K., & Grau, D. (2017). The need for detailed gender-specific occupational safety analysis. *Journal of Safety Research*, 62, 53–62. <https://doi.org/10.1016/j.jsr.2017.06.002>
- Das, S., & Sun, X. (2014). Investigating the pattern of traffic crashes under rainy weather by association rules in data mining. In *Transportation Research Board 93rd Annual Meeting*. Washington DC, USA.
- Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., & Jalayer, M. (2018). Supervised association rules mining on pedestrian crashes in urban areas: identifying patterns for appropriate countermeasures. *International Journal of Urban Sciences*, 23(1), 30–48. <https://doi.org/10.1080/12265934.2018.1431146>
- DiDomenico, A., McGorry, R. W., Huang, Y. H., & Blair, M. F. (2010). Perceptions of postural stability after transitioning to standing among construction workers. *Safety Science*, 48(2), 166–172. <https://doi.org/10.1016/j.ssci.2009.07.006>
- Ding, L., Fang, W., Luo, H., Love, P. E. D., Zhong, B., & Ouyang, X. (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in Construction*, 86, 118–124. <https://doi.org/10.1016/j.autcon.2017.11.002>
- Dong, X. S., Choi, S. D., Borchardt, J. G., Wang, X., & Largay, J. A. (2013). Fatal falls from roofs among U.S. construction workers. *Journal of Safety Research*, 44(1), 17–24. <https://doi.org/10.1016/j.jsr.2012.08.024>
- Esmaeili, B., Hallowell, M. R., & Rajagopalan, B. (2015a). Attribute-based safety risk assessment. I: Analysis at the fundamental level. *Journal of Construction Engineering and Management*, 141(8), 04015021. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000980](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000980)
- Esmaeili, B., Hallowell, M. R., & Rajagopalan, B. (2015b). Attribute-based safety risk assessment. II: Predicting safety outcomes using generalized linear models. *Journal of Construction Engineering and Management*, 141(8), 04015022. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000981](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000981)
- Eteifa, S. O., & El-adaway, I. H. (2018). Using social network analysis to model the interaction between root causes of fatalities in the construction industry. *Journal of Management in Engineering*, 34(1), 04017045. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000567](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000567)
- Eurostat. (2015). *Accidents at work statistics*. https://ec.europa.eu/eurostat/statistics-explained/index.php/Accidents_at_work_statistics
- Evanoff, B., Dale, A. M., Zeringue, A., Fuchs, M., Gaal, J., Lipscomb, H. J., & Kaskutas, V. (2016). Results of a fall prevention educational intervention for residential construction. *Safety Science*, 89, 301–307. <https://doi.org/10.1016/j.ssci.2016.06.019>
- Fang, Q., Li, H., Luo, X., Ding, L., Rose, T. M., An, W., & Yu, Y. (2018). A deep learning-based method for detecting non-certified work on construction sites. *Advanced Engineering Informatics*, 35, 56–68. <https://doi.org/10.1016/j.aei.2018.01.001>
- Gao, R., Chan, A. P. C., Lyu, S., Zahoor, H., & Utama, W. P. (2018). Investigating the difficulties of implementing safety practices in international construction projects. *Safety Science*, 108, 39–47. <https://doi.org/10.1016/j.ssci.2018.04.018>
- Gerassis, S., Martín, J. E., Garcia, T. J., Saavedra, A., García, J. T., & Taboada, J. (2016). Bayesian decision tool for the analysis of occupational accidents in the construction of embankments. *Journal of Construction Engineering and Management*, 143(2), 04016093. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001225](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001225)
- Gerassis, S., Albuquerque, M. T. D., García, J. F., Boente, C., Giráldez, E., Taboada, J., & Martín, J. E. (2019). Understanding complex blasting operations: A structural equation model combining Bayesian networks and latent class clustering. *Reliability Engineering and System Safety*, 188, 195–204. <https://doi.org/10.1016/j.ress.2019.03.032>
- Geurts, K., Wets, G., Brijs, T., & Vanhoof, K. (2012). Profiling high frequency associations rules accident locations using as-

- sociation rules. *Transportation Research Record: Journal of the Transportation Research Board*, 1840(1), 123–130. <https://doi.org/10.3141/1840-14>
- Gephi 0.9.2. *An open source software for exploring and manipulating networks* (n.d.). <https://gephi.org/>
- Glinskiy, V., Serga, L., Khvan, M., & Zaykov, K. (2016). Fuzzy neural networks in the assessment of environmental safety. *Procedia CIRP*, 40, 614–618. <https://doi.org/10.1016/j.procir.2016.01.143>
- Grill, M., & Nielsen, K. (2019). Promoting and impeding safety – A qualitative study into direct and indirect safety leadership practices of constructions site managers. *Safety Science*, 114, 148–159. <https://doi.org/10.1016/j.ssci.2019.01.008>
- Guo, S., Zhang, P., & Ding, L. (2019). Time-statistical laws of workers' unsafe behavior in the construction industry: A case study. *Physica A: Statistical Mechanics and its Applications*, 515, 419–429. <https://doi.org/10.1016/j.physa.2018.09.091>
- Heinrich, H. (1959). *Industrial accident prevention*. New York: McGraw-Hill.
- Jebelli, H., Ahn, C. R., & Stentz, T. L. (2016). Fall risk analysis of construction workers using inertial measurement units: Validating the usefulness of the postural stability metrics in construction. *Safety Science*, 84, 161–170. <https://doi.org/10.1016/j.ssci.2015.12.012>
- Kaskutas, V., Dale, A. M., Lipscomb, H., & Evanoff, B. (2013). Fall prevention and safety communication training for foremen: Report of a pilot project designed to improve residential construction safety. *Journal of Safety Research*, 44(1), 111–118. <https://doi.org/10.1016/j.jsr.2012.08.020>
- Kazan, E., & Usmen, M. A. (2018). Worker safety and injury severity analysis of earthmoving equipment accidents. *Journal of Safety Research*, 65, 73–81. <https://doi.org/10.1016/j.jsr.2018.02.008>
- Kheni, N. A., Gibb, A. G. F., & Dainty, A. R. J. (2010). Health and safety management within small- and medium-sized enterprises (SMEs) in developing countries: Study of contextual influences. *Journal of Construction Engineering and Management*, 136(10), 1104–1115. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000218](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000218)
- Kim, Y. A., Ryoo, B. Y., Kim, Y.-S., & Huh, W. C. (2012). Major accident factors for effective safety management of highway construction projects. *Journal of Construction Engineering and Management*, 139(6), 628–640. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000640](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000640)
- Liao, C. W., & Chiang, T. L. (2016). Reducing occupational injuries attributed to inattentive blindness in the construction industry. *Safety Science*, 89, 129–137. <https://doi.org/10.1016/j.ssci.2016.06.010>
- Liao, C. W., & Perng, Y. H. (2008). Data mining for occupational injuries in the Taiwan construction industry. *Safety Science*, 46(7), 1091–1102. <https://doi.org/10.1016/j.ssci.2007.04.007>
- Liao, P. C., Chen, H., & Luo, X. (2019). Fusion model for hazard association network development: A case in elevator installation and maintenance. *KSCE Journal of Civil Engineering*, 23(4), 1451–1465. <https://doi.org/10.1007/s12205-019-0646-5>
- Lin, C.-L., & Fan, C.-L. (2018). Examining association between construction inspection grades and critical defects using data mining and fuzzy logic. *Journal of Civil Engineering and Management*, 24(4), 301–317. <https://doi.org/10.3846/jcem.2018.3072>
- Loosemore, M., & Malouf, N. (2019). Safety training and positive safety attitude formation in the Australian construction industry. *Safety Science*, 113, 233–243. <https://doi.org/10.1016/j.ssci.2018.11.029>
- Melo, R. R. S. de, Costa, D. B., Álvares, J. S., & Irizarry, J. (2017). Applicability of unmanned aerial system (UAS) for safety inspection on construction sites. *Safety Science*, 98, 174–185. <https://doi.org/10.1016/j.ssci.2017.06.008>
- Mistikoglu, G., Gerek, I. H., Erdis, E., Mumtaz Usmen, P. E., Cakan, H., & Kazan, E. E. (2015). Decision tree analysis of construction fall accidents involving roofers. *Expert Systems with Applications*, 42(4), 2256–2263. <https://doi.org/10.1016/j.eswa.2014.10.009>
- Mirabadi, A., & Sharifian, S. (2010). Application of association rules in Iranian Railways (RAI) accident data analysis. *Safety Science*, 48, 1427–1435. <https://doi.org/10.1016/j.ssci.2010.06.006>
- Mohammadi, A., Tavakolan, M., & Khosravi, Y. (2018). Factors influencing safety performance on construction projects: A review. *Safety Science*, 109, 382–397. <https://doi.org/10.1016/j.ssci.2018.06.017>
- Mohandes, S. R., & Zhang, X. (2019). Towards the development of a comprehensive hybrid fuzzy-based occupational risk assessment model for construction workers. *Safety Science*, 115, 294–309. <https://doi.org/10.1016/j.ssci.2019.02.018>
- Ning, X., Qi, J., & Wu, C. (2018). A quantitative safety risk assessment model for construction site layout planning. *Safety Science*, 104, 246–259. <https://doi.org/10.1016/j.ssci.2018.01.016>
- Olson, R., Varga, A., Cannon, A., Jones, J., Gilbert-Jones, I., & Zoller, E. (2016). Toolbox talks to prevent construction fatalities: Empirical development and evaluation. *Safety Science*, 86, 122–131. <https://doi.org/10.1016/j.ssci.2016.02.009>
- Patel, D. A., & Jha, K. N. (2014). Neural network approach for safety climate prediction. *Journal of Management in Engineering*, 31(6), UNSP 05014027. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000348](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000348)
- Patel, D. A., & Jha, K. N. (2016). Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model. *Safety Science*, 89, 240–248. <https://doi.org/10.1016/j.ssci.2016.06.020>
- Project Management Institute (PMI). (2008). *A guide to the project management body of knowledge (PMBOK Guide)* (4th ed.). Newtown Square, PA, USA.
- Rapidminer Studio 9.2.0. *Data science, machine learning, predictive analytics*. (n.d.). <https://rapidminer.com/>
- Reason, J., (1990). *Human error*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139062367>
- Rubio-Romero, J. C., Rubio, M. C., & García-Hernández, C. (2012). Analysis of construction equipment safety in temporary work at height. *Journal of Construction Engineering and Management*, 139(1), 9–14. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000567](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000567)
- RStudio. (2019). *Integrated development environment for R*. (Computer Software). <http://www.rstudio.com/>
- Schoenfisch, A., Lipscomb, H., Silverstein, B., Cameron, W., & Adams, D. (2014). Rates of and circumstances surrounding work-related falls from height among union drywall carpenters in Washington State, 1989–2008. *Journal of Safety Research*, 51, 117–124. <https://doi.org/10.1016/j.jsr.2014.09.007>
- Shao, B., Hu, Z., Liu, Q., Chen, S., & He, W. (2019). Fatal accident patterns of building construction activities in China. *Safety Science*, 111, 253–263. <https://doi.org/10.1016/j.ssci.2018.07.019>
- Shin, D. P., Park, Y. J., Seo, J., & Lee, D. E. (2018). Association rules mined from construction accident data. *KSCE Journal of Civil Engineering*, 22(4), 1027–1039. <https://doi.org/10.1007/s12205-017-0537-6>

- Stiles, S., Ryan, B., & Golightly, D. (2018). Evaluating attitudes to safety leadership within rail construction projects. *Safety Science*, 110, 134–144. <https://doi.org/10.1016/j.ssci.2017.12.030>
- Suárez-Cebador, M., Rubio-Romero, J. C., & López-Arquillos, A. (2014). Severity of electrical accidents in the construction industry in Spain. *Journal of Safety Research*, 48, 63–70. <https://doi.org/10.1016/j.jsr.2013.12.002>
- Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B., & Bowman, D. (2017). Construction safety clash detection: Identifying safety incompatibilities among fundamental attributes using data mining. *Automation in Construction*, 74, 39–54. <https://doi.org/10.1016/j.autcon.2016.11.001>
- Tokdemir, O. B., & Ayhan, B. U. (2019). The analysis of accidents with contact of sharp objects by using analytic hierarchy process and artificial neural networks. *DÜMF Journal of Engineering*, 10(1), 323–334. <https://doi.org/10.24012/dumf.466493>
- U.S Bureau of Labor Statistics (2017). *Injuries, illnesses, and fatalities*. <https://www.bls.gov/iif/>
- Verma, A., Khan, S. Das, Maiti, J., & Krishna, O. B. (2014). Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports. *Safety Science*, 70, 89–98. <https://doi.org/10.1016/j.ssci.2014.05.007>
- Weng, J., Zhu, J. Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis and Prevention*, 92, 43–52. <https://doi.org/10.1016/j.aap.2016.03.017>
- Wimer, B., Pan, C., Lutz, T., Hause, M., Warren, C., Dong, R., & Xu, S. (2017). Evaluating the stability of a freestanding Mast Climbing Work Platform. *Journal of Safety Research*, 62, 163–172. <https://doi.org/10.1016/j.jsr.2017.06.014>
- Winge, S. & Albrechtsen, E. (2018). Accident types and barrier failures in the construction industry. *Safety Science*, 105, 158–166. <https://doi.org/10.1016/j.ssci.2018.02.006>
- Winge, S., Albrechtsen, E., & Mostue, B. A. (2019). Causal factors and connections in construction accidents. *Safety Science*, 112, 130–141. <https://doi.org/10.1016/j.ssci.2018.10.015>
- Xu, C., Bao, J., Wang, C., & Liu, P. (2018). Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. *Journal of Safety Research*, 67, 65–75. <https://doi.org/10.1016/j.jsr.2018.09.013>
- Yao, Z., Deng, W., & Wu, D. (2018). Association rule analysis of contributory factors to severe traffic accidents. In *18th COTA International Conference of Transportation Professionals* (pp. 1886–1873). <https://doi.org/10.1061/9780784481523.186>
- Yiu, N. S. N., Sze, N. N., & Chan, D. W. M. (2018). Implementation of safety management systems in Hong Kong construction industry – A safety practitioner’s perspective. *Journal of Safety Research*, 64, 1–9. <https://doi.org/10.1016/j.jsr.2017.12.011>
- Zhang, P., Lingard, H., Blismas, N., Wakefield, R., & Kleiner, B. (2015). Work-health and safety-risk perceptions of construction-industry stakeholders using photograph-based Q methodology. *Journal of Construction Engineering and Management*, 141(5), 04014093. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000954](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000954)
- Zhang, X., & Liu, Z. (2011). Analysis of multi-dimensional association rule in marine casualties. In *First International Conference on Transportation Information and Safety (ICTIS)* (pp. 2697–2705). [https://doi.org/10.1061/41177\(415\)340](https://doi.org/10.1061/41177(415)340)
- Zhao, D., McCoy, A. P., Kleiner, B. M., Mills, T. H., & Lingard, H. (2016). Stakeholder perceptions of risk in construction. *Safety Science*, 82, 111–119. <https://doi.org/10.1016/j.ssci.2015.09.002>