PAN
ARCHIVES OF ACOUSTICS
Vol. **45**, No. 2, pp. 303–311 (2020)
DOI: 10.24425/aoa.2020.133150

Research Paper

Modeling and Predicting the Changes in Hearing Loss of Workers with the Use of a Neural Network Data Mining Algorithm: A Field Study

Sajad ZARE[(1)], Mohammad Reza GHOTBIRAVANDI[(1)], Hossein ELAHISHIRVAN[(2)*]
Mostafa Ghazizadeh AHSAEED[(3)], Mina ROSTAMI[(2)], Reza ESMAEILI[(4)]

[(1)] *Department of Occupational Health*
School of Public Health
Kerman University of Medical Sciences and Health Services
Kerman, Iran; e-mail: ss_zare87@yahoo.com, ravandir@yahoo.co.uk

[(2)] *Student Research Committee, Kerman University of Medical Sciences*
Kerman, Iran; e-mail: mina_rostami1370@yahoo.com
*Corresponding Author e-mail: hosseynelahi12@gmail.com

[(3)] *Department of Computer Engineering*
Shahid Bahonar University of Kerman
Kerman, Iran; e-mail: GhazizadehAhsaee.most@stu-mail.um.ac.ir

[(4)] *School of Public Health*
Hamedan University of Medical Sciences
Hamedan, Iran; e-mail: hosseinelahi555@yahoo.com

(*received September 23, 2018; accepted February 24, 2020*)

The aim of the study study was to model, with the use of a neural network algorithm, the significance of a variety of factors influencing the development of hearing loss among industry workers. The workers were categorized into three groups, according to the A-weighted equivalent sound pressure level of noise exposure: Group 1 (L_{Aeq} < 70 dB), Group 2 (L_{Aeq} 70–80 dB), and Group 3 (L_{Aeq} > 85 dB). The results obtained for Group 1 indicate that the hearing thresholds at the frequencies of 8 kHz and 1 kHz had the maximum effect on the development of hearing loss. In Group 2, the factors with maximum weight were the hearing threshold at 4 kHz and the worker's age. In Group 3, maximum weight was found for the factors of hearing threshold at a frequency of 4 kHz and duration of work experience. The article also reports the results of hearing loss modeling on combined data from the three groups. The study shows that neural data mining classification algorithms can be an effective tool for the identification of hearing hazards and greatly help in designing and conducting hearing conservation programs in the industry.

Keywords: noise; modeling; noise induced hearing loss (NIHL); neural network algorithm.

1. Introduction

Exposure to noise is a largely prevalent occupational health risk factor. Occupational noise exposure is usually characterized by frequency-weighted sound level, normalized to an 8-hour working day (LEX,8h). It is estimated that over 60 million of people around the world are exposed to noise exceeding the 85 dB permissible limit in the workplace, for an eight-hour working day (Zare *et al.*, 2019). Exposure to excessive noise levels affect the workers' health and causes various occupational safety hazards (Nassiri *et al.*, 2016; Safari Variani *et al.*, 2018). One of the most adverse health effects of noise exposure is noise-induced hearing loss (NIHL).

Worldwide, NIHL is a largely prevalent occupational hearing hazard. The World Health Organization (WHO) has estimated that over 12% of the world population are affected by NIHL and NIHL is the second most common cause of hearing loss among older adults (Zare *et al.*, 2019). In the USA, 7.4–10.2 million industrial workers are at risk of NIHL. In Sweden, 100 million US dollars is annually spent on compensation for the hearing-impaired (Ahmed *et al.*, 2001). In Malasya, 26.9% of industry workers have a hearing loss in a frequency range of 3000–6000 Hz (Leong, 2003) and in Iran, at least 1 million workers are subjected to excessive noise exposure (Golmohammadi *et al.*, 2006).

In the industry, noise pollution control is mainly based on the measurement of $L_{EX,8h}$ (Golmohammadi *et al.*, 2006). The estimation of occupational health hazards caused by noise exposure includes a variety of factors, such as the worker's age and the duration of work experience, the type of noise, the noise exposure duration, the noise frequency spectrum, the number of workers affected by noise, and the duration of the working shift (Zamanian *et al.*, 2012; 2013; Tajic *et al.*, 2008). The knowledge of the interrelationships between individual hearing risk factors greatly helps in designing and undertaking hearing protection activities in the workplace. The present study had the following main objectives: (1) to identify the predictor factors of hearing loss in the industry, (2) to determine the hearing loss of workers across both ears, (3) to model and predict the changes of hearing loss with the use of neural network algorithms, (4) to assess the accuracy and error rates of the hearing loss models developed in the study.

The study was conducted with the use of data mining and artificial neural network modelling. Data mining (DM) is the process of extracting valid, reliable information from databases and transforming the information of interest into a form suitable for use in a given application or activity. Data mining is the analysis step of the Knowledge Discovery in Databases (KDD) process (Badr *et al.*, 2009). Artificial neural networks are computational, information processing paradigms modeled after the human brain, designed to recognize patterns and the relationships between the components and parameters of the system under investigation (Kohzadi *et al.*, 1995). The advantage of artificial neural networks is direct learning from the data without any need to estimate their statistical characteristics. Neural networks make it possible to uncover the relationship between the set of inputs and outputs, to predict every output corresponding to a desired input with no need of any initial assumptions and knowledge about the relationships between the studied parameters (Golabi *et al.*, 2013).

2. Method

2.1. Objective of the study

The study was a cross-sectional investigation aimed at monitoring and predicting – with the use of an artificial neural network modeling – the development of hearing loss of workers in a mineral and in an industrial company. After determining the factors influencing the hearing loss we sought to determine the impact and the weight of each individual factor. The successive stages of the study were as follows:

1) selection of predictor factors for hearing loss modeling,
2) audiometric testing of both ears,
3) calculation of permanent threshold shift (PTS) in each ear,
4) calculation of total PTS across both ears,
5) classification of hearing loss degrees,
6) modeling the hearing loss changes with the use of an artificial neural network algorithm,
7) estimation of the error rate and the accuracy of the model (ISO 1999, 2013; Ramos-Miguel *et al.*, 2015).

2.2. Predictor factors

The workers tested in the study were assigned into three groups: one control group and two case groups, classified according to the noise exposure level. The factors of age, duration of work experience, A-weighted equivalent sound pressure level related to an 8-hour working day ($L_{EX,8h}$) and frequency were used to model the development of hearing loss (Nawi *et al.*, 2011; Ramos-Miguel *et al.*, 2015; ISO 1999, 2013). The subjects were divided into three age groups: 20–35, 35–50, and +50 year-old (Ramos-Miguel *et al.*, 2015). The subjects in each group were categorized into three ranges of the duration of work experience: less than 10 years, 10–20 years, and more than 20 years of service (Majumder, Sharma, 2014).

Equivalent sound pressure level was measured using a TES-1345 dosimeter. The dosimeter was calibrated with a CEL 110.2 calibrator (ISO 9612, 2009; Golmohammadi, Aliabadi, 1999). The modeling and prediction of hearing loss changing were made for frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz (ISO 1999, 2013; Schlauch, Nelson, 2009).

2.3. Audiometric measurements and calculation of NIHL

Pure-tone hearing thresholds were measured at 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz frequencies with a CA 120 audiometer (Gubbels *et al.*, 2017). Noise induced hearing loss (NIHL) was calculated for each ear from Eq. (1) (Golmohammadi, Aliabadi, 1999):

$$\mathrm{NIHL} = \frac{a^*}{4}, \tag{1}$$

where

$$a^* = (\mathrm{TL}_{500\ \mathrm{Hz}}) + (\mathrm{TL}_{1000\ \mathrm{Hz}}) + (\mathrm{TL}_{2000\ \mathrm{Hz}}) + (\mathrm{TL}_{4000\ \mathrm{Hz}}),$$

and TL – hearing loss at a given frequency [dB], NIHL – noise-induced hearing loss [dB].

After calculating NIHL for each ear, the total NIHL was calculated for both ears, from Eq. (2) (Golmohammadi, Aliabadi, 1999):

$$\mathrm{NIHL}_t = \frac{(\mathrm{NIHL}_b \times 5) + (\mathrm{NIHL}_p)}{6}, \tag{2}$$

where NIHL_t – total PTS in both ears [dB], NIHL_b – PTS in the better ear [dB], NIHL_p – PTS in the weaker ear [dB].

2.4. Classification of hearing loss degrees

The degree of hearing loss was classified according to WHO (1991). Hearing losses of 26–40 dB were considered mild, 41–55 dB moderate, 56–70 dB moderately severe, 71–90 dB severe, and > 91 dB profound (WHO, 1991).

2.5. Modeling the development of hearing loss on the basis of an artificial neural network

2.5.1. Input and output encoding

The robustness of neural networks to unforeseen pattern variations in new data set is regarded as their positive side. Their downside however is that, for encoding attribute values, they follow a standardized procedure, meaning that all the attributes, including the categorical ones, are granted a value ranging from 0 to 1. The following equation (Eq. (3)) indicates the calculation algorithm (Larose, Larose, 2014):

$$X^* = \frac{X - \min(X)}{\mathrm{range}(X)} = \frac{X - \min(X)}{\max(X) - \min(X)}. \tag{3}$$

Upon clear ordering of the classes, single output nodes can be utilized. For instance, one can imagine a categorization of reading prowess in elementary schools based on a collection of student attributes. The successive reading level categories may be defined as follows: first grade category – output from 0 to 0.25, second grade category: $0.25 \leq$ output < 0.50, third grade category: $0.50 \leq$ output < 0.75, fourth grade category: output ≥ 0.75 (Larose, Larose, 2014).

2.5.2. Neural networks for estimation and prediction

Since neural networks produce a continuous output, they are typically exploited for running estimation and prediction, using Eq. (4):

$$\textit{prediction} = \textit{output}\ (\text{data range}) + \textit{minimum}, \tag{4}$$

where *output* is the neural network output in (0,1) range, *data range* is the range of the original attribute values on a non-normalized scale, and *minimum* indicates the lowest attribute value on the non-normalized scale (Larose, Larose, 2014).

A neural network comprises a *layered*, *feedforward*, *completely connected* network of artificial neurons, or *nodes*. The *feedforward* nature of the network is limited to a single flow direction, whereby no looping or cycling is permitted. The neural network consists of two or more layers; nonetheless, the majority of networks comprise three layers: an *input layer*, a *hidden layer*, and an *output layer*. The number of hidden layers may not exceed one. In most of the cases, however, the networks encompass only one layer, which is sufficient for most applications. Furthermore, the neural network is *completely connected*, which means that each node in a particular layer is associated with each node in the next layer. On the other hand, the nodes of the same layer are not connected to each other. The weight of the connection between nodes is indicated by W_{1A}. At the stage of initialization, the weights are given random values of 0 and 1 (Larose, Larose, 2014).

The number and the type of data set attributes commonly determine the number of input nodes. The number of hidden layers and the number of nodes in every hidden layer can be identified by the user. Based on the classification task, the output layer may possess more than one node (McCullagh, 2010; Larose, Larose, 2014). The neural network structure is displayed in Fig. 1.

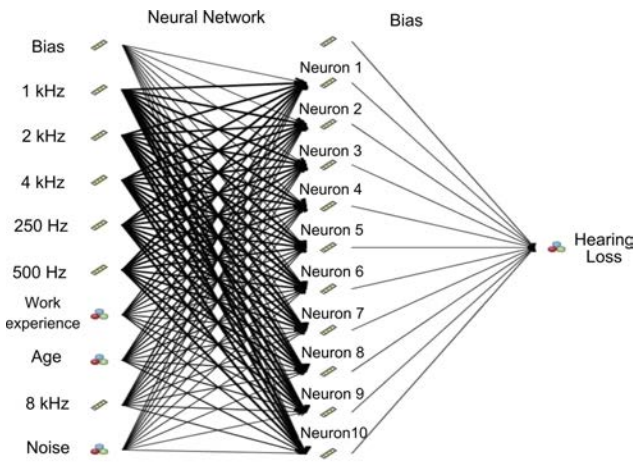

Fig. 1. Neural network structure.

The power and the flexibility of the network are related with the number of nodes in the hidden layer. A large number of hidden layers may cause overfitting which results in memorizing the training set at the expense of generalizability to the validation set. If overfitting occurs, one may reduce the number of hidden layers. On the other hand, when the training accuracy is too low, the number of nodes may be increased in the hidden layer (Larose, Larose, 2014).

Data set inputs (e.g. attribute values) are fed into the input layer and pass through the hidden layer with no processing. As a result, the input layer nodes do not possess the same node structure as that of the hidden layer and output layer nodes. The node inputs and the connection weights are combined into a single scalar value through a combination function (which typically is summation Σ), known as net (Eq. (5)) (Larose, Larose, 2014):

$$\text{net}_j = \sum_{ij}^{i} x_{ij} = W_{0_j} x_{0_j} + W_{1_j} x_{1_j} + ... + W_{Ij} x_I, \tag{5}$$

where x_{ij} indicates the i-th input to node j, W_{ij} refers to the weight connected with the ith input to node j, and there are $I+1$ inputs to node j.

It is worth mentioning that inputs from upstream nodes are illustrated by $x_1, x_2, ..., x_I$, whereas x_0 is a constant input which conventionally takes the value of 1 and is similar to the constant factor in regression models. Therefore, every hidden or output layer node (represented by j) comprises an extra input which is equal to a specific weight $W_{0_j} x_{0_j} = dW_{0_j}$ (e.g. W_{0_B} for node B). Equation (6) presents the calculation algorithm (Larose, Larose, 2014):

$$\begin{aligned} \text{net}_A &= \sum_i W_{iA} x_{iA} \\ &= W_{0A}(1) + W_{1A} x_{1A} + W_{2A} x_{2A} + W_{3A} x_{3A}. \end{aligned} \tag{6}$$

In neurons functioning in biology, signals are transmitted between two neurons only if the combination of inputs to a neuron exceeds a certain threshold level, which results in firing of the neuron. The firing response is not necessarily linearly related to the input stimulation increment. This behavior of neurons in biology is simulated in artificial neural networks by a nonlinear activation function. The sigmoid function (Eq. (7)) is the most typical activation function (Krogh, Vedelsby, 1995; Larose, Larose, 2014):

$$y = \frac{1}{1 + e^{-x}}, \tag{7}$$

where e is the natural logarithm base.

Prior to computing net_Z, the node contribution should be gauged, as shown in Eq. (8):

$$\begin{aligned} \text{net}_B &= \sum_i W_{iB} x_{iB} \\ &= W_{0B}(1) + W_{1B} x_{1B} + W_{2B} x_{2B} + W_{3B} x_{3B}. \end{aligned} \tag{8}$$

In the next step, node Zcombines the outputs from nodes A and B through net_Z, a weighted sum. This is carried out by the use of the weights related to the connections between these nodes, as shown in Eq. (9) (Larose, Larose, 2014).

$$\begin{aligned} \text{net}_Z &= \sum_i W_{iZ} x_{iZ} \\ &= W_{0Z}(1) + W_{AZ} x_{AZ} + W_{BZ} x_{BZ}. \end{aligned} \tag{9}$$

It should be noted that the inputs x_i to node Z are not data attribute values but outputs from the sigmoid functions from upstream nodes.

2.5.3. Assessment of the accuracy and the error rates of the models

The accuracy and the error rates of the models were determined from the confusion matrix. Confusion matrix is a square matrix whose dimensions are equal to the number of the output factor groups. In the matrix, the main diameter represents the percentage of cases predicted properly. According to Eq. (10), the model accuracy is the ratio of positive cases to the total number of cases (Larose, Larose, 2014)

$$\text{Accuracy} = \frac{\text{True Postive cases} + \text{True Negative cases}}{\text{All cases}}. \tag{10}$$

2.6. Ethical considerations

Ethical approval was obtained from the Ethics Committee of Kerman University of Medical Sciences (ID: IR.KMU.REC.1396.2458). All participants signed an informed consent form.

2.7. Data analysis

The data were analyzed with the use of SPSS software, version 18. The mean, standard deviation, correlation coefficient, and regression diagrams were analyzed by linear regression and a paired t-test. Modeling of the hearing loss changes was made with the use of IBM SPSS Modeler 18.0 software.

3. Results

3.1. Predictive factors

3.1.1. Age and work experience duration

The age and work experience duration of the workers tested in the study are shown in Table 1.

Table 1. Age and work experience duration of workers in Groups 1–3.

Groups	Factors	Mean	Standard deviation
Group 1 (n = 50)	Age	37.66	9.91
	Work experience duration	9.1	4.9
Group 2 (n = 50)	Age	35.56	11.45
	Work experience duration	8.48	5.38
Group 3 (n = 50)	Age	41.76	10.93
	Work experience duration	11.34	5.32

3.1.2. Equivalent sound pressure level

Group 1 was exposed to A-weighted equivalent sound pressure level (L_{Aeq}) of less than 70 dB. The participants in Group 2 and Group 3 were exposed to L_{Aeq} of 70–80 dB and >85 dB, respectively. The mean L_{eq} and standard deviation across participants in the groups were 70.0 ± 3 dB (Group 1), 77.6 ± 4.4 dB (Group 2), and 89.7 ± 3.0 dB (Group 3).

3.2. Hearing loss (target factor)

The hearing loss degree across both ears of the participants is shown, in terms of hearing loss severity, in Table 2. In Group 1 ($L_{\mathrm{Aeq}} < 70$ dB) 80% of the participants had normal hearing and 20% had mild hearing loss. In Group 2 (L_{Aeq} 70–80 dB), 74% had normal hearing and 20% mild hearing loss. In Group 3 ($L_{\mathrm{Aeq}} > 85$ dB), 58%, 20%, 8%, and 4% of the participants had, respectively, normal hearing, mild, moderate, and severe hearing loss, according to WHO (1991) classification. The data in Table 2 indicate that the severity of hearing loss increases with L_{Aeq}. The results of a paired t-test have shown that there was no significant difference ($p > 0.05$) between the mean hearing loss in the left and in the right ear of the participants, in the three groups.

3.3. Correlation between age, duration of work experience, and hearing loss

Table 3 shows the correlation between age, duration of work experience and hearing loss, determined by linear regression for the three groups. The data indicate that there was a statistically significant difference between age and hearing loss as well as between duration of work experience and hearing loss, in Groups 1 and 2.

3.4. Relationship between L_{Aeq} and hearing loss

Figure 2 shows correlation and linear regression between L_{Aeq} and hearing loss for all participants in Groups 1–3.

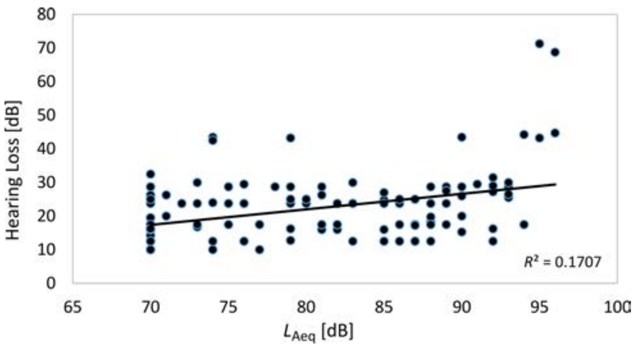

Fig. 2. Correlation between L_{Aeq} and hearing loss across Groups 1–3.

Table 2. Hearing loss of workers in Groups 1–3.

	Normal (0–25 dB)	Mild (26–40 dB)	Moderate (41–60 dB)	Severe (61–80 dB)	Profound (>80 dB)
Group 1 ($n = 50$) ($L_{\mathrm{Aeq}} < 70$ dB)	40 participants (80%)	10 participants (20%)	–	–	–
Group 2 ($n = 50$) (L_{Aeq} 70–80 dB)	37 participants (74%)	10 participants (20%)	3 participants (6%)	–	–
Group 3 ($n = 50$) ($L_{\mathrm{Aeq}} > 85$ dB)	29 participants (58%)	15 participants (30%)	4 participants (8%)	2 participants (4%)	–

Table 3. Correlation of age and work experience duration with hearing loss of workers in Groups 1–3.

	Group	Regression statistics		
		Coefficient of determination, R^2*	Correlation, R	p
Age	Group 1	0.148	0.385	0.008
	Group 2	0.155	0.394	0.008
	Group 3	0.036	0.189	0.277
Work experience duration	Group 1	0.131	0.362	0.014
	Group 2	0.102	0.32	0.038
	Group 3	0.079	0.28	0.076

3.5. Modeling of hearing loss changes

Five different models (Model 1 – Model 5) of hearing loss changes were used for hearing loss prediction and modeling. Model 1 (People working in the office) was used for Group 1, Model 2 – for Group 2, Model 3 – for Group 3, Model 4 (people working in the operating area) for Groups 2 and 3, Model 5 for Groups 1, 2, and 3. The results of modeling obtained with the use of IBM SPSS Modeler 18.0, are shown in Figs 3–7. The plots show, for a given model, the weight percentage of hearing loss of each predictor factor. Tables 4–8 present the confusion matrix of the neural network algorithm for the hearing loss models.

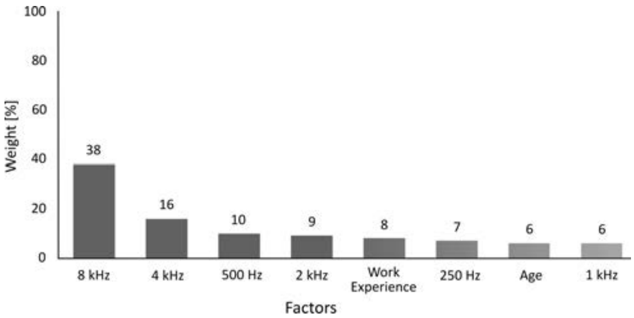

Fig. 3. Model 1: Weight percentage of hearing loss predictor factors for Group 1 (n = 50).

Table 4. Confusion matrix of data determined by the neural network algorithm in Model 1.

Measured severity	Predicted NIHL severity	
	Normal	Mild
Normal	100%	0.0%
Mild	0.0%	100%

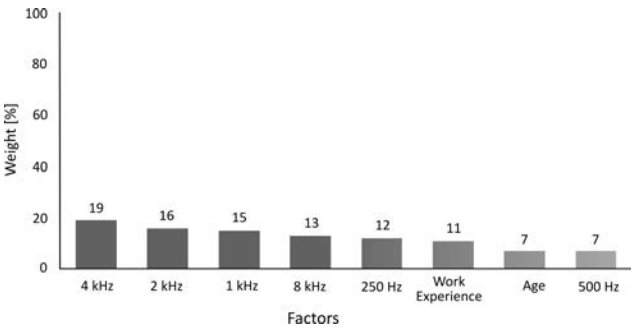

Fig. 4. Model 2: Weight percentage of hearing loss predictor factors for Group 2 (n = 50).

Table 5. Confusion matrix of data determined by the neural network algorithm in Model 2.

Measured severity	Predicted NIHL severity		
	Normal	Mild	Moderate
Normal	100%	0.0%	0.0%
Mild	0.0%	100%	0.0%
Moderate	0.0%	0.0%	100%

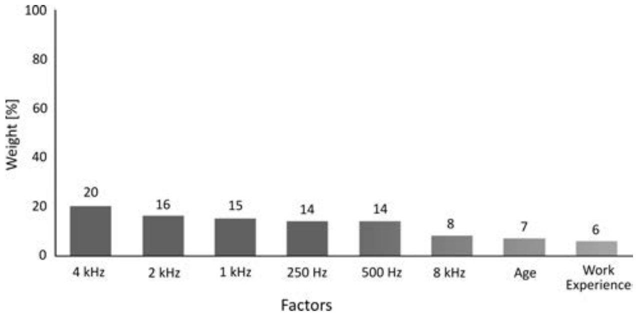

Fig. 5. Model 3: Weight percentage of hearing loss predictor factors for Group 3 (n = 50).

Table 6. Confusion matrix of data determined by the neural network algorithm in Model 3.

Measured severity	Predicted NIHL severity			
	Normal	Mild	Moderate	Severe
Normal	100%	0.0%	0.0%	0.0%
Mild	40%	53.3%	0.0%	6.7%
Moderate	0.0%	75%	25%	0.0%
Severe	0.0%	0.0%	0.0%	100%

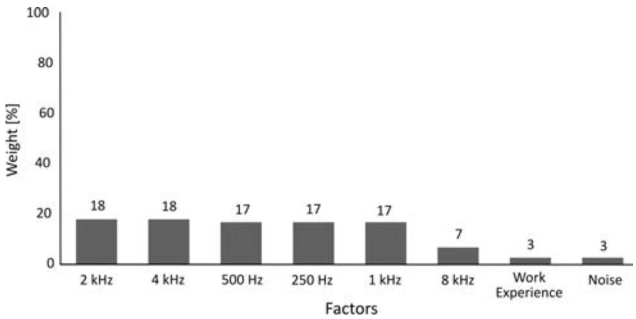

Fig. 6. Model 4: Weight percentage of hearing loss predictor factors for Groups 2 and 3 (n = 100).

Table 7. Confusion matrix of data determined by the neural network algorithm in Model 4.

Measured severity	Predicted NIHL severity			
	Normal	Mild	Moderate	Severe
Normal	98.5%	1.5%	0.0%	0.0%
Mild	0.0%	96%	4%	0.0%
Moderate	0.0%	0.0%	100%	0.0%
Severe	0.0%	0.0%	0.0%	100%

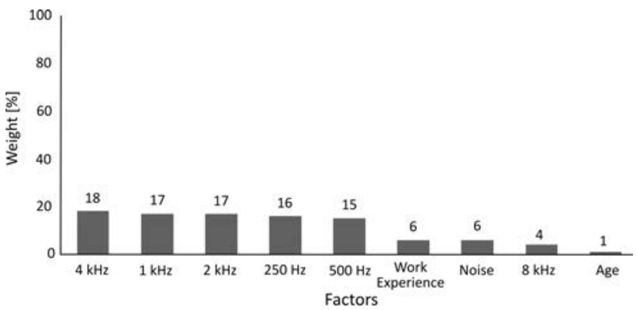

Fig. 7. Model 5: Weight percentage of hearing loss predictor factors for Groups 1–3 (n = 150).

Table 8. Confusion matrix of data determined by the neural network algorithm in Model 5.

Measured severity	Predicted Severity of NIHL			
	Normal	Mild	Moderate	Severe
Normal	100%	1.5%	0.0%	0.0%
Mild	2.9%	97.1%	0.0%	0.0%
Moderate	0.0%	0.0%	100%	0.0%
Severe	0.0%	0.0%	0.0%	100%

4. Discussion and conclusions

The results shown in Table 3 indicate that correlation between age and hearing loss as well as between duration of work experience and hearing loss of the workers were statistically significant in Groups 1 and 2. The data in Fig. 2 show that there was a significant difference between the A-weighted sound pressure level of exposure and hearing loss across Groups 1–3.

The data presented in Figs 3–7 show the weight of individual predictor factors in the development of hearing loss in Models 1–5. In Model 1 ($L_{\text{Aeq}} < 70$ dB), the hearing threshold at 8 kHz, with a 38% weight, had the maximum impact on hearing loss, while the factors of age and hearing threshold at 1 kHz (6% weight), were the least influential ones (Fig. 3). The prediction accuracy of the neural network algorithm was 100% (Table 4). In Model 2 (L_{Aeq} 70–80 dB), the factor with maximum weight was the hearing threshold at 4 kHz (19%) while the factors of age and threshold at the 500-Hz frequency (weights of 7%) had the minimum impact (Fig. 4). The prediction accuracy of the neural network algorithm was 100% (Table 5). In Model 3 ($L_{\text{Aeq}} > 85$ dB) the threshold at 4 kHz, with a 20% weight, had the maximum impact while the factor of duration of work experience had the minimum impact, with a 6% weight (Fig. 5). In this model, the prediction accuracy of the neural network algorithm was 80% (Table 6). In Model 4, determined for Groups 2 (L_{Aeq} 70–80 dB) and 3 ($L_{\text{Aeq}} > 85$ dB), the maximum impact was observed for the hearing thresholds at 4 kHz and 2 kHz frequencies, with 18% weights, and the minimum impact (3% weight) was found for the factors of duration of work experience and noise (Fig. 6). The prediction accuracy of the neural network algorithm was 98% (Table 7). In Model 5, determined for Groups 1–3, the threshold at 4 kHz had the maximum impact, with a weight of 18%, and the least impact (1% weight) was found for the factor of age (Fig. 7). The neural network algorithm prediction accuracy was 99.3% (Table 8).

The present findings, demonstrating the effects of individual predictor factors on the development of hearing loss in industry workers, are in agreement with published studies of hearing loss in steel industry workers which indicate that hearing loss increases with sound exposure level and duration of workers' service (Golmohammadi *et al.*, 2001; Masumi *et al.*, 2008). Golmohammadi *et al.* (2006) studied the effect of noise on the development of hearing loss of workers in stone cutting industry in Iran and reported that maximum hearing loss was observed in a frequency range around 4000 Hz (Golmohammadi *et al.*, 2006). Zare *et al.* (2019) used the C5 algorithm to determine the weight of factors affecting hearing loss, determined from audiometric data, in three groups of workers, classified on the basis of the exposure sound pressure level. The factor with the highest weight, in a group of machinery workers was the hearing threshold at 4 kHz frequency (Zare *et al.*, 2019).

The high prediction accuracy of the algorithms applied in the present study is a finding in agreement with reported studies. For example, in a study conducted to predict hearing loss symptoms from audiometry, using the FP-Growth (Frequent Pattern Growth) algorithm as a feature extraction technique, Noma *et al.* (2013) reported that the error rate ranged from 0 to 5.4%. In a recent study Zare *et al.* (2019) obtained prediction accuracy from 94% to 100% for different models.

Majumder and Sharma (2014) used machine learning and data classification models to investigate hearing hazards of professional drivers. The study was conducted with the use of unsupervised (Expectation Maximization, *k*-means, Linear Vector Quantization, Self Organization Map) and supervised (Naïve Bayes, Instance Based, Back Propagation Network, Radial Basis Function) learning techniques. The results have demonstrated that all the techniques, except the Radial Basis Function classifier, have shown high performance in terms of classification accuracy (Majumder, Sharma, 2014).

Nawi *et al.* (2011) applied a Gradient Descent with Adaptive Momentum (GDAM) algorithm to predict noise induced hearing loss in workers, using age, duration of work experience, and noise exposure as the main factors involved in hearing loss. The accuracy obtained in the present study in the prediction of hearing loss is close that reported by Nawi *et al.* (2011).

The present study added new data to a large body of investigations of hearing hazards in the industry. Most published studies on the modeling of hearing loss were based on audiometric data. The findings reported here show that neural data mining classification algorithms can be an effective tool for hearing hazard identification and greatly help in designing and conducting hearing conservation programs in the industry.

Acknowledgements

This study was supported by the Institute For Futures Studies In Health, Kerman University of Medical Sciences, Iran, research project number 96000838. We are grateful to the Institute for the support.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Ahmed H.O. *et al.* (2001), Occupational noise exposure and hearing loss of workers in two plants in eastern Saudi Arabia, *Annals of Occupational Hygiene*, **45**(5): 371–380, doi: 10.1093/annhyg/45.5.371.

2. Badr A., Mohammad Esmaeil S., Heidari H. (2009), Applying the data-mining technique in order to categorize the target users of the central library of isfahan university of technology (studying the motives and information seeking behaviors of them), *Iranian Journal of Information Processing and Management*, **33**(1): 275–298.

3. Golabi M.R., Akhondali A.M., Radmanesh F. (2013), Comparison of the performace of different neural networks algorithm functions in simulation of seasonal precipitation case study: selected stations of Khuzestan province [in Persian], *Journal of Geographical Sciences*, **13**(30): 151–169.

4. Golmohammadi R., Aliabadi M. (1999), *Noise and vibration engineering*, Daneshju, Hamedan.

5. Golmohammadi R., Zaman Parvar A.R., Khalili A. (2001), Assessment of the relationship between noise and hearing loss in workers Isfahan Steel Rolling Workshop, *Scientific Journal of Hamadan University of Medical Sciences*, **8**(1): 35–38.

6. Golmohammadi R., Ziad M., Atari S. (2006), Assessment of noise pollution and its effects on stone cut industry workers of Malayer District, *Iran Occupational Health Journal*, **3**(1–2): 23–27.

7. Gubbels S.P., Gartrell B.C., Ploch J.L., Hanson K.D. (2017), Can routine office-based audiometry predict cochlear implant evaluation results?, *Laryngoscope*, **127**(1): 216–222, doi: 10.1002/lary.26066.

8. ISO 1999 (2013), *Acoustics – Estimation of noise-induced hearing loss.*

9. ISO 9612 (2009), *Acoustics – Determination of occupational noise exposure – Engineering method.*

10. Kohzadi N., Boyd M.S., Kaastra I., Kermanshahi B.S., Scuse D. (1995), Neural networks for forecasting: an introduction, *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, **43**(3): 463–474, doi: 10.1111/j.1744-7976.1995.tb00135.x.

11. Krogh A., Vedelsby J. (1995), Neural network ensembles, cross validation, and active learning, *Proceedings of the 7th International Conference on Neural Information Processing Systems*, pp. 231–238, MIT Press, Cambridge, MA.

12. Larose D.T., Larose C.D. (2014), *Discovering knowledge in data: an introduction to data mining*, John Wiley & Sons.

13. Leong M.S. (2003), *Noise and vibration problems how they affects us and the industry in the Malaysian context*, Penerbit Universiti Teknologi Malaysia.

14. Majumder J., Sharma L.K. (2014), Application of data mining techniques to audiometric data among professionals in India, *Journal* of *Scientific Research and Reports*, **3**(23): 2960–2971, doi: 10.9734/JSRR/2014/12700 .

15. Masumi A., Rekabi H., Bayat A., Abshirini H. (2008), *Comparison of noise induced hearing loss and impact noise*, Med Sci, 7.

16. Mccullagh J. (2010), Data mining in sport: A neural network approach, *International Journal of Sports Science and Engineering*, **4**(3): 131–138.

17. Nassiri P., Zare S., Monazzam Mr., Pourbakht A., Azam K., Golmohammadi T. (2016), Modeling signal-to-noise ratio of otoacoustic emissions in workers exposed to different industrial noise levels, *Noise and Health*, **18**(85): 391–398, doi: 10.4103/1463-1741.195808.

18. Nawi N.M., Rehman M.Z., Ghazali M.I. (2011), Noise-induced hearing loss prediction in Malaysian industrial workers using gradient descent with adaptive momentum algorithm, *International Review on Computers and Software*, **6**(5): 740–749.

19. Noma N., Khanapi M., Ghani A., Mohamad Khir A., Noorizan Y. (2013), Predicting hearing loss symptoms from Audiometry data using FP-Growth Algorithm and Bayesian Classifier, *Australian Journal of Basic and Applied Sciences*, **7**(9): 35–43.

20. Ramos-Miguel A., Perez-Zaballos T., Perez D., Falconb J.C., Ramosb A. (2015), Use of data mining to predict significant factors and benefits of bilateral cochlear implantation, *European Archives of Oto-Rhino-Laryngology*, **272**(11): 3157–3162, doi: 10.1007/s00405-014-3337-3.

21. Safari Variani A., Ahmadi S., Zaroshani V., Ghorbanideh M. (2018), Water pump noise control using designed acoustic curtains in a residential building of Qazvin city, *Iran Occupational Health Journal*, **15**(1): 126–135.

22. Schlauch R.S., Nelson P. (2009), Puretone evaluation, [In:] *Handbook of Clinical Audiology*, 6th ed., pp. 30–49, Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia.

23. Tajic R., Ghadami A., Ghamari F. (2008), The effects of noise pollution and hearing of metal workers

in Arak, *Zahedan Journal of Research in Medical Sciences*, **10**(4): e94504.

24. WHO (1991), *Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment Programme Planning*, Geneva, June 18–21.

25. Zamanian Z., Mohammadi H., Rezaeeyani M.T., Dehghany M. (2012), An investigation of shift work disorders in security personnel of 3 hospitals of Shiraz University of Medical Sciences, 2009, *Iran Occupational Health*, **9**(1): 52–57.

26. Zamanian Z., Dehghani M., Hashemi H. (2013), Outline of changes in cortisol and melatonin circadian rhythms in the security guards of shiraz university of medical sciences, *International Journal of Preventive Medicine*, **4**(7): 825–830.

27. Zare S., Ghotbi-Ravandi M.R., Elahishirvan H., Ahsaee M.G., Rostami M. (2019), Predicting and weighting the factors affecting workers' hearing loss based on audiometric data using C5 algorithm, *Annals of Global Health*, **85**(1): 88, doi: 10.5334/aogh.2522.