Journal of Human, Environment, and Health Promotion. 2019; 5(3): 121-126

Journal of Human, Environment, and Health Promotion

Journal homepage: www.zums.ac.ir/jhehp

Accident Modeling in Small-Scale Construction Projects Based on Artificial Neural Networks

Behrouz Alizadeh Savareh [a] | Mohsen Mahdinia [b] | Samira Ghiyasi [c] | Jmshid Rahimi [d] | Ahmad Soltanzadeh [b*]

[a] Department of Medical Informatics, School of Management and Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[b] Department of Occupational Safety and Health Engineering, Health School and Research Center for Environmental Pollutants, Qom University of Medical Sciences, Qom, Iran.
[c] Department of Environmental Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran.
[d] Department of Occupational Safety and Health Engineering, Health School, Alborz University of Medical Sciences, Karaj, Iran.

***Corresponding author: Ahmad Soltanzadeh**
Department of Occupational Safety and Health Engineering, Health School and Research Center for Environmental Pollutants, Qom University of Medical Sciences, Qom, Iran. Postal code: 3715614566.
E-mail address: soltanzadeh.ahmad@gmail.com

ARTICLE INFO

Article type:
Original article

Article history:
Received 8 June 2019
Revised 30 July 2019
Accepted 26 August 2019

DOI: 10.29252/jhehp.5.3.5

Keywords:
Construction projects
Accident
Modeling

ABSTARACT

Background: Several factors contribute to accidents in small-scale construction projects (SSCPs). The present study aimed to assess the influential factors in SSCP accidents and introduce a model to predict their frequency.
Methods: In total, 38 SSCPs were within the scope of this investigation. The safety index of accident frequency rate (AFR) causing 452 injury construction accidents during 12 years (2007-2018) was analyzed and modeled. Data analysis was performed based on feature selection using Pearson's χ^2 coefficient and SPSS modeler, as well as the artificial neural networks (ANNs) in MATLAB software.
Results: Mean AFR was estimated at 26.32 ± 14.83, and the results of both approaches revealed that individual factors, organizational factors, training factors, and risk management-related factors could predict the AFR involved in SSCPs.
Conclusion: The findings of this research could be reliably applied in the decision-making regarding safety and health construction issues. Furthermore, Pearson's correlation-coefficient and ANN modeling are considered to be reliable tools for accident modeling in SSCPs.

1. Introduction

The assessments performed in small-scale construction projects (SSCPs) are associated with significant challenges in terms of safety and health performance problems, which is mainly due to the multidimensionality of safety and health (S&H) issues, which often yield unreliable outcomes [1,2]. In the past two decades, more than 26,000 individuals employed in construction companies have been reported to die in the United States, which accounts for an average of five deaths per day [3]. In addition, reports from the European Union have revealed that 30 million working days are lost due to construction accidents [4]. Although only 12% of Iranian workers are active in the SSCPs, statistics uncover the high rate of the accidents faced by this population [5, 6]. Therefore, special attention should be paid to the recognition of the most important influential factors in such accidents in an attempt to find/model the correlations between accidents, as well as the contributing factors in this regard.

Accident analysis and modeling is primarily aimed at determining the main factors that control accidents [7]. Accident frequency rate (AFR) is a basic index to quantitatively analyze occupational accidents [8].

How to cite: Alizadeh Savareh B, Mahdinia M, Ghiyasi S, Rahimi J, Soltanzadeh A. Accident Modeling in Small-scale Construction Projects Based on Artificial Neural Networks. *J Hum Environ Health Promot.* 2019; 5(3): 121-6.

Identification of AFR-associated factors is of utmost importance in analyzing construction accidents and proffering the optimal strategies for the reduction and prevention of accidents [9].

Researchers have proposed various models to determine and analyze the factors involved in accidents and possibly predict construction accidents [10-13]. For instance, Mitropoulos et al. developed a model for the analysis of accidents, predicting the correlations between accidents, and identifying the infleuntial factors in construction accidents [14].

Artificial neural networks (ANNs) are among intelligent stochastic approaches, which are considered to be practical tools for the prediction of accidents. In a study in this regard, ANN methodology was reliably applied in the modeling and prediction of the contributing factors to accidents [15].

According to the literature, data is unclear regarding the correlation between the influential and predictive factors involved in AFR, as well as the use of findings in the preventive and strategic plans proposed to reduce construction accidents. The present study aimed to determine the AFR-related factors involved in SSCPs and use them to construct an intelligent model based on ANN for the prediction of AFR. With the complementary use of Pearson's χ^2 coefficient and SPSS, it was attempted to propose a practical solution for the safety and health problems in SSCPs.

2. Materials and Methods

This analytical study was conducted on 38 Iranian SSCPs recorded during 2007-2018. The projects with ≤35 workers were considered as small-scale construction projects.

2.1. Data Collection

Data were collected from the local documents of the SSCPs. In total, more than 500 documents on accidents, which were registered within 12 years, were selected for the modeling and statistical analyses. After the meticulous screening of the documents, incomplete data were eliminated. Finally, data of 452 accidents that caused injuries in the workers were selected for further investigation.

2.2. Determination of Factors

The collected data were classified into two categories of independent and dependent factors.

2.2.1. Independent Factors

The following data categories were defined as the independent factors in this study: Individual factors (IFs), including age, work experience, and education level of the injured workers;

Organizational factors (OFs), such as the type of job, type of activity, and average number of workers;

Training factors (TFs), including pre-employment, periodic, and post-accident factors, personal protective equipment (PPE), housekeeping, duration of training, and training contents;

Risk management factors (RMFs), such as the establishment of risk management systems, incident investigation, hazard identification studies (HAZID), periodic risk assessments, risk control, S&H reporting systems, tool box meetings (TBMs), housekeeping, and S&H audit and inspection.

2.2.2. Dependent Factors

AFR was used as the dependent factor and considered as the output layer of the developed ANN. AFR is an index used for the quantitative analysis of occupational accidents, which has been applied in several investigations to quantify the S&H performance [8]. In this study, AFR was calculated using equation 1, as follows [16]:

$$AFR = \frac{\text{Total number of accidents} \times 200000}{\text{Total working hours}} \quad (1)$$

2.3. Feature Selection

Actual problems involve a large number of explanatory variables, which govern one or more outputs. Therefore, substantial time and effort should be dedicated to finding new tools to manage large volumes of data, which might be associated with various challenges, such as the curse of dimensionality. In order to overcome such limitations, the algorithms of feature selection were used in our analysis so as to identify the most influential characteristics in this regard [17]. In this study, feature selection was performed using the IBM SPSS Modeler 14.2 as one of the most potent data mining software. Since the entry included a wide variety of data (e.g., continuous, nominal, flag, and ordinal data), Pearson's χ^2 coefficient was used to assess their differences. In addition, the significant cutoff point for feature selection was considered to be 0.95 [18].

2.4. Artificial Neural Networks (ANNs)

ANNs are computational tools that are capable of solving numerous problems in various fields. They are a network of artificial neurons and a system based on the biological neural network, which emulates the behavior of the system. The high complexity of natural neurons is summarized in a computational mathematical model assigned to artificial neurons [19]. ANN uses the artificial neurons with a combination of capabilities for the sensing of the process behavior as inspired by the pattern that is based upon experimental data.

It has been well documented that neural networks are capable of data classification, prediction, approximation, and clustering. As intelligent tools, ANNs are widely applied in predicting the behavior of various systems through modeling the medical diagnostics, sales forecasting, control over industrial activities, and research on customers, data validation, and risk management. Particularly, ANNs have been occasionally applied in the modeling and prediction of the incidence and severity of complex accidents in industries in order to determine the influential factors [15]. Multilayer perceptron (MLP) has numerous merits in the modeling and prediction of the events that are based on ANN models. The MLP network receives the inputs, analyses them with the aid of one or more hidden layers, and correlates them with the outputs. As a result, the MLP network learns to emulate the correlation between the input and output variables through adjusting the weights and biases in an iterative manner until reaching the permissible error. Back-propagation is considered to be the most commonly used algorithm to obtain parameters, which propagates the inputs by taking forward step in the network and calculating the error in a backward manner iteratively [19, 20].

Design of the neural network architecture in an ANN structure strongly affects the accuracy of the modeling. Normally, the MLP structure consists of an input layer, an output layer, and one or two intermediate of hidden layers. Accordingly, various structures could be achieved depending on the number and arrangement of the hidden layers. In this regard, Linoff and Berry defined some instructions so as to find the optimal MLP structure. According to the literature, neural networks have been developed for the prediction of accidents, and the optimal structure of the neural networks was attained to meet the objectives of the current research. It was concluded that one or two hidden layers, neurons 1-3 times (×3) of the input neurons could be sufficient to rely on this modeling.

Accordingly, the proposed structures were implemented using the MATLAB software [21].

Moreover, various models that were obtained in this study were compared in order to attain the optimal model based on the MSE criterion, as follows:

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 \qquad (2)$$

where MSE is the mean square error, y_i shows the network output of the i^{th} data, $\tilde{y}_i$ is the desired output of the i^{th} data, and n represents the number of the dataset.

3. Results and Discussion

3.1. Descriptive Statistics

The results of the statistical studies are presented in table 1. The mean AFR value was calculated to be 26.32 ± 14.83, and the age and work experience of the injured workers were estimated at 29.88 ± 7.70 and 4.98 ± 3.97 years, respectively. More than 70% of the injured individuals were construction workers, and 25% were technicians. Meanwhile, 51.4% of the injured workers had pre-employment S&H training, 24.1% gained periodic S&H training, and 19.4% passed S&H training after an accident.

The proportion of the risk management factors (RMFs), including HAZID, periodic risk assessment, and S&H control measures, was determined to be 18.1%, 18.5%, and 11.3%, respectively.

3.2. Feature Selection

Feature selection was performed using Pearson's χ^2 coefficient, indicating that IFs (age, work experience, and education level), OFs (activity type and average number of workers), TFs (duration, content, and periodic training), and RMFs (HAZID, periodic risk assessment, risk control, and housekeeping) were the most important contributing factors to accidents. Additionally, these factors were qualified to the ANN in order to analyze and predict the influential factors in AFR (Table 2).

Table 1: Descriptive Results of Influential Factors in Construction Accidents

Factors		Values
Individual Factors (IFs)		
Age (Years) (M ± SD)		29.88 ± 7.70
Work Experience (M ± SD)		4.98 ± 3.97
Education	Sub Diploma	162 (35.8%)
	Diploma	203 (44.8%)
	Academic	88 (19.4%)
Organizational Factors (OFs)		
Average of workers (M ± SD)		95.50 ± 77.63
Type of Job	Construction Workers	337 (74.4%)
	Technicians	104 (23.0%)
	Drivers	12 (2.6%)
Activity Type	Construction Work	312 (68.9%)
	Mechanical	31 (6.8%)
	Installation	102 (22.5%)
	Electricity	8 (1.8%)
S&H training factors (TFs)		
Pre-employment training		233 (51.4%)
Periodic Training		109 (24.1%)
After Accident Training		88 (19.4%)
PPE Training		95 (21.0%)
Housekeeping Training		15 (3.3%)
Duration of Training		131 (28.9%)
Content of Training		70 (15.5%)
Risk Management Factors (RMFs)		
Establishment of risk management system		92 (20.3%)
Incident/accident investigation		83 (18.4%)
HAZID		82 (18.1%)
Periodic risk assessment		84 (18.5%)
Risk control measures		51 (11.3%)
S&H reporting system		48 (10.6%)
Toolbox meeting		52 (11.5%)
Housekeeping		29 (6.4%)
S&H Checklist		262 (57.8%)
S&H audit and inspection		41 (9.1%)

Table 2: Results of Feature Selection to Determine Influential Factors in Construction Accidents

Selected factors	Value & importance rate
Risk control	0.998
Education	0.997
HAZID	0.996
Age	0.995
Job Experience	0.989
Housekeeping	0.987
Periodic risk assessment	0.982
Content of training	0.975
Duration of training	0.972
Activity type	0.966
Periodic training	0.963
Average of workers	0.955

3.3. ANN Modelling

Based on the applied method in the design of the neural network structures, 1,331 different structures of the neural networks were achieved.

For 1^{st} hidden layer: 1 → 3×number of input factors → 36 networks

For 2^{nd} hidden layer: 0 → 3×number of input factors → 37 networks

Total networks: 36 × 37 = **1331 networks**

Following that, the networks were compared based on the mean square error (MSE) criterion, and the screening of

1,331 simulated networks resulted in the selection of five optimum neural networks (Table 3).

According to the findings, the lowest MSE was measured to be 0.0049, corresponding to the optimal neural network that enabled the prediction of the nonlinear behavior of the system (Figure 1-4). Based on the outcomes of the ANN model (Table 2; Figure 1), the factors were properly predicted by the developed ANN model, which was fed by the outputs of feature selection.

Figure 2 shows the MSE values and number of the iterations (epochs) in training, indicating that the designed ANN could appropriately mimic the AFR value as learned from the developed model. Figure 3 depicts the histogram of error between the actual AFR values and predicted values based on the ANN model. Accordingly, the absolute minimum error value was 0.02 from one point to another, where a slight absolute error value could be detected as well. Therefore, the histogram analysis demonstrated that the samples experiencing the least error in the network were trained properly.

As is shown in figure 4, the regression value (R), which was defined as the measure of the correlations between the actual outputs of AFR and ANN predictions, was 0.8843; the obtained value was reasonable considering the diversities in the factors and their variation intervals. These findings confirmed the reliability and significance of the developed ANN in predicting the AFR value by receiving the mentioned input variables.

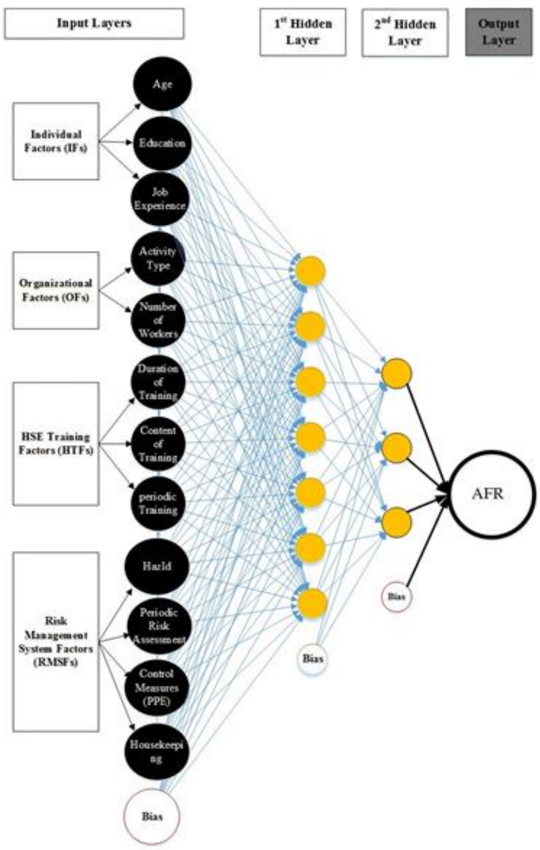

Figure 1: Optimum ANN Structure of AFR in Construction

Table 3: MSE Statistical Parameters

Opt. Network	Test MSE	Training MSE	The number of neurons of the 1st Hidden Layer	The number of neurons of the 2nd Hidden Layer
1	0.0049	0.0094	7	3
2	0.0067	0.0095	11	10
3	0.0069	0.0101	19	12
4	0.0073	0.0099	21	24
5	0.0080	0.0085	17	31

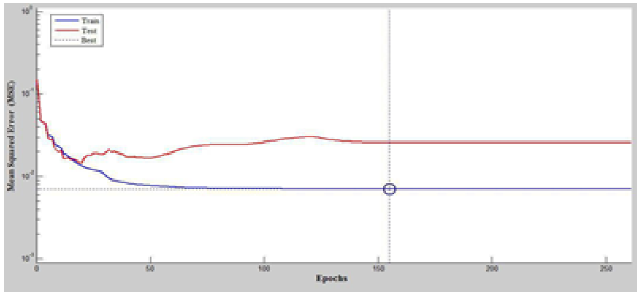

Figure 2: MSE versus Learning

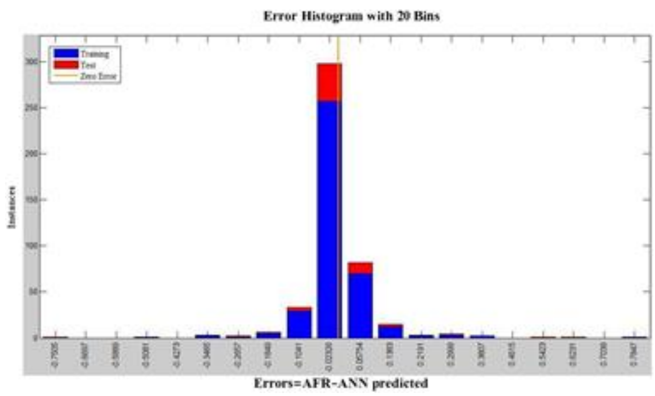

Figure 3: Error Histogram

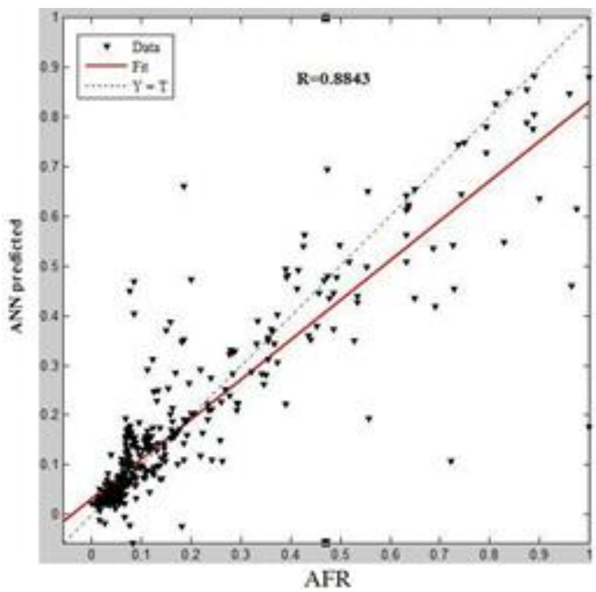

Figure 4: Correlations between ANN Predictions and AFR Actual Values

According to the results of feature selection and ANN model, 12 out of 23 studied factors were the optimal predictors of AFR in SSCPs. In general, the modeling process encompasses several factors, such as IFs, OFs, S&H TFs, and RMFs.

Previous investigations have indicated that construction accidents could be influenced by several factors [10,22]. In the present study, the statistical analysis and stochastic modeling revealed that construction accidents could be properly estimated with relatively high degrees of reliability and perception in terms of the contributing factors. Such insights could help develop predictive algorithms and intelligent models for the prediction of construction accidents in the future. Due to the multifaceted complexities associated with the recognition of all the aspects of modern constructive industries, the analysis of construction accidents using common analytical methods for accidents remains a matter of debate [22,23]. As such, some studies have denoted that the application of ANN modeling could be a promising strategy for the prediction of accidents in construction fields and determining the most significant influential factors in such accidents [15,24,25].

According to the results of the present study, IFs and OFs could be considered as direct, mediating influential factors in various S&H problems, such as construction accidents [12,22,26]. Furthermore, the ANN model revealed that the age, work experience, and education level of the injured workers as the IFs and average number of workers and activity type as the OFs could predict AFR in SSCPs.

Workplace hazards are of crucial importance, and various strategies have been implemented to prevent or minimize the probability of accidents. S&H training plays a key role in decreasing the rate of accidents, and inadequate training has been reported to be a major cause of accidents [11,25,27]. The findings of the current research demonstrated that some TFs could contribute to construction accidents, such as periodic training and duration and the contents of training.

SSCPs are associated with S&H risks [13,28], and the results of ANN modeling in the present study indicated that some RMFs could also predict AFR, including HAZID, periodic risk assessment, development and implementation of risk control measures, and housekeeping. Moreover, some researchers have reported that the most important influential factors in SSCP accidents are the weaknesses of the S&H risk management systems, such as the deficiencies in HAZID, risk assessment, and risk control [22,29,30]. Some of the limitations of this analytical study were the small number of the considered accidents-related factors in the model and geographical sprawl of the SSCPs. It seems that the development and implementation of a comprehensive study based on large data banks on accidents and the related factors is necessary in order to generalize the model to more cases and achieve more reliable and comprehensive data.

4. Conclusion

According to the results, the ANN model revealed that stochastic modeling is an effective and reliable tool for the analysis and prediction of accidents in SSCPs. Furthermore, the findings of the research confirmed that the ANN model of AFR provided a valid solution to the problem of occupational accident analysis. It is believed that this methodology could be applied in the preventive measures in this regard. The reinterpretation of the findings also confirmed that AFR could be predicted with 88.43% precision. Therefore, it could be inferred that the modeling of the AFR based on ANN enables the defining and developing of new patterns to promote the optimum performance of S&H structures in SSCPs.

On the other hand, the obtained results suggested that changes in the AFR do not necessarily occur due to a single factor, but they may occur when two or more factors are simultaneously manipulated. As such, preventive measures should be focused on minimizing the limitations of the current procedure in S&H management systems; therefore, a comprehensive investigation is required to enhance the training process in terms of the individual and organizational factors.

Authors' Contributions

This article was carried out by all the authors. A.S., and B.A.S, designed the manuscript and contributed to carry out data collection and data analysis and M.M., and J.R., wrote the manuscript.

Conflict of Interest

The Authors declare that there is no conflict of interest.

Acknowledgments

Hereby, we extend our gratitude to the Vice Chancellor of Research and Technology at Qom University of Medical Sciences and Health Services for the partial support of this study (Project No. 971000). We would also like to than the mangers and safety engineers of the studied large construction projects for their invaluable and skillful assistance in this research project.

References

1. Farshadnia SM, Arghami S, Shahab SA, Majidi F. A Budgeting Model for the Safety Unit of an Under ConstructionMetro Station in Tehran Using a Robust Optimization. *J Hum Environ Health Promot*, 2018; 4(3): 131-7.

2. Soltanzadeh A, Mohammadfam I, Moghimbeygi A, Ghiasvand R . Exploring Causal Factors on the Severity Rate of Occupational Accidents in Construction Worksites. *Int J Civ Eng*, 2017; 15(7): 959-65.

3. Hinze JW, Teizer J. Visibility-Related Fatalities Related to Construction Equipment. *Saf Sci*. 2011; 49(5): 709-18.

4. Kines P, Spangenberg S, Dyreborg J. Dyreborg, Prioritizing Occupational Injury Prevention in the Construction Industry: Injury Severity or Absence? *J Saf Res*, 2007; 38(1): 53-8.

5. Mohammadfam I, Soltanzadeh A, Mahmoudi S, Moghimbeigi A. P154 Analytical Modelling of Occupational Accidents' Size Using Structural Equation Modelling approach (SEM); A Field Study in Big Construction Industries. *BMJ Publishing Group Ltd* . 2016; 73: A172.

6. Soltanzadeh A, Mohammadfam I, Moghimbeigi A, Ghiasvand R. Key Factors Contributing to Accident Severity Rate in Construction Industry in Iran: A Regression Modelling Approach. *Arh Hig Rada Toksikol*, 2016; 67(1): 47-53.

7. Saffarzadeh M, Pooryari M. Accident Prediction Model Based on Traffic and Geometric Design Characteristics. *Int J Civ Eng*, 2005; 3(2): 112-9.

8. Sari M, Selcuk AS, Karpuz C, Duzgun HS. Stochastic Modeling of Accident Risks Associated with an Underground Coal Mine in Turkey. *Saf Sci*. 2009; 47(1): 78-87.

9. Venkataraman N. Safety Performance Factor. *Int J Occup Saf Ergon*, 2008; 14(3): 327-31.

10. Carrillo Castrillo JA, Trillo Cabello AF, Rubio Romero JC. Construction Accidents: Identification of the Main Associations between Causes, Mechanisms and Stages of the Construction Process. *Int J Occup Saf Ergon*. 2017; 23(2): 240-50.

11. Mohammadfam I, Soltanzadeh A, Moghimbeigi A, Savareh BA. Analysis and Modeling of Threatening Factors of Workforce's Health in Large-Scale Workplaces: Comparison of Four-Fitting Methods to Select Optimum Technique. *Electron Physician*. 2016; 8(2): 1918-26.

12. Mohammadfam I, Soltanzadeh A, Moghimbeigi A, Akbarzadeh M. Modeling of Individual and Organizational Factors Affecting Traumatic Occupational Injuries based on the Structural Equation Modeling: A Case Study in Large Construction Industries. *Arch Trauma Res*. 2016; 5(3): e33595.

13. Chong HY, Low TS. Accidents in Malaysian Construction Industry: Statistical Data and Court CasesInt. *J Occup Saf Ergon*, 2014; 20(3): 503-13.

14. Mitropoulos P, Abdelhamid TS, Howell GA. Systems Model of Construction Accident Causation. *J Constr Eng Manag*. 2005; 131(7): 816-25.

15. Moghaddam FR, Afandizadeh S, Ziyadi M. Prediction of Accident Severity Using Artificial Neural Networks. *Int J Civ Eng*, 2011; 9(1): 41-8.

16. OSHA. Safety and Health Managment System eTool. *OSHA*. 2012.

17. Hinrichs A, Novak E, Ullrich M, Woźniakowski H. The Curse of Dimensionality for Numerical Integration of Smooth Functions. *Math Comput*. 2014; 83(290): 2853-63.

18. Biesiada J, Duch W. Feature Selection for High-Dimensional Data—A Pearson Redundancy Based Filter, InComputer Recognition Systems 2. *Springerlink*. 2007: 242-9.

19. Awodele O, Jegede O. Neural Networks and Its Application in Engineering. *Sci IT*, 2009: 83-95.

20. Czarnecki WM, Osindero S, Jaderberg M, Swirszcz G, Pascanu R. Sobolev Training for Neural Networks. *arXiv Preprint arXiv*:1706.04859; 2017.

21. Berry MJ, Linoff GS. Data Mining Techniques: for Marketing, Sales, and Customer Relationship Management. 2nd ed. *USA: John Wiley & Sons*; 2011.

22. Soltanzadeh A, Mohammadfam I, Moghimbeygi A, Ghiasvand R. Exploring Causal Factors on the Severity Rate of Occupational Accidents in Construction Worksites. *Int J Civ Eng*, 2017; 15(7): 959-65.

23. Mohammadfam I, Soltanzadeh A, Moghimbeigi A, Akbarzadeh M. Confirmatory Factor Analysis of Occupational Injuries: Presenting an Analytical Tool. *Trauma Mon*, 2017; 22(2): e33266.

24. Mohammadfam I, Soltanzadeh A, Moghimbeigi A, Savareh BA. Use of Artificial Neural Networks (ANNs) for the Analysis and Modeling of Factors That Affect Occupational Injuries in Large Construction Industries. *Electron Physician*, 2015; 7(7): 1515-22.

25. Soltanzadeh A, Mohammadfam I, Moghimbeigi A. P153 Predicting and Determining Factors of Occupational Accidents Severity Rate (ASR) Using Artificial Neural Networks (ANN); A Case Study in Construction Industry. *BMJ Publishing Group Ltd*. 2016; 73(1):A171-2.

26. Piranveyseh P, Motamedzade M, Osatuke K, Mohammadfam I, Moghimbeigi A, Soltanzadeh A, et al. Association between Psychosocial, Organizational and Personal Factors and Prevalence of Musculoskeletal Disorders in Office Workers. *Int J Occup Saf Ergon*. 2016; 22(2): 267-73.

27. Samadi H, Rostami M, Bakhshi E, Garosi E, Kalantari R. Can Educational Intervention be Useful in Improvement of Body Posture and Work Related Musculoskeletal Symptoms? *J Hum Environ Health Promot*. 2018; 4(2): 81-6.

28. Alizadeh SS, Mortazavi SB, Mehdi Sepehri M. Assessment of Accident Severity in the Construction Industry Using the Bayesian Theorem. *Int J Occup Saf Ergon*. 2015; 21(4): 551-7.

29. Seifi Azad Mard HR, Estiri A, Hadadi P, Seifi Azad Mard M. Occupational Risk Assessment in the Construction Industry in Iran. *Int J Occup Saf Ergon*. 2017; 23(4): 570-7.

30. Baratchi M, Mansouri N, Ahmadi A. Hazard Assessment Matrix; Results of a Delphi Study. *J Hum Environ Health Promot*. 2018; 4(3): 121-5.