The Journal of
Laryngology & Otology

cambridge.org/jlo

Main Article

Emre Soylemez takes responsibility for the integrity of the content of the paper

Cite this article: Soylemez E, Avci I, Yildirim E, Karaboya E, Yilmaz N, Ertugrul S, Tokgoz-Yilmaz S. Predicting noise-induced hearing loss with machine learning: the influence of tinnitus as a predictive factor. *J Laryngol Otol* 2024;1–6. https://doi.org/10.1017/S002221512400094X

Received: 20 June 2023
Revised: 4 March 2024
Accepted: 13 March 2024

Keywords:
machine learning; noise-induced hearing loss; occupational diseases; occupational groups

Corresponding author:
Emre Soylemez;
Email: emresoylemez@karabuk.edu.tr

Predicting noise-induced hearing loss with machine learning: the influence of tinnitus as a predictive factor

Emre Soylemez[1,2], Isa Avci[3], Elif Yildirim[3], Engin Karaboya[4], Nihat Yilmaz[5], Süha Ertugrul[5] and Suna Tokgoz-Yilmaz[6,7]

[1]Department of Audiometry, Vocational School of Health Services, Karabuk University, Karabuk, Türkiye, [2]Audiology and Speech Pathology Ph.D. Program, Health Sciences Institute, Ankara University, Ankara, Türkiye, [3]Department of Computer Engineering, Karabuk University, Karabuk, Türkiye, [4]Department of Audiology, Karabuk Training and Research Hospital, Karabuk, Türkiye, [5]Department of Otorhinolaryngology, Karabuk University, Karabuk, Türkiye, [6]Department of Audiology, Faculty of Health Sciences, Ankara University, Ankara, Türkiye and [7]Audiology, Balance and Speech Disorders Unit, Medical Faculty, Ankara University, Ankara, Türkiye

Abstract

Objectives. This study aimed to determine which machine learning model is most suitable for predicting noise-induced hearing loss and the effect of tinnitus on the models' accuracy.
Methods. Two hundred workers employed in a metal industry were selected for this study and tested using pure tone audiometry. Their occupational exposure histories were collected, analysed and used to create a dataset. Eighty per cent of the data collected was used to train six machine learning models and the remaining 20 per cent was used to test the models.
Results. Eight workers (40.5 per cent) had bilaterally normal hearing and 119 (59.5 per cent) had hearing loss. Tinnitus was the second most important indicator after age for noise-induced hearing loss. The support vector machine was the best-performing algorithm, with 90 per cent accuracy, 91 per cent F1 score, 95 per cent precision and 88 per cent recall.
Conclusion. The use of tinnitus as a risk factor in the support vector machine model may increase the success of occupational health and safety programmes.

Introduction

Despite being preventable, noise-induced hearing loss is one of the most common types of sensorineural hearing loss. Noise-induced hearing loss refers to damage to the inner ear caused by prolonged exposure to high levels of noise. The estimated worldwide prevalence of noise-induced hearing loss is 16 per cent, with a 7 per cent prevalence in Western countries and 21 per cent in developing countries.[1]

After presbycusis, noise-induced hearing loss is the second most common cause of sensorineural hearing loss.[2] The severity of noise-induced hearing loss depends on both the intensity and duration of exposure to noise, as well as individual factors. Although hearing loss typically progresses slowly, it can eventually reach moderate or even severe levels over time.

Noise-induced hearing loss can have negative impacts on workers' communication skills, work performance and quality of life. In addition, exceeding the hearing level of 40 dB is classified as a disability.[3] As an occupational disease, noise-induced hearing loss affects not only workers and employers but also government budgets. Since there is currently no medical or surgical treatment available for noise-induced hearing loss,[4,5] early diagnosis is critical in preventing some of the adverse effects. The potential usefulness of methods such as otoacoustic emissions in the early detection of noise-induced hearing loss is currently a topic of research.

Machine learning algorithms, a new group of statistical methods primarily used in software and engineering fields, are preferred for analysing non-linear multidimensional complex events and uncertain information.[6] Machine learning algorithms have the ability to automatically generate new rules based on input data and can estimate unknown data that may be difficult to define manually.[7] In other words, a dataset with known risk factors (input) and outcomes (output) can be taught to machine learning algorithms.

After training, machine learning algorithms can estimate outputs for new inputs that are presented to them. Some studies have investigated the use of machine learning algorithms in the detection of occupational diseases.[6,8] Environmental and individual factors (such as age and noise intensity) that play a role in the formation of noise-induced hearing loss can be used to predict hearing loss using machine learning algorithms.[8,9]

Few studies in the literature predict noise-induced hearing loss with machine learning algorithms.[6,8,9] Unlike these studies, we included tinnitus as an input in our study. This study aimed to determine which machine learning algorithm is more suitable for predicting noise-induced hearing loss and the effect of tinnitus on the models' accuracy.

According to our hypothesis, using tinnitus, one of the early markers of noise-induced hearing loss, as an input may increase the accuracy of machine learning algorithm models.

Materials and methods

Participant selection

This prospective study was carried out on metal industry workers who presented at the otolaryngology out-patient clinic and were referred for hearing tests. All workers had been working in the machinery area for at least one year and were exposed to noise at a minimum level of 85 dB (A).

During the audiological examination, a detailed anamnesis was obtained from each participant. The questionnaire included questions on the following parameters: age (years), duration of exposure to noisy environments (years), frequency of ear protection equipment use (never, sometimes or continuously), smoking status (yes or no, and if yes, how many years) and the presence of tinnitus (right ear, left ear or bilateral). Workers with perforation of the eardrum, type B and C tympanograms, conductive and mixed hearing loss, and hearing loss due to another reason (congenital hearing loss, sudden hearing loss, etc.) were not included in the study.

A pure tone audiometry test was administered to all 200 male workers to determine their hearing thresholds. The anamnesis data, which included risk factors for noise-induced hearing loss, were used as inputs to train the machine learning algorithms to estimate the probability of hearing loss (output) in these workers (as shown in Figure 1).

We obtained both verbal and written consent from all participants in accordance with the Declaration of Helsinki. The study was approved by the ethics committee of Karabuk University (approval number: 2022/838).

Audiological evaluation

A pure tone audiometry test was administered bilaterally to all workers using the Madsen Astera (GN Otometrics, Taastrup, Denmark) in a soundproof room. The air-conduction hearing thresholds in the range of 250–6000 Hz were determined using TDH 39 supraaural headphones (Telephonics Corp., New York City, USA), while the bone-conduction hearing thresholds in the range of 500–4000 Hz were determined using the Radioear B71 (Radioear Corp., USA) bone vibrator. The tympanometric examination was performed with an Interacoustics AZ 26 (Interacoustics, Middelfart, Denmark) with a 226-Hz probe tone. The pure tone average (PTA) was calculated by taking the arithmetic mean of the frequency band thresholds (500, 1000, 2000 and 4000 Hz). A PTA greater than 20 dB in at least one ear was considered to indicate hearing loss.

Statistical analysis and machine learning models

The International Business Machines Statistical Package for the Social Science 21 (IBM SPSS Corporation, Armonk, NY, USA) was used for statistical analysis. Variables that met the normality assumption were presented as mean ± standard deviation, and variables that did not meet the normality assumption were presented as median (minimum–maximum). The compliance of the variables with normality distribution was checked with the Shapiro–Wilk test. The Mann–Whitney U test and chi-square test were used to compare hearing loss groups and risk factors. In all statistical analyses, p less than 0.05 was accepted as the statistical significance level.

Python programming language (Version 3.7) was used to develop the machine learning algorithm. Machine learning algorithms can be classified as supervised, unsupervised or reinforced reinforcement learning. Supervised learning algorithms are used in classification and regression problems.[10] In this study, k-nearest neighbour, decision tree, random forest, support vector machine, logistic regression and XGBoost algorithms, which are considered supervised algorithms, were used. The performance of these models was evaluated using accuracy, precision, F1 score, recall and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.[11]

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

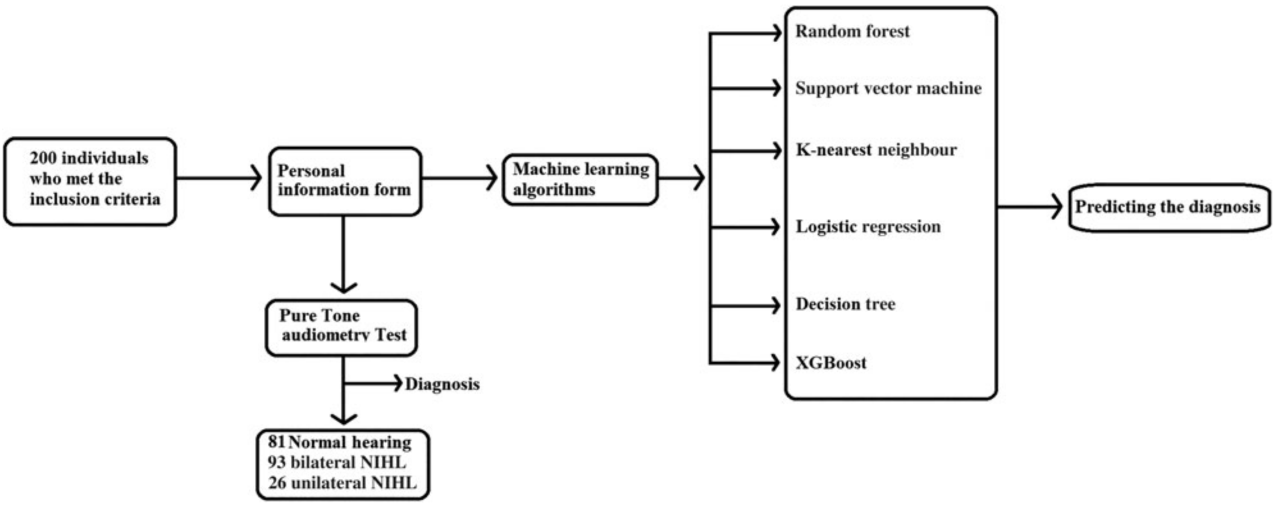

Figure 1. Flowchart detailing test steps and machine learning algorithms applied to workers. Two hundred workers were included in the study. A personal information form was applied to these workers. Questions in the fact sheet included risk factors for noise-induced hearing loss: age (years), working duration in noisy environments (years), using hearing protection apparatus (never, sometimes, continuously), smoking status (yes or no, if yes, how many years) and tinnitus (right, left or bilateral). NIHL = noise-induce hearing loss.

https://doi.org/10.1017/S002221512400094X Published online by Cambridge University Press

$$\text{F1 score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

where TP denotes true positive, TN denotes true negative, FN denotes false negative and FP denotes false positive.

These algorithms were trained using 160 data points (80 per cent of the data) and the success of the algorithms was tested with the remaining 40 data points (20 per cent of the data).

K-nearest neighbour

The k-nearest neighbour algorithm classifies newly obtained data by assigning it to the class of the nearest similar neighbours. It uses two basic metrics: distance and K neighbourhood ratios.[12]

Decision tree

This is a supervised learning algorithm used in classification and regression problems. The decision tree algorithm tries to solve the problem by representing the data in tree form. Each decision node corresponds to a variable and each leaf node corresponds to its target tag. The sequence followed while creating the tree is: (1) the most suitable variable is put at the root of the tree and should be as simple as possible; (2) the dataset is divided into subsets; and (3) these two operations are continued until the leaf reaches the nodes.[13]

Random forest

This ensemble learning algorithm combines the decisions of many independent multivariate trees. Random forest is a classification model that attempts to create more accurate and compatible models using multiple decision trees.[14] It uses averaging to improve forecast accuracy and control overfitting.

Support vector machine

The support vector machine is a learning algorithm that can be used for classification and regression analysis. Data points are separated by a line or hyperplane to divide them into two or more classes. The gap between the two classes should be as large as possible to reduce errors during classification.[14]

Logistic regression

Logistic regression is a regression method for classification. It is used to reveal the effect of one or more variables on the overall outcome. Due to this feature, it is the preferred research approach for identifying the most influential variable among the independent variables and predicting the output variable. While performing logistic regression, attention should be paid to the variables' independence and the validity of assumptions to ensure appropriate modelling.[15]

XGBoost

XGBoost is a reinforced tree algorithm model based on gradient-boosting principles. Compared with other approaches, XGBoost applies more systematic model reinforcement to control overfitting, thus aiming to improve performance.[16] This algorithm makes corrections to errors after making predictions. The performance of XGBoost depends on parallelism and hardware optimisation.

Results

The workers had a mean age of 39.96 ± 10.96 years (range, 19–57 years). Eighty-one workers (40.5 per cent) had normal bilateral hearing and 119 (59.5 per cent) had hearing loss. Among those with hearing loss, 93 (78.15 per cent) had bilateral hearing loss and 26 (21.84 per cent) had unilateral hearing loss. The hearing thresholds of workers with and without hearing loss are presented in Table 1.

We compared age, exposure duration to industrial noise, smoking, use of ear protection equipment and the presence of tinnitus between workers with and without hearing loss. Ageing, duration of exposure to industrial noise, smoking, not using ear protection equipment and tinnitus were risk factors for noise-induced hearing loss ($p < 0.05$). Conditions that are risk factors for noise-induced hearing loss are presented in Table 2.

The support vector machine was the best-performing algorithm, with 90 per cent accuracy, 91 per cent F1 score, 95 per cent precision, 88 per cent recall and 90.6 per cent area under the receiver operating characteristic curve. The confusion matrix and the receiver operating characteristic curve of the support vector machine model are presented in Figure 2a,b. Our study showed that age and tinnitus contributed the most to the overall result in the support vector machine model (Figure 2c). The decision tree algorithm was the second-best-performing algorithm, with 87.5 per cent accuracy, 89 per cent F1 score, 100 per cent precision, 79 per cent recall and 84.8 per cent area under the receiver operating characteristic curve.

Table 1. Pure tone hearing thresholds (mean ± standard deviation) for the right and left ears according to frequencies ($N = 200$)

Frequency	With HL	Without HL
250 Hz		
– Left (dB)	17.26 ± 7.55	12.67 ± 5.70
– Right (dB)	18.36 ± 7.78	12.40 ± 4.40
500 Hz		
– Left (dB)	18.36 ± 9.09	11.23 ± 4.14
– Right (dB)	18.52 ± 8.52	11.48 ± 4.21
1000 Hz		
– Left (dB)	18.27 ± 10.42	9.81 ± 4.83
– Right (dB)	17.81 ± 10.20	9.25 ± 5.13
2000 Hz		
– Left (dB)	23.41 ± 17.44	10.18 ± 6.29
– Right (dB)	21.42 ± 15.10	8.02 ± 4.78
4000 Hz		
– Left (dB)	55.50 ± 18.66	21.23 ± 8.85
– Right (dB)	52.31 ± 20.47	21.11 ± 11.56
6000 Hz		
– Left (dB)	51.42 ± 23.08	20.61 ± 11.49
– Right (dB)	51.68 ± 24.29	21.79 ± 13.63
Pure tone average		
– Left (dB)	28.88 ± 10.00	13.11 ± 3.46
– Right (dB)	27.51 ± 10.13	12.46 ± 3.86

HL = hearing loss

https://doi.org/10.1017/S002221512400094X Published online by Cambridge University Press

Table 2. Conditions that are risk factors for noise-induced hearing loss ($N = 200$)

Risk factor	With HL	Without HL	*p*
Age (median (range); years)	48.0 (19.0–57.0)	30.0 (19.0–50.0)	<0.001[a]
Working duration (median (range); years)	16.0 (1.0–37.0)	5.0 (1.0–30.0)	<0.001[a]
Smoking (median (range); years)	10.0 (0–37.0)	0.5 (0–23.0)	0.002[a]
Using hearing protection apparatus (*n* (%))			<0.001[b]
– Never	64 (53.8)	32 (39.5)	
– Sometimes	45 (37.8)	24 (29.6)	
– Continuously	10 (8.4)	25 (30.9)	
Tinnitus (*n* (%))	61 (51.3)	3 (3.7)	<0.001[b]

HL = hearing loss; *a* = Mann–Whitney *U* test; *b* = chi-square test

The k-nearest neighbour algorithm was the worst-performing algorithm, with 80 per cent accuracy, 83 per cent F1 score, 86 per cent precision, 79 per cent recall and 80.4 per cent area under the receiver operating characteristic curve. The performances of the logistic regression, random forest, support vector machine, decision tree, k-nearest neighbour and XGBoost models in test sets are presented in Table 3.

Discussion

Long-term exposure to workplace noise affects the inner ear in three stages.[17,18] The first stage involves minor damage to hair cells, which occurs with the first exposure to noise. This damage cannot be detected with a pure tone audiometry test. However, individuals may experience auditory disturbances, such as tinnitus and hyperacusis, as well as non-auditory disorders, including headaches, fatigue and stress.[17]

The second stage occurs when the noise exposure continues for months or years and damages the basal part of the cochlea due to the resonance frequency effect in the external auditory canal. This damage can be detected as acoustic notches on the audiogram at 3, 4 or 6 kHz. Speech intelligibility is usually not severely affected at this stage, and the damage may go unnoticed without a hearing test. The severity of noise-induced hearing loss may rapidly increase and reach a plateau at the end of this stage.

The third stage occurs with long-term exposure to chronic noise, which often leads to a decline in communication skills, and seeking treatment for hearing loss becomes necessary.[17] The goal of workplace hearing screenings, which are mandated by occupational health and safety regulations, is to detect

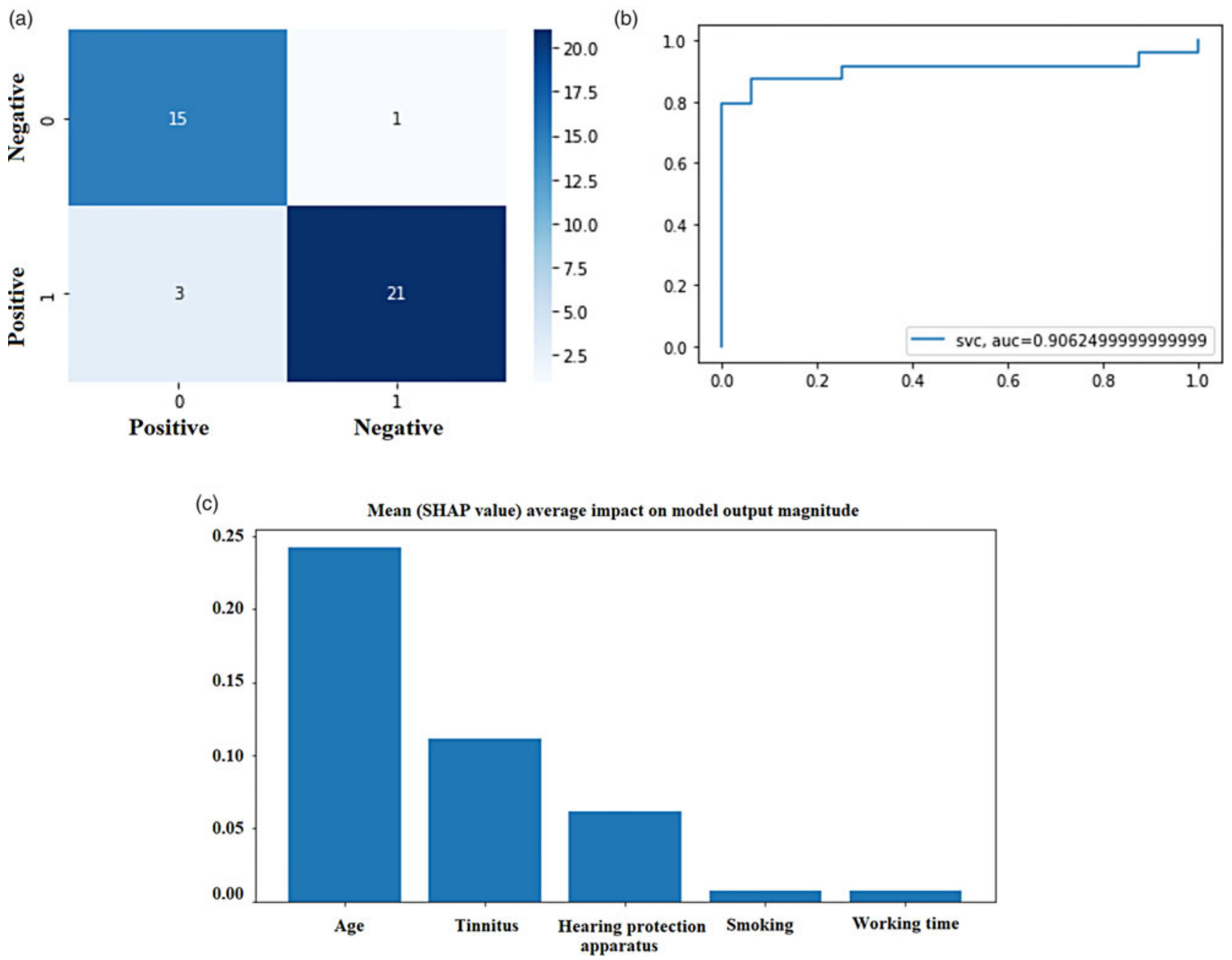

Figure 2. (a) The confusion matrix of the support vector machine model. (b) Receiver operating characteristic curve of the support vector machine model. (c) SHAP analysis of the support vector machine. SVC = Support Vector Classification; AUC = Area Under the Curve; SHAP = SHapley Additive exPlanations

Table 3. Test performances of logistic regression, random forest, support vector machine, decision tree, k-nearest neighbours and XGBoost

Models	Precision	Recall	F1 score	Accuracy	AUC-ROC
Logistic regression	0.91	0.83	0.87	0.85	0.885
Random forest	0.91	0.83	0.87	0.85	0.863
Support vector machine	0.95	0.88	0.91	0.90	0.906
Decision tree	1.00	0.79	0.89	0.875	0.848
K-nearest neighbour	0.86	0.79	0.83	0.80	0.804
XGBoost	0.84	0.88	0.86	0.825	0.899

AUC-ROC = The Area Under the Curve of the Receiver Operating Characteristic Curve

noise-induced hearing loss in its early stages and implement necessary measures promptly.

This study utilised data from 200 workers, with noise-induced hearing loss risk factors as input and hearing test results as output, to predict the likelihood of noise-induced hearing loss using machine learning algorithms. Of the 200 workers, the data of 160 (80 per cent) were used as training data and the data of 40 (20 per cent) as test data. Six machine learning algorithms (k-nearest neighbour, decision tree, random forest, support vector machine, logistic regression and XGBoost) were trained using the training data. The overall accuracy of the six models ranged from 80 to 90 per cent, with the support vector machine performing best, with accuracy of 90 per cent.

Noise-induced hearing loss is a multifactorial disease that arises from the interplay of genetic, individual and environmental factors. Nevertheless, the biological damage incurred by individuals is linked to the total amount of noise (the fundamental energy level).[18] The equal-energy principle posits that equal energy exposure leads to an equal amount of biological damage, which is determined by the sound pressure level and duration of noise exposure, therefore the sound pressure level and exposure time are crucial risk factors for hearing loss.

In our study, all participants were machinery area workers in the metal industry and their exposures to noise levels were similar. We therefore evaluated the use of ear protection equipment (input) as a potential risk factor. Ear protection equipment is designed to reduce the intensity of noise before it reaches the inner ear. Regular and continuous use of ear protection equipment can prevent up to 30 per cent of hearing loss.[19] Ramakers *et al.* showed that individuals who used ear protection equipment during outdoor music festivals reported less temporary hearing loss and tinnitus than those who did not.[20]

The other risk factors that we used as inputs in our study to train machine learning algorithms and predict noise-induced hearing loss are age and smoking. Ageing causes degenerations in the peripheral and central auditory systems as well as in all tissues and cells. Chronic workplace noise does not directly damage the cochlea but leads to the production of reactive oxygen species and other free radical molecules in the cochlea, the possible cause of which is metabolically overactive cochlear mitochondria, ionic fluxes and ischaemic reperfusion.[21]

Nicotine in cigarettes increases the amount of free radicals and reactive oxygen species, triggering oxidative damage similar to the effect of chronic noise exposure. It also stimulates the production of nuclear factor kappa B, which plays a role in inflammatory processes and cell damage.[22] Consequently, smoking and noise exposure act synergistically to increase the risk of hearing loss. Tao *et al.* reported that the mean hearing thresholds at 4 and 6 kHz were higher in smokers than non-smokers, and the incidence of high-frequency hearing loss was higher in smokers (48.9 per cent) than in non-smokers (33.8 per cent).[23]

There are several studies in the literature that have used machine learning algorithms to predict noise-induced hearing loss.[6,24] Zhao *et al.* estimated the hearing test results of 1113 workers in 17 different factories using four machine learning models (support vector machine, neural network multilayer perceptron, random forest and adaptive boosting).[6] The researchers used the age of the workers, exposure time to noise, A-weighted equivalent sound pressure level and median kurtosis as inputs. The best-performing algorithm in the study was the support vector machine model, with an accuracy of 80.1 per cent, while the other three algorithms had accuracies ranging from 78 to 79 per cent.

Similarly, Farhadian *et al.* estimated the hearing test results of 210 workers in a steel factory using artificial neural networks and logistic regression.[24] In this study, the age of the workers, noise exposure level, work experience, use of ear protection equipment and smoking status were used as inputs. The authors reported that the accuracy of artificial neural networks was 88.6 per cent in predicting hearing loss and was better than logistic regression.

In our study, we aimed to detect hearing loss in metal industry workers using the k-nearest neighbour, decision tree, random forest, support vector machine, logistic regression and XGBoost algorithms. Similar to the findings of Zhao *et al.*,[6] the support vector machine algorithm performed best and the accuracy rate was 90 per cent. The accuracies of the other algorithms were between 80 and 87.5 per cent. The performance of the algorithms can vary based on factors such as the number and type of inputs used, the weight ratios of the inputs and their correlation with the outputs. Unlike previous studies, we also used the presence of tinnitus in workers as an input.

We found hearing loss in 61 (95.3 per cent) of 64 (32 per cent) workers with tinnitus. This finding is consistent with previous reports that the prevalence of tinnitus is higher in workers exposed to excessive noise and can reach up to 80 per cent in military personnel.[25] Indeed, tinnitus was the second most significant variable affecting the success rate in our study, and adding tinnitus as an input may have increased the accuracy rate of the support vector machine model.

Another study utilized the C5 algorithm to estimate the degree of noise-induced hearing loss and investigate the factors affecting it.[26] The authors reported that the 4 kHz frequency had the highest impact, accounting for 22 per cent of the estimated hearing loss degree according to the C5 algorithm. In our study, age had the most significant effect weight in predicting noise-induced hearing loss at 24 per cent, while working time in noisy environments had the lowest effect weight at 0.7 per cent.

Recently developed machine learning algorithms and artificial neural networks have become very interesting when applied to occupational diseases, such as noise-induced hearing loss. Noise-induced hearing loss is one of the most common occupational diseases and is mainly influenced by environmental factors. Our study demonstrated that the risk of noise-induced hearing loss can be predicted cheaply and quickly using environmental factors and workers' characteristics in a machine learning algorithm (support vector machine model).

- Machine learning can be used to predict diseases
- Noise-induced hearing loss can also be predicted with machine learning
- Workers' age, working duration, smoking and earplug usage status were used as inputs
- The study achieved 80.1 and 88.6 per cent accuracy with support vector machine and neural networks, respectively
- Tinnitus was used as an input in this study and 90 per cent accuracy was achieved with the support vector machine model

Furthermore, we showed that the onset age of noise-induced hearing loss can be detected approximately when the existing risk factors are implemented in a machine learning algorithm. Future studies could incorporate hearing screening scales and other diseases, such as metabolic diseases, that may affect noise-induced hearing loss as inputs and investigate the accuracy of the algorithms.

Conclusion

Incorporating early markers of hearing loss, such as tinnitus, into machine learning algorithms may enhance the prediction ratio of the models. The support vector machine algorithm, which has the highest accuracy, can be used in the early detection of noise-induced hearing loss. Thus, the success of occupational health and safety programmes for employees exposed to occupational noise can be increased.

Acknowledgments. We acknowledge the volunteers for their contribution to this study.

Competing interests. None declared

References

1 Nelson DI, Nelson RY, Concha-Barrientos M, Fingerhut M. The global burden of occupational noise-induced hearing loss. *Am J Ind Med* 2005;**48**:446–58
2 Rabinowitz PM. Noise-induced hearing loss. *Am Fam Physician* 2000;**61**:2749–56
3 Schmucker C, Kapp P, Motschall E, Loehler J, Meerpohl JJ. Prevalence of hearing loss and use of hearing aids among children and adolescents in Germany: a systematic review. *BMC Public Health* 2019;**19**:1277
4 Manar MK, Shukla SP, Mohan U, Singh SK, Verma V. Validation of a questionnaire to identify noise-induced hearing loss among drivers. *J Family Med Prim Care* 2019;**8**:1196–201
5 Lapsley Miller JA, Marshall L, Heller LM, Hughes LM. Low-level otoacoustic emissions may predict susceptibility to noise-induced hearing loss. *J Acoust Soc Am* 2006;**120**:280–96
6 Zhao Y, Li J, Zhang M, Lu Y, Xie H, Tian Y *et al.* Machine learning models for the hearing impairment prediction in workers exposed to complex industrial noise: a pilot study. *Ear Hear* 2019;**40**:690–9
7 McKearney RM, MacKinnon RC. Objective auditory brainstem response classification using machine learning. *Int J Audiol* 2019;**58**:224–30
8 Chen F, Cao Z, Grais EM, Zhao F. Contributions and limitations of using machine learning to predict noise-induced hearing loss. *Int Arch Occup Environ Health* 2021;**94**:1097–111
9 Mrena R, Ylikoski M, Mäkitie A, Pirvola U, Ylikoski J. Occupational noise-induced hearing loss reports and tinnitus in Finland. *Acta Otolaryngol* 2007;**127**:729–35
10 Hasan M, Balbahaith Z, Tarique M. Detection of SQL injection attacks: a machine learning approach. In: *2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*. Ras Al Khaimah: IEEE, 2019;1–6.
11 Hosam E, Hosny H, Ashraf W, Kaseb AS. SQL injection detection using machine learning techniques. In: *2021 8th International Conference on Soft Computing & Machine Intelligence (ISCMI)*. Cairo: IEEE, 2021;15–20.
12 Bui XN, Jaroonpattanapong P, Nguyen H, Tran QH, Long NQ. A novel hybrid model for predicting blast-induced ground vibration based on k-nearest neighbors and particle swarm optimization. *Sci Rep* 2019;**9**:13971
13 Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry* 2015;**27**:130–5
14 Bandyopadhyay N, Jadhav A. Churn prediction of employees using machine learning techniques. *Tehnički Glasnik* 2021;**15**:51–9
15 Stoltzfus JC. Logistic regression: a brief primer. *Acad Emerg Med* 2011;**18**:1099–104
16 Jain R, Nayyar A. Predicting employee attrition using XGBoost machine learning approach. In: *International Conference on System Modeling & Advancement in Research Trends (SMART)*. Moradabad: IEEE, 2018;113–20.
17 Hirose K, Discolo CM, Keasler JR, Ransohoff R. Mononuclear phagocytes migrate into the murine cochlea after acoustic trauma. *J Comp Neurol* 2005;**489**:180–94
18 Le TN, Straatman LV, Lea J, Westerberg B. Current insights in noise-induced hearing loss: a literature review of the underlying mechanism, pathophysiology, asymmetry, and management options. *J Otolaryngol Head Neck Surg* 2017;**46**:41
19 Davies H, Marion S, Teschke K. The impact of hearing conservation programs on incidence of noise-induced hearing loss in Canadian workers. *Am J Ind Med* 2008;**51**:923–31
20 Ramakers GG, Kraaijenga VJ, Cattani G, van Zanten GA, Grolman W. Effectiveness of earplugs in preventing recreational noise-induced hearing loss: a randomized clinical trial. *JAMA Otolaryngol Head Neck Surg* 2016;**142**:551–8
21 Kamogashira T, Fujimoto C, Yamasoba T. Reactive oxygen species, apoptosis, and mitochondrial dysfunction in hearing loss. *Biomed Res Int* 2015;**2015**:617207
22 Istrate M, Hasbei-Popa M, Iliescu DA, Ghita AC, Ghita AM. Effects of cigarette smoking on sensorineural hearing impairment and age related macular degeneration. *Tob Prev Cessat* 2021;**7**:55
23 Tao L, Davis R, Heyer N, Yang Q, Qiu W, Zhu L *et al.* Effect of cigarette smoking on noise-induced hearing loss in workers exposed to occupational noise in China. *Noise Health* 2013;**15**:67–72
24 Farhadian M, Aliabadi M, Darvishi E. Empirical estimation of the grades of hearing impairment among industrial workers based on new artificial neural networks and classical regression methods. *Indian J Occup Environ Med* 2015;**19**:84–9
25 Yankaskas K. Prelude: noise-induced tinnitus and hearing loss in the military. *Hear Res* 2013;**295**:3–8
26 Zare S, Ghotbi-Ravandi MR, ElahiShirvan H, Ahsaee MG, Rostami M. Predicting and weighting the factors affecting workers' hearing loss based on audiometric data using C5 algorithm. *Ann Glob Health* 2019;**85**:88