Use of Artificial Neural Networks (ANNs) for the Analysis and Modeling of Factors That Affect Occupational Injuries in Large Construction Industries

Iraj Mohammadfam[1], Ahmad Soltanzadeh[2], Abbas Moghimbeigi[3], Behrouz Alizadeh Savareh[4]

[1] Ph.D. of Occupational Hygiene Engineering, Associate Professor, Department of Occupational Hygiene Engineering, School of Public Health and Research Center for Health Sciences, Hamadan University of Medical Sciences, Hamadan, Iran
[2] Ph.D. Candidate of Occupational Hygiene Engineering, Department of Occupational Hygiene Engineering, School of Public Health and Research Center for Health Sciences, Hamadan University of Medical Sciences, Hamadan, Iran
[3] Ph.D. of Biostatistics, Associate Professor, Department of Biostatistics and Epidemiology and Modeling of Non-communicable Disease Research Canter, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran
[4] Ph.D. Candidate of Medical Informatics, Department of Medical Informatics, Allied Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Type of article: Original

Abstract

Introduction: Occupational injuries as a workforce's health problem are very important in large-scale workplaces. Analysis and modeling the health-threatening factors are good ways to promote the workforce's health and a fundamental step in developing health programs. The purpose of this study was ANN modeling of the severity of occupational injuries to determine the health-threatening factors and to introduce a model to predict the severity of occupational injuries.

Methods: This analytical chain study was conducted in 10 large construction industries during a 10-year period (2005-2014). Nine hundred sixty occupational injuries were analyzed and modeled based on feature weighting by the rough set theory and artificial neural networks (ANNs). Two analytical software programs, i.e., RSES and MATLAB 2014 were used in the study.

Results: The severity of occupational injuries was calculated as 557.47 ± 397.87 days. The findings of both models showed that the injuries' severity as a health problem resulted in various factors, including individual, organizational, health and safety (H&S) training, and risk management factors, which could be considered as causal and predictive factors of accident severity rate (ASR).

Conclusion: The results indicated that ANNs were a reliable tool that can be used to analyze and model the severity of occupational injuries as one of the important health problems in large-scale workplaces. Additionally, the combination of rough set and ANNs is a good and proper chain approach to modeling the factors that threaten the health of workforces and other H&S problems.

Keywords: workforce's health, occupational injury, accident severity rate (ASR), artificial neural networks (ANN), rough set theory

1. Introduction

Despite the extensive efforts to reduce occupational injuries, the World Health Organization (WHO) has introduced it as a public health epidemic (1-2). In addition, the International Labor Organization (ILO) stated that, annually, 268 million occupational accidents cause harm and injuries in workplaces and that the occupational injuries resulted in more than 2.2 million deaths among the workforce (1, 3). Developing and implementing health and safety (H&S)

programs have positive effects on people's, society's, and economic welfares, and they result in preventive H&S and control programs, thereby improving the health of the workforce. H&S planning requires proper and applicable data about the problems and the threatening factors that affect health. As is well known, occupational injuries are one of the most important health threats and problems in large-scale workplaces (4-6). The accident severity rate (ASR) is one of the significant indices in the analysis and modeling of the severity of occupational injuries (7, 8). Therefore, analysis and modeling of this index using practical modeling techniques can provide accurate identification of health-threatening factors, as well as help managers and engineers to design and develop various H&S programs for the prevention and reduction of occupational injuries (4, 9).

Some studies have shown that the analysis findings of occupational injuries were reliable and valid, making them useful as a basis for developing health management programs (4, 9, 10). Use of fitting methods is a good approach to analyze and modeling the severity of occupational injuries and health-threatening factors. Linear regression, generalized linear models, regression analysis, and artificial neural networks (ANNs) are examples of modeling techniques that can be used to model health problems, such as occupational injuries (11-13). The comparison of the fitting methods in order to select the optimum technique for the function approximation of ASR in part one of this study indicated that ANNs had the highest correlation coefficient (R = 0.968) and the lowest relative error (re = 0.063); thus, they were deemed to be the best fitting method for analyzing and modeling the severity of occupational injuries. Therefore, artificial neural networks (ANNs), as a best and optimum fitting technique, are one of the most significant, reliable, and practical tools used to model the factors that cause occupational injuries and to predict the severity of injuries (11-13). The results of the Carrillo-Castrillo research team confirmed that ANNs are useful in estimating occupational accident risk factors (12). In addition, Moghaddam et al. showed that ANNs can be used to predict the severity of occupational injuries (13). Therefore, in this study, we aimed to determine the causal factors of occupational injuries and further introduce a model for the prediction and reduction of the severity of occupational injuries and to improve the health of workforces based on a chain approach using rough set theory and artificial neural networks modeling

2. Material and Methods

2.1. Setting
This analytical chain study was carried out in 10 Iranian large construction industries from 2005 to 2014. In this study, 980 people who were injured in occupational accidents were analyzed and modeled. The following steps were employed to complete the study.

2.2. Data Collection
Data were collected from accident investigation forms, interviews (injured workforce, witnesses, supervisors, managers, and other personnel), documents of H&S training system and risk management. All registered documents of occupational injuries were peer reviewed over the 10-year period. Occupational injuries with incomplete data were excluded from the study. A total of 980 occupational accidents that resulted in injury were selected for the study. The studied and assessed factors included individual factors (i.e., average age, work experience, educational level, and marital status), organizational factors or characteristics of the workplace management and structure (i.e., job title, activity types during accidents, the number of workers, existing time pressure, and types of employment, including formal and contractors), H&S training factors (, i.e., pre-employment training, periodic training, past injury training, training about personal protective equipment (PPE) and housekeeping, quantity and quality of training), and H&S risk management factors (RMF: hazard identification (HAZID), periodic risk assessment, accident investigation, risk control measures, such as PPE, implementing tool box meeting (TBM), and housekeeping, applying H&S checklists, establishment of incident reports, and H&S audits). Furthermore, the severity of occupational injuries was analyzed based on the accident severity rate (ASR) (14). ASR-OSHA was calculated as follows (15): ASR = (total number of lost work days due to occupational accidents and injures × 200,000/total number of hours worked).

2.3. Feature Weighting
Due to the fact that each of the factors had a specific weight and effect on the severity of occupational injuries, therefore, determining the weight of each of these factors on ASR is important. Therefore, rough set theory was applied to determine the weight of each of these factors. Rough set theory is known as a new classification method in data mining and a mathematical tool for exploring data patterns (16). Moreover, it copes with separate factors through multi-level categories. What makes this method a good choice for feature selection is its ability to use an

explicit mathematical manner to ascertain latent patterns of data, redundancies, and dependencies between data features (17). In order to determine the weight of factors, rough set exploration system (RSES Software) was applied (18).

2.4. Artificial Neural Networks (ANNs)
ANNs are considered as predicting-analytical models that are used extensively in research fields. Some of them are industrial risk management control activities, accidents' consequences and severity, and data validation (19). Moreover, they can be helpful in predicting the severity of complicated occupational accidents and injuries and determining factors that affect such injuries (11-13). Multi-layer perceptron (MLP) is one of the well-known ANN models that is composed of an input, an output, and, usually, one or more latent layers. Input-output data are used to determine weights and biases while learning the MLP process (12, 13, 20). Additionally, the back-propagation technique, which generates the input forward in a network and calculates the error backward in an iterative manner, is used to obtain these parameters (13).

ANNs' architecture mostly influences the accuracy of the structure. According to MLP's hidden layer, different type of structures can be defined. Moreover, there are several sources for determining optimized structures of MLP (21-24). Besides, on the basis of MSE criterion, different models were compared. $MSE = 1/n \Sigma(y_i - \tilde{y}_i)^2$, where MSE: Mean Square Error, y_i: Network output of ith data, $\tilde{y}_i$: Desired output of ith data, and n: number of datasets. Also, "tan sig(n)= $2/(1+\exp(-2*n))-1$" was used as the activation function. This is mathematically equal to tanh (N). Notably, this function differs and runs faster than MATLAB implementation of tanh. Furthermore, the results can have small discrepancies. Tansig(n) is a good and proper trade-off between neural networks where speed is important and the exact shape of the transfer function is not. In addition, the Bayesian regularization learning function was used in constructing the network. This learning function uses the Jacobian matrix, which assumes that performance is a mean or sum of squared errors for calculations. As a result, trained networks with this function must use the MSE performance function. Finally, in order to avoid overfitting in the learning data, the neural network was divided as follows: 80% of the data were considered as the training set, and 20% were the test data; data classification in each set was assigned randomly (25). In addition, we used MATLAB 2014a software to implement the ANN.

3. Results
ASR of occupational injuries was calculated as 557.47 ± 397.87. The descriptive findings of IFs, OFs, TFs, and RMFs are shown in Table 1 and 2. The mean of age and work experience of injured workforces were 27.82 ± 5.23 and 4.39 ± 3.65 years, respectively. Higher than 70% of injured workforces were simple workers, and 24.2% were technicians; and 17.6% of injuries occurred in maintenance activities. In addition, the findings showed that time pressure was involved in 75% of occupational injuries; also 75% of the injuries occurred to the contractor workforces. Quality and quantity of training were 25.0% and 26.3%, respectively. Hazard identification (HAZID), periodic risk assessment, and accident investigation were estimated to be 18.6, 47.6, and 17.0%, respectively. Furthermore, H&S control measures, including PPE, TBM, and housekeeping, were 27.8, 16.2, and 23.6%, respectively. The findings of feature weighting by rough set theory are shown in Table 3. The results of applying rough set by RSES shows that among the 25 studied factors, 13 factors, i.e., age, number of workers, quality of training, HAZID, working experience, activity, housekeeping, accident investigation, past injury training, periodic risk assessment, contractor, periodic training, PPE, and time pressure, had weights above zero and were used as input factors in the implementation of the neural network. The results of ANN are presented in Figures 1-4. This ANN included 13 factors as input, one hidden layer (containing 20 neurons), and one output layer (ASR) (Figure 1). As mentioned earlier, the 13 output factors of feature weighting were considered as input factors of ANN. Figure 2 shows the MSE values versus learning epochs. The results showed that the lowest MSE was 0.0055, which indicates that the designed ANN was able to predict the exact ASR value according to the input factors. In addition, this finding confirmed the findings of feature weighting. Figure 3 shows the histogram of error between the actual value (ASR) and the value predicted by ANN. As can be seen, the absolute minimum value was 0.0005,3 and, as it moves away from the point, the value of the absolute error increases. As observed in Figure 4, the regression value (R) is a measure of the correlation between actual output (ASR) and the predicted value by the ANN (R = 0.9465). This significant result shows that the trained neural network was able to predict the value of ASR according to these13 factors.

Table 1. Descriptive results of Individual Factors (IF) and Organizational Factors (OF)

Studied Factors			Descriptive Values
Individual Factors (IF)	Age (years) (M ± SD)		27.82 ± 5.23
	Work Experience (M ± SD)		4.39 ± 3.65
	Education	Sub Diploma[1]	325 (33.2%)
		Diploma	398 (40.6%)
		Upper diploma	190 (19.4%)
		Above B.Sc.	67 (6.8%)
	Marital Status	Single	481 (49.1%)
		Married	499 (50.9%)
Organizational Factors (OF)	Average of workers (M ± SD)		41.45 ± 22.99
	Job Title	Simple Workers	719 (73.4%)
		Technicians	237 (24.2%)
		Supervisor	24 (2.4%)
	Activity Type	Normal Work	641 (65.4%)
		Installation	84 (8.6%)
		Maintenance	172 (17.6%)
		Material Handling	83 (8.5%)
	Time Pressure		721 (73.6%)
	Contractor		724 (73.9%)

1: Diploma: high school diploma

Table 2. Descriptive results of Training Factors (TF) and Risk Management Factors (RMF)

Studied Factors		Descriptive Values
H&S Training Factors (TF)	Pre-employment training	225 (23.0%)
	Periodic Training	381 (38.9%)
	Past Injury Training	135 (13.8%)
	PPE Training	253 (25.8%)
	Housekeeping Training	141 (14.4%)
	quantity of Training	258 (26.3%)
	Quality of Training	245 (25.0%)
H&S Risk Management Factors (RMF)	HAZID	182 (18.6%)
	Periodic risk assessment	466 (47.6%)
	accident investigation	167 (17.0%)
	Risk control measures; PPE	272 (27.8%)
	TBM	159 (16.2%)
	Housekeeping	231 (23.6%)
	Checklist	809 (82.6%)
	Incident Report	128 (13.1%)
	H&S Audit	197 (20.1%)

Table 3. Results of feature weighting by rough set

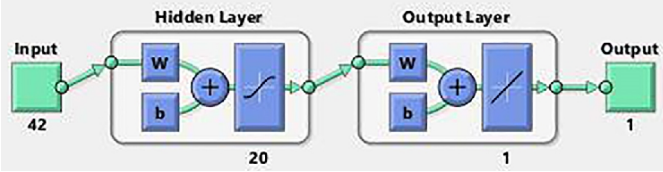| Selected factors	Factors Importance	Selected factors	Factors Importance
Age	100.0	Accident investigation	30.0
Number of workers	100.0	Past accident training	25.0
Quality of training	100.0	Periodic risk assessment	25.0
HAZID	80.0	Contractor	20.0
Working experience	75.0	Periodic training	15.0
Activity	75.0	PPE	5.0
Housekeeping	40.0	Time pressure	5.0

Figure 1. Algorithm of ANN to predict ASR
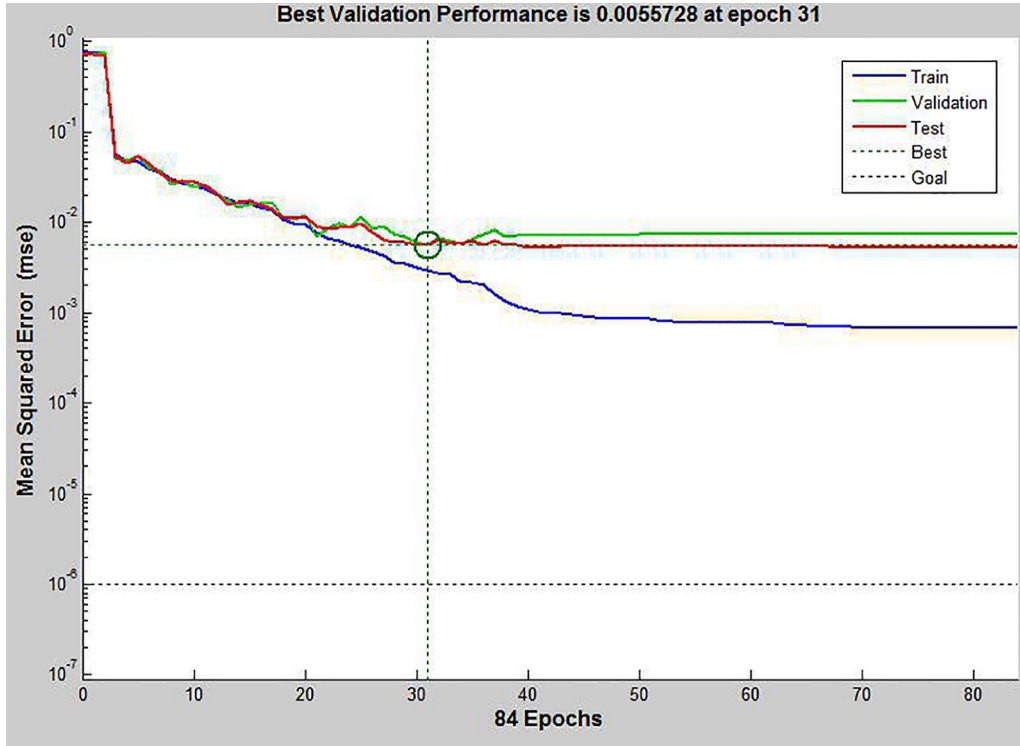

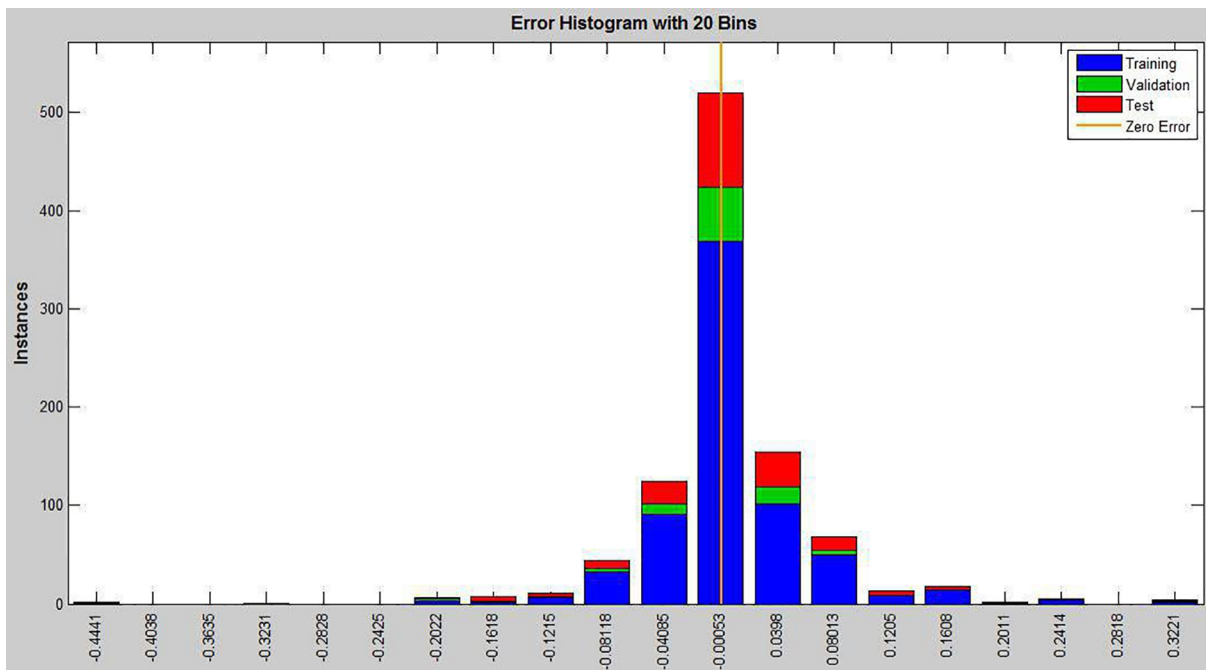**Figure 2.** MSE versus learning epochs
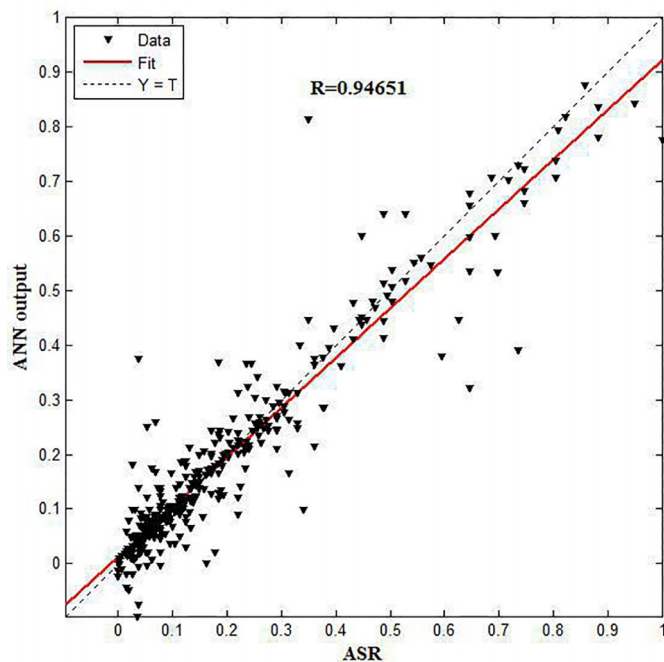

Figure 3. Error histogram

Figure 4. Relationship Model of ANN and ASR

4. Discussion

The present study, which was aimed to ANN modeling of the severity of occupational injuries to determine the H&S threatening factors, introduced a model to predict ASR for occupational injuries, and it was conducted based on using a chain approach including rough set and ANN. The findings of part two confirmed the results of part one, which implicated that ANN is an optimum technique for approximation function of occupational injuries' severity and other health problems in large-scale workplaces. Analysis and modeling of occupational injuries will result in accurate perception of health threatening factors and can be used as predictive algorithms and to model injuries in the future (11-13). Because of the great complexity of workplaces, analysis of occupational injuries using common analytical methods is very difficult. The findings of part 1 and some studies have shown that applying artificial neural networks (ANN) is proper and applicable to estimate injury factors (12, 13, 20). Carrillo-Castrillo et al. indicated that artificial neural network is a most reliable and practical tool that can analyze and model severity rate of occupational accidents and injuries (12). Also, Moghaddam et al. predicted highway crash severity using ANN (13).

Consistent with results of some studies, the findings of the present study indicated that occupational injuries are one of the most important threats to workforce's health, and they are induced by the combination of several factors. The results of both rough set and ANN indicated that the average age of workforce and work experience were individual factors, and the organization factors included average number of workers in each activity, activity type, time pressure, and using contractor workforce. Also, the quality of H&S training, periodic and past injury training as training factors, and H&S risk management factors, including HAZID, periodic risk assessment, accident investigation, risk control measures, such as using PPE and implementing housekeeping, were the most important factors in causal analysis and in predicting occupational injuries and their severity (4, 26-28).

ANN modeling of ASR revealed that this technique is an effective and reliable tool for analyzing and modeling health problems, such as occupational injuries. This modeling affirmed that the created and validated ANN model of ASR is a valid solution approach in the analysis of the severity of occupational injuries. Thus, the findings can be applied for predicting and preventing health-related problems. Additionally, in re-interpreting the results of this modeling, it can be shown that with the 13-factors, ASR in the large-scale workplaces can be 94.651% predicted. In other words, ASR modeling by ANN should be defined and developed to promote optimum performance of H&S structures in the workplaces.

The findings of this analytical study, which were presented in two parts, indicated that applying the chain approach in analyzing and modeling occupational accidents and injuries as one of the most important threats to the health of the workforces is a good and applicable method. Furthermore, this would lead to identification and determining combinations of factors that affect the health and safety of the workforce in large-scale workplaces, and it is a proper way to predict and reduce occupational injuries.

5. Conclusions

The findings affirmed that the chain analytical approach, including rough set theory and ANN modeling, is a good and applicable method to analyze and model the factors affecting health of workforce and predicting severity of occupational injuries. Furthermore, the results indicated that health threats, such as occupational injuries, are a result of combination factors, including individual, organization, H&S training, and risk management factors. Therefore, health decision makers should plan comprehensive health programs that focus on all of the mentioned factors.

Conflict of Interest:
There is no conflict of interest to be declared.

Authors' contributions:
All authors contributed to this project and article equally. All authors read and approved the final manuscript.

References

1) Fernández-Muñiz B, Montes-Peón JM, Vázquez-Ordás CJ. Relation between occupational safety management and firm performance. Safety science. 2009; 47(7): 980-91. doi: 10.1016/j.ssci.2008.10.022.
2) Hämäläinen P, Takala J, Saarela KL. Global estimates of occupational accidents. Safety science. 2006; 44(2): 137-56. doi: 10.1016/j.ssci.2005.08.017.
3) Fernández-Muñiz B, Montes-Peón JM, Vázquez-Ordás CJ. Safety management system: Development and validation of a multidimensional scale. Journal of Loss Prevention in the process Industries. 2007; 20(1): 52-68. doi: 10.1016/j.jlp.2006.10.002.
4) Amini SM, Gilaki M, Karchani M. Safety of Nanotechnology in Food Industries. Electron Physician. 2014; 6(4): 962. doi: 10.14661/2014.962-968, PMID: 25763176, PMCID: PMC4324265
5) Teymuri GH, Sadeghian M, Kangavari M, Asghari M, Madrese E, Abbasinia M, et al. Review the number of accidents in Tehran over a two-year period and prediction of the number of events based on a time-series model. Electron Physician. 2013; 5(3): 698-705. doi: 10.14661/2013.698-705, PMID: 26120405. PMCID: PMC4477752
6) Mearns K, Rundmo T, Flin R, Gordon R, Fleming M. Evaluation of psychosocial and organizational factors in offshore safety: a comparative study. J Risk Res. 2004; 7(5): 545-61. doi: 10.1080/1366987042000146193.
7) Coleman PJ, Kerkering JC. Measuring mining safety with injury statistics: Lost workdays as indicators of risk. J Safety Res. 2007; 38(5): 523-33. doi: 10.1016/j.jsr.2007.06.005, PMID: 18023637.
8) Rikhardsson PM, Impgaard M. Corporate cost of occupational accidents: an activity-based analysis. Accid Anal Prev. 2004;36(2):173-82. doi: 10.1016/S0001-4575(02)00147-1, PMID: 14642872.
9) Cowlagi RV, Saleh JH. Coordinability and consistency in accident causation and prevention: formal system theoretic concepts for safety in multilevel systems. Risk Anal. 2013; 33(3): 420-33. doi: 10.1111/j.1539-6924.2012.01886.x, PMID: 22967134.
10) Aghakhani N, Rahbar N, Feizi A. Epidemiology of burn injuries in west Azerbaijan province, western Iran. Iran Red Crescent Med J. 2009; 11(1): 85-9.
11) Khelassi A. Explanation-aware computing of the prognosis for breast cancer supported by IK-DCBRC: Technical innovation. Electron Physician. 2014; 6(4): 947–954. doi: 10.14661/2014.947-954, PMID: 25763174, PMCID: PMC4324263.

12) Carrillo-Castrillo JA, Martín JG, de la Vega RG, Onieva L. Neural Network Application for Risk Factors Estimation in Manufacturing Accidents. Managing Complexity: Springer; 2014. p. 277-83.

13) Moghaddam FR, Afandizadeh S, Ziyadi M. Prediction of accident severity using artificial neural networks. International Journal of Civil Engineering. 2011; 9(1): 41.

14) Sari M, Selcuk AS, Karpuz C, Duzgun HSB. Stochastic modeling of accident risks associated with an underground coal mine in Turkey. Safety sci. 2009; 47(1): 78-87. doi: 10.1016/j.ssci.2007.12.004.

15) (OSHA) OSHA. Safety & health managment system etool. 2012.

16) Greco S, Matarazzo B, Slowinski R. Rough sets theory for multicriteria decision analysis. Eur J Oper Res. 2001; 129(1): 1-47. doi: 10.1016/S0377-2217(00)00167-3.

17) Pawlak Z. Rough sets: Theoretical aspects of reasoning about data: Springer Science & Business Media; 2012.

18) Zhao S. Intrusion Detection Using the Support Vector Machine Enhanced with a Feature-weight Kernel: University of Regina; 2007.

19) Awodele O, Jegede O, editors. Neural networks and its application in engineering. Proceedings of Informing Science & IT Education Conference (InSITE); 2009.

20) Chang L-Y. Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. Safety sci. 2005; 43(8): 541-57. doi: 10.1016/j.ssci.2005.04.004.

21) Matviykiv O, Faitas O. Data classification of spectrum analysis using neural network. 2012.

22) Linoff GS, Berry MJ. Data mining techniques: for marketing, sales, and customer relationship management: John Wiley & Sons; 2011.

23) Swingler K. Applying neural networks: a practical guide: Morgan Kaufmann; 1996.

24) Xu S, Chen L. A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. 2008.

25) Prechelt L. Proben1: A set of neural network benchmark problems and benchmarking rules. 1994.

26) Kalte HO, Hosseini AH, Arabzadeh S, Najafi H, Dehghan N, Akbarzadeh A, et al. Analysis of electrical accidents and the related causes involving citizens who are served by the Western of Tehran. Electron Physician. 2014; 6(2): 820-6. doi: 10.14661/2014.820-826, PMID: 25763153, PMCID: PMC4324276.

27) Rahmani A, Mahmodkhani S, Moslemi M, Gharagozlou F, Ahmadnezhad I, Karchani M, et al. Investigation of Occupational Stress and its Relationship with the Demographic Characteristics of Workers in Ilam, Iran. Electron Physician. 2013; 5(1): 611-5. doi: 10.14661/2013.611-615, PMID: 26120391, PMCID: PMC4477777.

28) Kalte HO, Ziaei M, Najafi H, Karami A, Akbarzadeh A, Ahmadnezhad I, et al. Use of the Finnish method to quantify the ergonomic properties in an office environment among the workplaces in an Electricity Distribution Company in West Tehran Province. Electron Physician. 2014; 6(1): 779-85. doi: 10.14661/2014.779-785, PMID: 25763145, PMCID: PMC4324280.