Descriptive and Prescriptive Analysis of Construction Site Incidents Using Decision Tree Classification and Association Rule Mining

Ozgur Ugur
Department of Research and Development
ENKA SYSTEMS
Istanbul, Turkey
ozgur.ugur@enkasystems.com

Ali Atilla Arisoy
Department of Research and Development
ENKA SYSTEMS
Istanbul, Turkey
ali.arisoy@enkasystems.com

Berkay Bolac
Department of Research and Development
ENKA SYSTEMS
Istanbul, Turkey
berkay.bolac@enkasystems.com

Murat Can Ganiz
Department of Computer Sciences
Marmara University
Istanbul, Turkey
murat.ganiz@marmara.edu.tr

Abstract— **Learning from previous incidents is essential for preventing future incidents and taking the necessary precautions. We analyze construction site incidents by employing data science process and machine learning algorithms such as decision trees and Apriori. Patterns that are extracted using machine learning algorithms provides interesting insights on the causes of incidents and their relations with other factors. The dataset we use is a novel dataset containing hundreds of construction site incidents between 2014 and 2020 from an international construction company. The data consist of a wide range of features such as activity during incident, incident condition, hazard source, incident severity, location, and time. The decision tree is used in a descriptive analytics setting to extract patterns in our dataset. Additionally, the Apriori algorithm is employed to extract patterns in the form of frequent itemsets and association rules. The patterns we extract using machine learning algorithms shed light on associations between different factors and different types of incidents. One of the interesting results of our study is that the patterns extracted from a supervised classifier in a descriptive analytics setting collides with the patterns extracted using the unsupervised machine learning algorithm of Apriori. The generated rules can be used for informing the health and safety experts by developing a decision support mechanism for taking necessary precautions for minimizing the risk of different types of incidents.**

Keywords— construction site incident analysis, descriptive analytics, association analysis, decision tree, Apriori

I. INTRODUCTION AND BACKGROUND

The safety and health of the workers at the construction sites is a major health concern. Even though there are many technological developments in the construction sector, it is largely a manual and labor-intensive process. This makes it prone to incidents mainly from human error [1]. By recording past incidents accurately and in detail, it is possible to analyze the data and learn from past accidents [2].

The construction industry is a high-risk sector. The regulations stated by the health and safety regulators can be ignored in many different construction locations [3]. In a study conducted in Kuwait, it is revealed that there is a significant lack of information in the recorded construction site accidents [4]. Statistical studies provide many different informative results that can be utilized for preventing these accidents. A survey argues that the main causes of these accidents are lack of safety knowledge and negligence of the workers such as not following safety rules or fail to use personal protective equipment [5]. Additionally, unskilled worker employment, and poor site management are also listed as important factors. In another study, eight causes of construction site accidents are identified, such as failure to enforce safety measures, lack of necessary training, and contractors that neglect to fulfill safety requirements [6]. It is important to prevent future accidents by taking precautions based on detailed analysis of the historical data and make predictions to regulate safety settings and may be raise alarms if certain conditions are encountered.

In the literature, several studies utilize machine learning models for predicting incidents that may occur in a construction project. Several construction incidents between the years 2011 and 2016 are analyzed in [7]. The dataset consists of 140.169 incident victims. Logistic regression (LR), decision tree (DT), random forest (RF), and AdaBoost models are used to predict the likelihood of mortality of workers by analyzing injury and death data from previous fatal accidents. Among these models, RF performs best in terms of prediction accuracy. Another study utilized 2.8 million accidents that have occurred in Japan over twelve years period and trained a LR model [8]. In [9], RF and stochastic gradient tree boosting (SGTB) models are utilized to predict injury type, energy hazard source type, and injured body type. The constructed models perform successfully in terms of rank probability skill score and outperform parametric models that are found in the literature [10]. Another study focuses on high-risk accidents in the construction industry. They use a dataset that is collected between 2007 and 2011 [11]. They use several machine learning models including decision tree, and association rule mining. Findings show that there are significant correlations between time of accident, location of the accident, body part affected, final consequence of accident, and lost workdays. One of the interesting results they find is that the frequency of accidents during night shifts is smaller compared to other shifts. However, during the night shifts, injuries to the head, back, spine, and limbs occur more frequently compared to other shifts. These and many other findings are consistent with the previous studies in the literature [12] - [14]. A different study aims to identify workers who are at risk of an accident that can result in severe consequences and classify these workers to determine adequate control measures [15]. Their results provide a

978-1-6654-3603-8/21/$31.00 ©2021 IEEE

guideline for more efficient safety strategies, occupational accident prevention, and programs for emergency scenarios.

We analyze construction site incidents using decision tree classifier and association rule mining, the Apriori algorithm. Frequent itemsets and association rules are mined to get insight into the causes of a given incident type. Several sets of rules are extracted for the most frequent incident types to determine hidden patterns and associations between different factors affecting different types of incidents. The mined patterns can be used for informing the health and safety experts (HSEs) by developing a decision support mechanism for taking necessary precautions to minimize the risk of different types of incidents.

In the following section a brief description for the machine learning models is presented. Afterwards, in section 3, the properties of the used dataset and the preprocessing steps are explained. The experiment results are provided in section 4, along with the discussion of the results. Our paper is concluded with section 4, including outlook for the future work.

II. Approach

Association rule mining (ARM) analysis is mainly conducted for the identification of significant associations between items and patterns that occur frequently, in a large dataset. Apriori algorithm is one of the most common and frequently used methods for association rule mining analysis. In the Apriori algorithm, all itemsets that occur above a certain threshold, namely minimum support, are efficiently generated [16]. After these itemsets are generated, based on a second threshold, which can be either the confidence or the lift values, association rules are formed. Confidence is defined as the likelihood of a consequent, occurring based on a given antecedent. Lift value is the increase of the probability of a consequent occurring, given a certain antecedent. The "mlxtend" package available in Python can be used for association rules analysis [17]. The Apriori function in this package supports one-hot-encoded data frames. Since our dataset consist of nominal features and their values, we first convert our dataset into a one-hot encoded format in which each nominal feature value is represented as a column taking 0 or 1 based on the occurrence in a specific incident instance or row. From this aspect each column, representing a nominal feature value, corresponds to an item. On the other hand, non-zero values in each row or instance can be considered as an itemset.

Decision trees are one of the most popular classification algorithms in machine learning [18]. This algorithm can also be used in a descriptive analytics setting as it extracts patterns in the form of a decision tree or classification rules that is easy for humans to interpret.

III. Experimental Setup

The dataset used in this study has been obtained from ENKA's Global Health and Safety Management System (EHSE) [19]. ENKA is a global construction company, which employed about 200,000 people in its projects abroad, headquartered in Istanbul, Turkey. The features in the dataset are as follows:

- Incident type: there are 17 different incident types available in the dataset: incision, burn, electric shock, luxation or broken bone, sprain or strain, crushing, shortness of breath, internal organ injury, trauma, puncture, limb loss, allergy, soft tissue injury, exposure to ultraviolet rays and others.
- Cause of incident: In total there are 145 different causes defined in the database. These causes are separated into 4 different groups as follows; immediate action cause, immediate condition cause, root cause personal factors, and root cause job factors.
- Activity during the incident: there are 30 different activities defined that are in progress during the incident.
- Incident condition: the defined conditions are separated into two groups, weather and ground conditions. For weather and ground conditions there are 14 and 12 options available, respectively.
- Hazard source: 23 different options are defined in the database for identifying the source of hazard of the incident.
- Incident location: the occurred incident can be either inside or outside of a construction site.
- Incident severity: the severity of an incident is categorized into 3 groups; serious, fatal, and not severe.
- Incident time: the day and time of the incident are recorded. The time of day has been divided into 6 equal size periods, starting from 00:00 AM

In the preprocessing step, nominal attributes are transformed to binary numeric by utilizing a one-hot encoding method. The raw dataset has 259 columns and about a thousand rows, where each row represents a single incident instance. The distribution of the incident types can be seen in Fig. 1. From the figure, it can be observed that the distribution is highly skewed.

The most frequent incident types are "near miss", "first aid", "property damage", and "medical treatment" covering 89% of all incident records. Preliminary analysis shows that the performance of supervised and unsupervised algorithms improves when only these most frequent four types of incidents are considered. Therefore, in this study, the top four most frequent incident types that are mentioned above are used.

The parameters that are mentioned in the previous section are merged into a single data frame. One-hot encoding technique is used to merge all parameters of a single incident on a single row. However, by doing so the sparsity of the data frame increases substantially. The raw dataset has 259 nominal or numeric features. Some of these features will take all zeros or ones. Therefore, to decrease the dimension and sparsity of the data frame, these columns are eliminated before the analysis. Some of the weather conditions, such as hail, and certain values that exist as an option in the software in initial root cause and incident type attributes are never recorded in the database. Therefore, these one-hot encoded columns are composed entirely of zeros and eliminated.

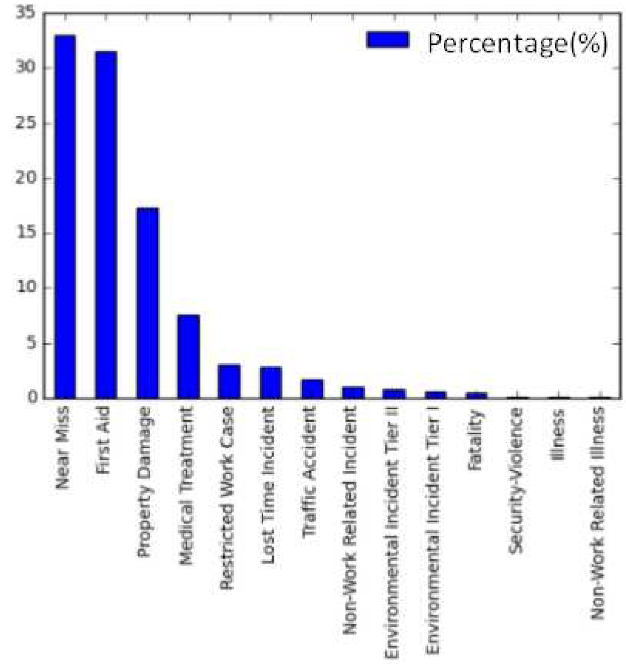

Fig. 1. Distribution of the incident types in the raw data set.

Finally, some low resolution and irrelevant features are eliminated from the dataset. These features are:

- Weather conditions: irrelevant and clear,
- Ground conditions: flat and dry,
- Incident severity: not severe and serious, and
- Incident location: inside the construction site.

Based on the preliminary analysis, these features that have high sparsity values do not provide any meaningful rules and are insignificant when classification models are developed, and thus are eliminated. The calculated sparsity value for the final dataset without the low-resolution features and the previous datasets are given in Table I.

TABLE I. SPARSITY VALUES OF THE RAW AND FILTERED DATASETS

Number of Features	Sparsity (%)
259	94.42
240	93.98
233	94.78

Descriptive analysis with 233 features and many values is not an easy task. Therefore, to reduce the complexity of the models and obtaining more easily interpretable patterns, the most influential features on the incident type are selected using the mutual information (MI) technique [20]. By calculating the MI, dependency between two given variables can be measured. If these selected variables are independent, then MI value will be zero. As dependency increases, the MI value also increases. MI is utilized as a feature selection method in a supervised setting. The incident type is selected as the class attribute. The relationship of each feature with the incident type is calculated. Highest MI values are presented in Fig. 2.

After sorting these values from largest to smallest, for different sized datasets, the mean score of MI is plotted and presented in Fig. 3.

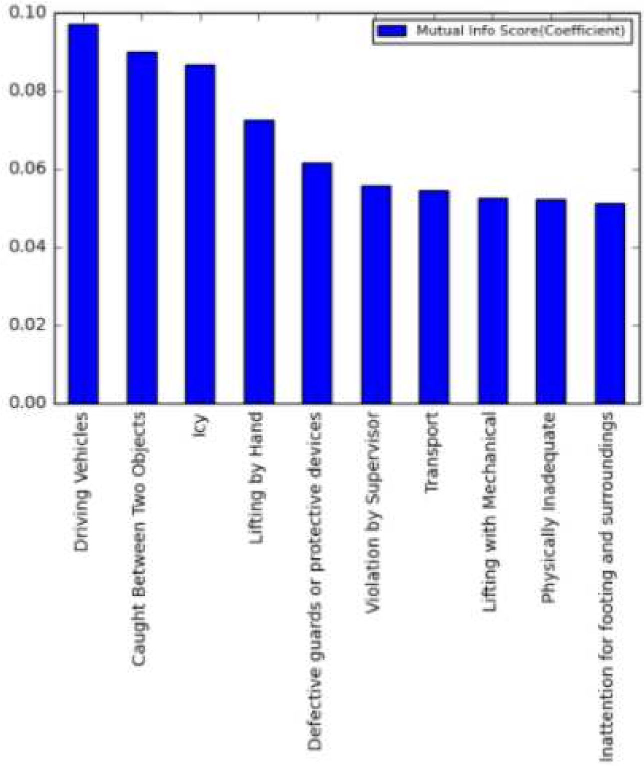

Fig. 2. Top 10 features with highest MI score.

Because the MI score decreases smoothly as the number of features in the dataset increases, different datasets with 20, 30, 40, 50, 60, and 70 features are constructed. The mean MI scores for these datasets are presented in Table II.

TABLE II. MEAN MI SCORES OF THE CONSTRUCTED DATASETS

Number of Features	Mean of MI Score
20	0.0546
30	0.0487
40	0.0445
50	0.0412
60	0.0383
70	0.0358

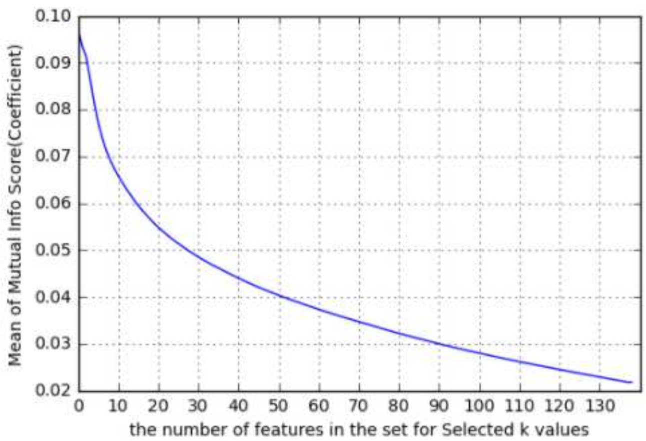

Fig. 3. Mean of MI score for the datasets with k-many features.

IV. Experiment Results and Discussion

A. Decision Tree Model

As mentioned in previous sections, incident type is the class attribute and used to train a decision tree model. J48 decision tree algorithm in Waikato Environment for Knowledge Analysis (WEKA) machine learning software is used for this purpose [17]. The default values of hyper-parameters are used, except for the confidence factor whose default value is changed from 0.25 to 0.01. As mentioned in the previous section, several datasets with different number of features are tested to observe which one performs the best. Based on the 10-fold cross-validation results, presented in Table III, the decision tree model that uses the dataset with 50 features as an input outperforms the other models.

The decision tree model is visualized in Fig. 4. According to this figure, the feature that provides the most information that is on the top of the tree is the hazard source of getting caught between two objects. On the right terminal nodes, it can be seen that if it is a serious accident then it will require medical treatment, else only a first aid will be necessary. This shows the decision tree model can learn common sense rules. We interpret this as an evidence for the soundness of the model for detecting patterns in incidents.

When we analyze the decision tree further, we can see that lack of health and safety expert (HSE) meetings, which is a job factor root cause, and driving vehicle, which is an activity during an incident, are features that are important for classifying for the four classes of incidents.

TABLE III. Decision Tree Models' Performance Summaries

Model Performance Criteria	20 Feature Dataset	30 Feature Dataset	40 Feature Dataset	50 Feature Dataset	60 Feature Dataset	70 Feature Dataset
Correctly classified instances (%)	57.90	59.53	59.38	60.08	58.78	59.08
Kappa statistic	0.3651	0.3908	0.3868	0.3623	0.3787	0.3861
Mean absolute error	0.2706	0.2577	0.2553	0.2571	0.2508	0.2398
Root mean squared error	0.3798	0.3764	0.3799	0.3911	0.3884	0.3927
Relative absolute error (%)	77.88	74.18	73.48	74.01	72.18	69.01
Root relative squared error (%)	91.07	90.33	91.16	93.85	93.22	94.24

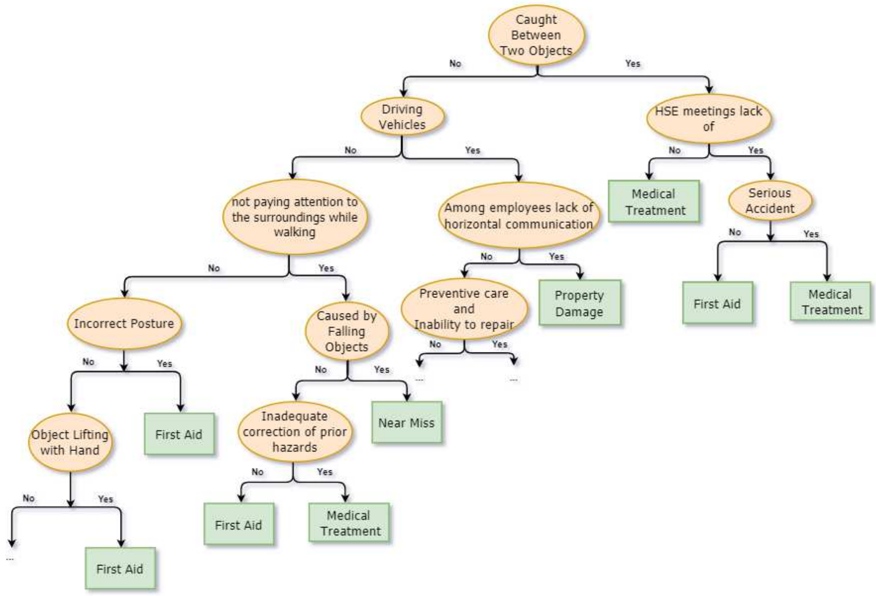

Fig. 4. Pruned decision tree model using 50 features.

B. Apriori Model

The objective of using ARM on the incident data is to detect interesting patterns in the form of itemsets or association rules. We mainly focus the frequent itemsets that includes an incident type and strong association rules that will conclude with an incident type. As mentioned, we use Apriori algorithm for this purpose. The dataset mentioned at the end of the "Data Preprocess" section is used as an input for this algorithm. A minimum support value of 2% is used in order to obtain as many rule combinations as possible. For extracting the association rules, the confidence parameter is set as 0.5. Consequently, Apriori generate 1799 association rules. The maximum number of items in antecedent and consequent of a rule is four and three, respectively. We apply several domain specific filters on the rules. For instance, association rules with antecedents containing incident types are eliminated. The total number of rules which has an incident type as a consequence is 131. The number of rules for "first aid", "near miss", and "property damage" type incidents is 82, 28, and 21, respectively. With the mentioned hyperparameter setting for the Apriori algorithm, no association rules for medical treatment type incidents are generated. This may be due to the scarcity of this type of incidents in the dataset.

The distribution of the features in the antecedents for the rule sets for "near miss" (NM), "first aid" (FA), "property damage" (PD), type incidents are presented in Table IV. Immediate action cause is the most frequent feature set for all types of incidents. However, no information about the incident location or severity can be detected for any type of incident.

TABLE IV. DISTRIBUTION OF FEATURES IN THE ANTECEDENTS FOR THE RULES OF FA, PD, AND NM TYPE INCIDENTS

Feature Sets in Antecedents	FA (%)	PD (%)	NM (%)
Immediate action causes	36.36	36.36	34.02
Personal factor root causes	22.50	22.50	20.50
Activity progress during incident	11.26	11.26	11.34
Hazard sources	9.53	9.53	9.62
Immediate condition causes	8.23	8.23	8.15
Incident time	4.33	4.33	5.31
Day of incident	3.46	3.46	3.89
Ground conditions	3.03	3.03	3.90
Weather conditions	0.87	0.87	0.71
Job factor root causes	0.43	0.43	2.56

In Table V, the most interesting rules for "first aid" type incidents, based on their lift values are presented. Based on these results it can be interpreted that the immediate action and condition causes, and personnel factor root causes have the highest influence on a "first aid" type of incident. It can also be interpreted that performing actions with less than sufficient attention and improper posture increases the potential of the occurrence of a first-aid type incident. The decision tree model is shown in Fig. 4 also supports this finding as well. HSE meetings should not be neglected so that risk of a severe accident that leads to a "first aid" or "medical treatment" type incident, can be minimized. Based on the association rule analysis, sufficient training must be provided for the worker so that they know proper posture for working and the awareness that working hastily can lead to injuries.

TABLE V. RULE SET FOR FIRST AID INCIDENT TYPE

Antecedents	Lift
● Immediate action cause: routine activity without thinking ● Immediate action cause: incorrect position/posture for work ● Personal factor root causes: training is not implemented in practice	2.53
● Immediate action cause: incorrect position/posture for work ● Personal factor root causes: employee perceived haste ● Immediate condition causes: congestion or restricted motion	2.46
● Personal factor root causes: employee perceived haste ● Day of incident: Tuesday ● Immediate condition causes: congestion or restricted motion	2.41
● Immediate action cause: incorrect position/posture for work ● Immediate action cause: violation by individual ● Personal factor root causes: employee perceived haste	2.34
● Immediate action cause: incorrect position/posture for work ● Immediate action cause: improper decision making or lack of judgment of the risk ● Personal factor root causes: training is not implemented in practice	2.34

The rule set obtained for "property damage" type of incidents is presented in Table VI. The lift values show that the rule set generated for "property damage" type incidents are much more important when compared with the rule set for "first aid" type incidents. Based on the results, it can be observed that driving a vehicle during the incident, increases the likelihood of a "property damage" type incident significantly. Additionally, the time of incident also seems to increase the likelihood of this type of incident as well. From the third row of Table VI, it can be interpreted that in the morning, a violation done by the worker who is driving a vehicle can lead to a "property damage" type incident. This can be the result of that worker not getting enough rest and therefore not being able to pay enough attention. HSEs should not allow the workers that are especially tired or cannot judge the risks or lack the knowledge of present hazards, to drive a vehicle, especially in the morning hours. HSEs also should inform the workers that individual violations and poor judgment skills can lead to even more severe consequences.

According to the most interesting rules for "near-miss" incidents, it can be said that the muddy ground condition is an important. Especially, if the worker is driving a vehicle on a muddy ground surface, then the likelihood of a "near-miss" type incident increases. Based on the root job cause factors, it can be interpreted that work planning must be done adequately to minimize the risk of an incident. Activities such as driving vehicles and mechanical lifting must be done much more carefully. Based on these outcomes, it can be interpreted that in case of poor ground conditions, the sufficiency of work planning becomes even more critical.

TABLE VI. RULE SET FOR PROPERTY DAMAGE INCIDENT TYPE

Antecedents	Lift
● Activity during incident: driving vehicle ● Immediate action cause: routine activity without thinking ● Immediate action cause: violation by individual	4.02
● Activity during incident: driving vehicle ● Immediate action cause: violation by individual ● Immediate action cause: Lack of knowledge of present hazards	3.99
● Activity during incident: driving vehicle ● Immediate action cause: violation by individual ● Incident hour: 08:00AM-12:00PM	3.47
● Activity during incident: driving vehicle ● Immediate action cause: Improper decision making or lack of judgment of the risk ● Personal factor root causes: poor judgment skills	3.21
● Activity during incident: driving vehicle ● Immediate action cause: violation by individual ● Immediate action cause: Improper decision making or lack of judgment of the risk	3.12

TABLE VII. RULE SET FOR NEAR MISS INCIDENT TYPE

Antecedents	Lift
● Ground condition: muddy ● Root cause job factors: insufficient work planning	2.55
● Ground condition: muddy ● Activity during incident: driving vehicle	2.27
● Ground condition: muddy ● Immediate action cause: violation by individual	2.06
● Activity during incident: lifting with mechanics ● Incident severity: serious	1.98
● Root cause job factors: inadequate workplace layout ● Root cause job factors: insufficient work planning	1.94

V. CONCLUSION

This study aims to develop a decision support mechanism for health and safety experts (HSEs) so that necessary precautions can be taken to minimize the risk of different types of incidents. To achieve this aim DT and Apriori algorithms are utilized in a descriptive analytics setting to analyze important patterns in construction site incident data. The four most frequently recorded incident types cover 89% of the dataset, therefore, these incident types are used.

Using 50 features a decision tree models is constructed and an accuracy up to 60% is obtained. The patterns observed at the end of decision tree modeling and Apriori analysis show overlapping results. Using these patterns HSEs can be informed so that the necessary precautions can be implemented to minimize the risk of potential incidents and losses. According to the discussions conducted with the ENKA's Health and Safety Department the outcomes of these results are reliable and can be generalized. Additionally, the department experts also declared that these outcomes can be utilized to predict the potential risk of a given type of incident.

The future work will include prediction of the potential risk of the investigated factors for the given incident type using different machine learning algorithms and provide insight for the site experts to take necessary precautions to minimize the risk. By coordinating with the HSE experts, the outcomes will be evaluated, and field implementation strategies will be discussed.

ACKNOWLEDGMENTS

This study is supported by ENKA Systems R&D Center under the "Multinational Occupational Health, Safety and Environmental Management Software for High Hazard Class" project (Project code: HSE) within the scope of 2020 the General Directorate of R&D Incentives of the Ministry of Industry and Technology operating cycle. Additionally, the authors would like to present their gratitude to the ENKA's Health and Safety Department for their sharing their expertise, and to ENKA Systems R&D Center researchers for their support.

REFERENCES

[1] Z. Zhou, Y. M. Goh, and Q. Li, "Overview and analysis of safety management studies in the construction industry," Safety Science, vol. 72, pp. 337–350, February 2015.

[2] D. K. Chua and Y. M. Goh, "Incident Causation Model for Improving Feedback of Safety Knowledge," Journal of Construction Engineering and Management, vol. 130, no. 4, pp. 542–551, August 2004.

[3] M. Elsebaei, O. Elnawawy, A. A. Othman, and M. Badawy, "Causes and impacts of site accidents in the Egyptian construction industry," International Journal of Construction Management, pp. 1–12, September 2020.

[4] N. A. Kartam and R. G. Bouz, "Fatalities and injuries in the Kuwaiti construction industry," Accident Analysis & Prevention, vol. 30, no. 6, pp. 805–814, November 1998.

[5] A. R. A. Hamid, M. Z. A. Majid, and B. Singh, "Causes of accidents at construction sites," Malaysian Journal of Civil Engineering, vol. 20, pp. 242–259, 2008.

[6] B. L. H. Chim, C. J. Chun, and F. K. Wah, "Accidents in construction sites: A study on the causes and preventive approaches to mitigate accident rate," INTI Journal, vol. 1, pp. 1–12, 2018.

[7] J. Choi, B. Gu, S. Chin, and J.-S. Lee, "Machine learning predictive model based on national data for fatal accidents of construction workers," Automation in Construction, vol. 110, p. 102974, February 2020.

[8] T. Nishimoto, K. Mukaigawa, S. Tominaga, N. Lubbe, T. Kiuchi, T. Motomura, and H. Matsumoto, "Serious injury prediction algorithm based on large-scale data and under-triage control," Accident Analysis & Prevention, vol. 98, pp. 266–276, January 2017.

[9] A. J. P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, "Application of machine learning to construction injury prediction," Automation in Construction, vol. 69, pp. 102–114, September 2016.

[10] B. Esmaeili, M. R. Hallowell, and B. Rajagopalan, "Attribute-Based Safety Risk Assessment. II: Predicting Safety Outcomes Using Generalized Linear Models," Journal of Construction Engineering and Management, vol. 141, no. 8, p. 04015022, August 2015.

[11] M. Amiri, A. Ardeshir, M. H. Fazel Zarandi, and E. Soltanaghaei, "Pattern extraction for high-risk accidents in the construction industry: a data-mining approach," International Journal of Injury Control and Safety Promotion, vol. 23, no. 3, pp. 264–276, May 2015.

[12] B. J. M. Ale, L. J. Bellamy, H. Baksteen, M. Damen, L. H. J. Goossens, A. R. Hale, M. Mud, J. Oh, I. A. Papazoglou, and J. Y. Whiston, "Accidents in the construction industry in the Netherlands: An analysis of accident reports using Storybuilder," Reliability Engineering & System Safety, vol. 93, no. 10, pp. 1523–1533, October 2008.

[13] G. H. Halvani, R. Jafarinodoushan, S. J. Mirmohammadi, and A. M. Mehrparvar, "A survey on occupational accidents among construction industry workers in Yazd city: Applying Time Series 2006-2011," Journal of Occupational Health and Epidemiology, vol. 1, no. 1, pp. 1–8, 2012.

[14] H. J. Im, Y. J. Kwon, S. G. Kim, Y. K. Kim, Y. S. Ju, and H. P. Lee, "The characteristics of fatal occupational injuries in Korea's construction industry, 1997–2004," Safety Science, vol. 47, no. 8, pp. 1159–1162, October 2009.

[15] S. S. Alizadeh, S. B. Mortazavi, and M. Mehdi Sepehri, "Assessment of accident severity in the construction industry using the Bayesian theorem," International Journal of Occupational Safety and Ergonomics, vol. 21, no. 4, pp. 551–557, December 2015.

[16] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," Proceedings of the 1993 ACM SIGMOD international conference on Management of data - SIGMOD '93, 1993.

[17] S. Raschka, "MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack," Journal of Open Source Software, vol. 3, no. 24, p. 638, April 2018.

[18] .N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on J48 algorithm for data mining," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 3, pp. 1114–1119, June 2013.

[19] "EHSE - Global HSE Management System (Health, Safety, Environment) ENKA Systems," ENKA Systems Yazilim A.S. [Online]. Available: https://www.enkasystems.com/solutions/ehse-global-hse-management-system/ (Accessed: 20-Apr-2021).

[20] E. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, 1948