



Machine learning based assessment of auditory threshold perception in human beings

H. S. Sanjay¹ · Basavaraj V. Hiremath¹ · B. S. Prithvi² · P. A. Dinesh³

Received: 29 August 2019 / Accepted: 20 December 2019 / Published online: 2 January 2020
© Springer Nature Switzerland AG 2020

Abstract

This study presents a novel approach to assess the perception of auditory Absolute threshold (ATTh) in healthy individuals exposed to noise and solvents in their occupational environment using machine learning approaches. 396 subjects with no known history of auditory pathology were chosen from three groups, namely, employees from Chemical Industries (CI), Fabrication Industries (FI), and professional Basketball Players (BP), with each category having 132 subjects. Absolute Threshold Test (ATT) was developed using MATLAB and the experiment was conducted in a silent, noise-free environment. ATTh was obtained twice, during the commencement and conclusion of the employees' workshift in CI and FI. For BP, ATTh was obtained before and after their basketball training sessions and was used as features for binary SVM classification approach, in which the RBF kernel-based technique was found to provide maximum accuracy as compared to linear and quadratic approach. For three-class classification, MLP neural network with Levenberg–Marquardt training function in the hidden layer and Mean Square Error function in the output layer was found to be optimal along with k-Nearest Neighbor (kNN) and Support Vector Machine (SVM) approach using Radial Basis Kernel Function (RBF), in which, an accuracy of 81.06% was observed in kNN approach and 92.4% using MLP neural network approach, whereas SVM yielded an accuracy of 93.94% in the classification of the subjects into CI, FI and BP, showing that the SVM outperformed kNN and MLP neural network for healthy subjects based on their occupational exposure/professional sports training. Such machine learning approaches could further be probed into, to improve the accuracy of classification. Also, such techniques can help in real-time classification of subjects based on their occupational exposure so as to predict and prevent plausible permanent hearing dysfunction due to occupational exposure as well as to aid in sports rehabilitation and training programs to assess the auditory perceptive abilities of the individuals.

Keywords Hearing perception · Absolute threshold test · k-Nearest neighbor · Multi-layer perceptron · Levenberg–Marquardt · Support vector machine · Radial basis kernel function

1 Introduction

Sound is the most important channel of communication between two living beings. Perception of sound is, thus, a prime factor in understanding the meaning of the information being communicated between the source and the listener. Sound is often defined in terms of both physical as well as various psychological factors such as

pitch, loudness, melody and gaps. Auditory dysfunction is a serious problem in humans but is seldom considered for preventive care [1]. This is more common in developing countries such as India. Although auditory diagnosis is performed using audiometry and related tests, they often fail to provide a preventive diagnosis. Also, these tests are subjective in nature. This problem could be tackled by the inclusion of auditory perceptual aspects into the tests

✉ H. S. Sanjay, sanjaymurthyhs@gmail.com | ¹Department of Medical Electronics, M S Ramaiah Institute of Technology, Bangalore, India. ²MediNXT Technologies, Bangalore, India. ³Department of Mathematics, M S Ramaiah Institute of Technology, Bangalore, India.



being developed to ascertain the hearing of a subject who is often exposed to occupational noises which could be a plausible cause for an irreversible auditory damage [2]. The present work emphasizes on one such approach with the aid of absolute threshold tests, and the results are assessed using machine learning approaches so as to arrive at a meaningful conclusion about the results obtained in terms of the onset of a plausible hearing deterioration in the subjects considered.

2 Background

Hearing and perception are often in synchrony for any process to be completed in terms of sensation in human beings. While aspects such as intensity and frequency are often used to quantify and localize the sound, perception of the same has a direct relationship with the understanding of the meaning of the sound perceived and is hence a qualitative aspect [3]. For instance, music, which is melodious to an individual may seem very absurd to another due to the variation in the perception of the sound. Hearing is disturbed by various environmental factors, one such example being the occupational exposure to noise in fabrication industries. Another important, but often neglected domain is the exposure to solvents in chemical industries which is known to cause permanent damage to the cochlea and thereby affecting the hearing of such employees. Hence, there lies an urgent need to address the concerns of such industries, wherein the detection of a plausible onset of auditory perception deterioration can be very helpful in rehabilitation of employees so as to avert permanent auditory dysfunction as an occupational hazard. This article emphasizes the assessment of the Absolute hearing threshold of the employees of chemical and fabrication industries to discern their abilities in terms of auditory perception. At the other end of the spectrum, the results were compared with professional basketball players, as literature suggests that regular training and practice of sports such as basketball and table tennis have resulted in an improvement of auditory perception in such players [4].

2.1 Absolute threshold

Absolute threshold (ATTh) is best measured on a decibel scale and can be varied with respect to the intensity using different patterns such as linear and staircase approaches. Other novel ways of varying the tones include parametric estimation by sequential testing and Maximum Likelihood Procedure, which is followed in the present work. Conventionally, pure tones of 500 ms duration is designed with a 1 kHz frequency. When such sounds are provided to the subjects, they are required to respond to the variations by

normal subjective approaches based on which the ATTh is obtained [5]. Detection based approaches confine themselves to find out whether a subject has heard the tone or not by incorporating a simple YES–NO approach, whereas discrimination-based techniques probe further to differentiate a given pair of tones as inputs. These paradigms are developed using adaptive techniques wherein the sound is varied based on the previous response of the individual, while the non-adaptive techniques do not use any such feedback to vary the sound and the same is achieved by a predefined protocol such as step, ramp or staircase approaches [6].

2.2 Occupational exposure

Exposure to noise in the work environment is perhaps one of the most important reasons for hearing dysfunctions in employees of fabrication industries and others who work with noise producing machines. In case of paint and chemical industries wherein machines such as centrifuges and boilers are common for numerous chemical processes, aromatic chemicals/solvents too play an important role in the deterioration of hearing abilities of such employees [7]. There is adequate proof to prove that a combined effect of exposure to noise as well as solvents cause a higher degree of hearing loss in the employees of solvent industries with constant exposure. This is due to the ill-effect of ototoxic chemicals (often divided into workplace chemicals and medications) which are proved to have a synergic effect on our hearing system. Solvents damage the neurological pathways as well as the cochlea resulting in an irreversible hearing loss which is more prominent if there happens to be an exposure to a combination of noise and solvents [8]. More than 700 chemicals are regarded to be ototoxic in nature and this fact is often neglected in case of cochleotoxic non-steroidal anti-inflammatory drugs which often cause sensorineural hearing loss [9].

While there have been discussions on plausible disadvantages of noise exposure, there is a positive dimension to a controlled exposure to sound, if not noise, which often results in an improvement in hearing perception which is observed in case of sports training and rehabilitation programs. For instance, in table tennis, the players seem to be ready to return the ball to the opponent while the ball has just been served even without visually perceiving the ball. This happens just by an auditory perception of the ball being hit by the racket of the opponent and this sound prepares the players to be ready to return it by anticipating its arrival trajectory and speed. Else, by the time the players prepare to return the serve, the ball would have gone past them and they will have lost a point in the game. This sense of preparedness comes due to the response to the sound first, and then to a visual response. This pattern

improves with regular training of sports. Such subjects are aptly considered as positively conditioned subjects in the present experiment [10].

2.3 Data analytics

Data analytics encompasses novel approaches such as cleansing, transforming and modelling of a given dataset. This helps in arriving at meaningful conclusions for decision making tasks. While approaches such as data mining help in developing prediction and assessment-based models, analytics follow more of a descriptive approach. This is supported by visualization to understand the results better. Various machine learning approaches are used to recognize the patterns to quantify the results, and neural network based techniques are useful in decision making by mimicking the natural patterns of the brain [11]. Often termed as artificial intelligence, this approach helps the model to learn, similar to biological neural networks. One such approach is the multilayer perceptron (MLP) which is a feed-forward artificial neural network consisting of a minimum of three nodes (input layer, hidden layer and output layer). Each node denotes a non-linear activation function and this involves a supervised learning approach with back propagation for the purpose of training. This can analyze the data which is not linearly separable and hence finds applications in various biological data analytics as well [12]. The present work confined to the usage of pattern recognition-based approaches in terms of kNN and SVM. Further, neural networks were incorporated with MLP for data analytics.

3 Materials and methods

The Absolute Threshold (ATTh) of hearing is the minimal intensity level of a pure tone sound which is heard by an average human being with normal hearing in the absence of any other sound [13]. The range of hearing is defined to be 20–20 kHz and any sound with frequency outside this range goes unperceived thereby not affecting the physiological aspects of hearing. The ATTh is frequency dependent and is highly sensitive for a range of 1–5 kHz during which the hearing threshold may extend up to –9 dB SPL [14].

3.1 Maximum likelihood procedure

The traditional approach to vary the input attributes of sound would be in staircase or ramp formats in incremental/decremental methods. Maximum likelihood procedure provides a novel approach to such variations wherein the response of the subject is used to obtain

the likelihood of the threshold value and the succeeding input is provided based on the ultimate estimate which is simply the position of the maximum likelihood. Maximum likelihood procedure commences by the provision of the sound above a predefined intensity threshold and the subjects are asked to respond if they were able to hear the sound at this level. The intensity level of the next auditory input is varied based on this current intensity level [15]. A psychometric function is then estimated based on the likelihood of the occurrence of the hearing function. The likelihood of occurrence of a k th hypothesized function ($L(A_k)$) in terms of a binomial log-likelihood approach is essential for the correct prediction. This likelihood of occurrence can be obtained as shown in (1).

$$L(A_k) = \sum_{i=1}^n [R \{ \ln A(x_n) \} + E \{ \ln \{ 1 - A(x_n) \} \}] \quad (1)$$

where $A(x_n)$ = Probability of the correct response for a given threshold level; n = The trial number; R = The occurrence of the correct answer (equated to 1); E = The occurrence of the wrong answer (equated to 0)

In the present context, (1) represents the likelihood of occurrence of the correct response for a predefined threshold level in terms of ATTh. In other words, $L(A_k)$ provides the correctness of the response of the subject in terms of the true positives of the predicted ATTh values. R corresponds to the number of occurrences of the correct prediction, which has been equated to 1 and E denotes the number of wrong responses which is equated to 0, while predicting the likelihood of occurrence of a correct response. The likelihood occurring the maximum number of times is concluded to be the likelihood of the psychometric function [16]. The threshold value for a given trial is found, based on the previous trial, with the aid of the psychometric function developed, as given in (2)

$$y = a - \left(\frac{1}{b} \right) \times \ln \left[\frac{1 - c - d}{p \text{ target} - c} - 1 \right] \quad (2)$$

where a = The midpoint of the likelihood; b = Slope of the psychometric function; c = False alarm rate; d = Rate of lapse of attention of the subject

In this case, y denotes the psychometric function, which is a simple inferential model observed in detection as well as discrimination-oriented applications. This provides an insight into the relationship between the stimulus as well as the response. Hence (2) could be considered as an application of a generalized linear model to the given psychophysical information, in the present context being ATTh.

3.2 Experimental paradigm

The ATT experiment developed finds its use in the auditory perception assessment while varying the intensity of the sound input. A pure tone of 1 kHz and 500 ms duration was developed to be the auditory stimulus and the intensity was varied using Maximum likelihood procedure with regard to the response dependent on the perceptive abilities of the subjects. The tone encompassed cosine ramps of 10 ms duration and 3 blocks of such tones were designed with 15 trials in every block. The subjects were asked to press “1” on the keyboard if the tone was heard and “0” otherwise. The maximum number of trials was fixed to be 100 and the range of intensity was 30 dB to 110 dB. At the beginning, the other constant values were $b = 1$, $c = 0$, $p\text{-target} = 0.631$. The initial block of sound was provided at 30 dB. Based on the subject response, the psychometric function was estimated using (1) and the intensity for the succeeding trial was found using (2). The experiment lasted for 3 min overall [17].

3.3 k-Nearest neighbor approach

The k-Nearest Neighbor (kNN) is perhaps the simplest classification approach in cases without any prior knowledge about the data distribution. It follows a non-parametric approach and is useful in case of classification as well as regression-oriented applications. Here, the input is composed of k nearest training samples in a given feature space. In case of kNN classifiers, a given object is classified based on the plurality vote of its neighbor along with the object being assigned to the class most common among its k nearest neighbors. kNN is a lazy learning approach where the function is approximated locally and all the computations are deferred till the classification is done. kNN stores all available cases and classifies new cases based on a similarity measure. It uses supervised learning technique. The final prediction of the type of data point is based on the majority type of its neighbors. kNN is often found to be sensitive to the local structure of the data.

A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. For continuous variables, Minkowski distance function is used, as shown in (3) with $q = 2$. Also If A and B would be represented by feature vectors $A(x_1, x_2, \dots, x_k)$ and $B(y_1, y_2, \dots, y_k)$, then x_i would be the *i*th feature of x and y_i would be the *i*th feature of y [18].

$$\left(\sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q} \tag{3}$$

where k = The dimensionality of the feature space; q = The order ($q = 2$ for the present experiment)

3.4 Support vector machine (SVM) classifier

SVM is a machine learning approach used to classify and categorize given data to be tested with the aid of a hyperplane from the training data. The training samples along the hyperplane near the boundary are termed as support vectors. The distance between the boundary hyperplanes and the support vectors are called as margins. The SVM approach conventionally incorporates the training and testing phases with corresponding data to be used as well. During the training phase, the target values (class labels) are defined along with the attributes. Although SVM is originally linear, usage of various kernel functions facilitates the utilization of SVM for non-linear applications as well. SVM is also known to have minimal generalization error and is constructed by identifying a set of planes used to differentiate multiple data classes. There exist various functions in SVM approach namely linear, quadratic and Radial Basis Functions (RBF) with the linear kernel being the simplest. The parameters of the RBF kernel are often varied based on the feature values. The linear kernel model is given in (4), the quadratic kernel in (5) and the RBF kernel is as shown in (6).

$$K(X_i, Y_i) = X_i^T Y_i + C \tag{4}$$

$$K(X_i, Y_i) = 1 - \frac{\|X_i - Y_i\|^2}{\|X_i - Y_i\|^2 + C} \tag{5}$$

$$d(X) = \sum_{N=1}^{svnum} \alpha y K(X_i, Y_i) + b \tag{6}$$

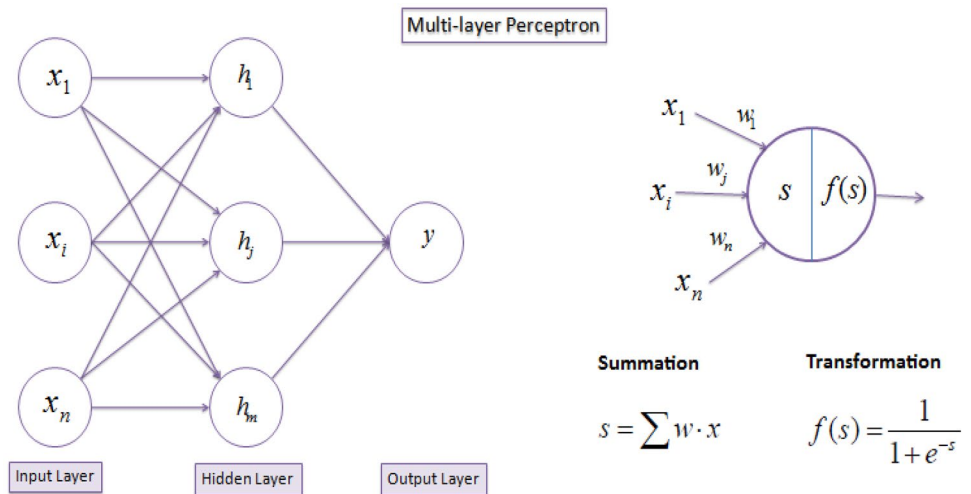
where $K(X_i, Y_i) = e^{-\frac{\|X_i - Y_i\|^2}{2\sigma^2}}$ and $X_i = i$ vector in the input (X is the input data vector); Y_i = the class to which the element belongs to; C = Capacity constant; α = hyperparameter; y = entire class; b = bias

The support vectors obtained using the SVM kernels are then fed into the SVM classifier. The prediction with respect to pre-defined features is then achieved using the SVM classifier [19].

3.5 MLP neural network

A Multilayer Perceptron (MLP) is a type of feed forward artificial neural network. It encompasses three layers of nodes namely input, output, and a hidden layer as shown in Fig. 1. Both hidden layer and the output nodes represent neurons using non-linear activation functions.

Fig. 1 Multilayer perceptron approach



MLP incorporates back-propagation for training. MLP can classify non-linear datasets as well. The backpropagation algorithm consists of two phases, namely, the forward phase (where the activations are propagated from the input to the output layer) and the backward phase (where the error between the observed actual and the requested nominal value in the output layer is propagated backwards in order to modify the weights and bias values). In forward propagation, the error propagates by adding all the weighted inputs and then computing outputs using sigmoid threshold, whereas in case of backward propagation, the error propagates backwards by allocating them to each unit according to the amount of this error the unit is responsible for. The mathematical approach for the functionality of MLP is as shown in (7)–(9)

$$d_o = y \times (1 - y) \times (t - y) \tag{7}$$

$$d_i = y_i \times (1 - y_i) \times (w_i \times d_o) \tag{8}$$

$$\Delta w = \eta \times d \times x \tag{9}$$

where d_o = error in the output neuron; d_i = error in the hidden neuron; x = input layer; y = Entire matrix of predicted value composed of the outputs of each neuron; t = target; w_i = updated weight value for each neuron when the backward propagation is performed; y_i = output for each neuron; η = hyperparameter; Δw = change in the weights

A generic workflow for this is shown in Fig. 1 [20].

The complete framework encompassing a step-by-step overview of the proposed methodology is presented in Fig. 2.

3.6 Subjects considered

396 subjects without any known auditory pathological history and of age group 30–50 years were chosen for the assessment with an informed consent, with 132 for each group CI, FI and BP. The class of subjects were defined as positively conditioned (BP) and negatively conditioned (CI and FI). The criteria for selection also included the condition that the subjects should have been working in chemical/fabrication industries or practicing basketball for a minimum of 3 years. The ATT was run twice a day once before their lunch break (LB) and then after the conclusion (CS) of their workshift. The nature of the job for the negatively conditioned subjects were such that there was a very minor exposure to noise and solvents before their lunch break. But this exposure significantly increased post LB after which the CS readings were planned to be obtained. The test was conducted in a silent noise free environment using a Dell Inspiron 15R laptop and the sound was provided using a boAt Rockerz 275 Bluetooth headset which was equipped with an adaptive noise cancelling feature as well. The generation and the presentation of sound was performed using MATLAB and the experiment was conducted with the aid of Psychoacoustics tool.

4 Results and discussions

The ATTh of the subjects were analyzed using basic statistical methods and then were subject to machine learning based data analytics for automated classification of subjects into their respective classes.

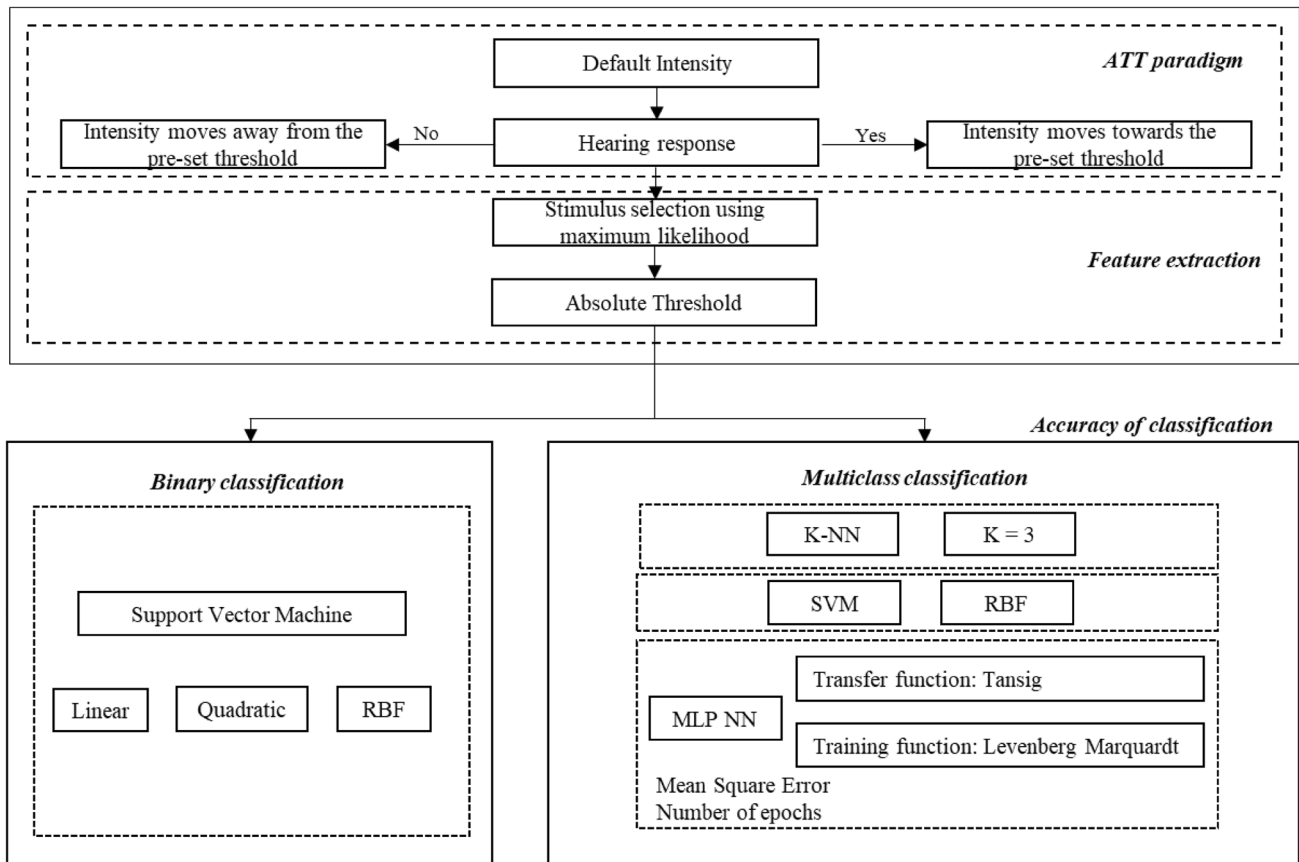


Fig. 2 Experimental framework

4.1 Statistical approach

When the ATTh was run for the first time (LB), the mean value was 39.52 dB, 58.26 dB and 49.78 dB for CI, FI and BP respectively. As hypothesized, the ATTh in BP was better than that of FI. This could have been attributed to the professional practice of basketball. However, the ATTh of CI was better than that of FI while the initial hypothesis was that FI was better than CI due to the fact that FI was exposed to noise alone whereas the CI was exposed to a combined effect of noise and solvents which should have resulted in a higher deterioration of ATTh. But after their workshift/conclusion (CS) of the practice sessions, the mean values were found to be 49.11 dB, 62.18 dB and 44.67 dB for CI, FI and BP. Upon observation of the change (CG) in the ATTh post workshift/practice sessions, a definite deterioration of the ATTh was found in case of CI and FI, with an improvement in BP. Also, the deterioration in CI was higher than FI due to a combined exposure to noise as well as solvents as shown in Table 1. Also, an average deviation of about $\pm 5\%$ from the mean value was seen in the data

Table 1 Mean ATTh for CI, FI and BP (dB)

	LB	CS
CI	39.52	49.11
FI	58.26	62.18
BP	49.78	44.67

acquired. This could be attributed to the fact that these were subjective recordings obtained, entirely based on the response of the subject and hence this variation.

A pictorial depiction of the variations in the mean values are provided in Fig. 3 for better representation as well as to compare the changes between each of the category of subjects for LB and CS.

Based on the results in Fig. 3, it is evident that ATTh could be useful to assess the positive and negative aspects of auditory conditioning in CI, FI and BP. Hence machine learning based approaches were incorporated to further probe into the variations in ATTh with respect to the considered subject categories and the results of the same are described in the succeeding section.

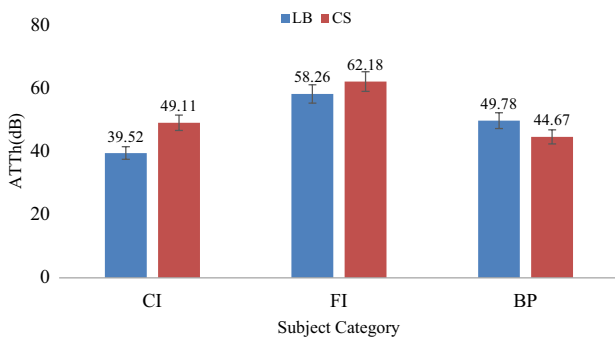


Fig. 3 Mean ATTh—a comparison

4.2 Classification based analysis

The ATTh obtained midway and at the conclusion phase were considered as features for the automatic classification of the subjects into CI, FI and BP respectively.

4.2.1 Binary classification

The SVM technique was used to achieve the binary classification of ATTh for CI, FI and BP. For each of the category, out of the 132 datasets, 88 were used for training and the rest 44 for testing. Linear, quadratic and RBF based methods were employed for classification for CI versus FI, FI versus BP and CI versus BP. The C value was fixed as 100 for RBF method. The confusion matrix for each of the technique (for testing dataset) is provided in Table 2 and the accuracy (in %) is provided in Table 3.

From Table 3, one can conclude that the RBF kernel-based approach performed the best, as compared to linear and quadratic approach in case of binary SVM. Another observation was that the linear and quadratic kernels performed better in classifying CI versus FI (85.22% and 86.36%). The linear kernel had a lower accuracy for CI versus BP (70.45%) and the quadratic seemed to perform low for FI versus BP (75%). Hence, one can infer the inferior performance of both linear as well as quadratic approaches in case of positively conditioned subjects (BP). But RBF portrayed the highest accuracy for every category of classification (CI vs. FI = 94%, FI vs. BP = 97.7% and CI vs. BP = 98.86%). Also, because both CI and BP lie at the opposite ends of the spectrum with respect to conditioning, the accuracy seemed to be the highest.

After a successful trial with the binary approach, a multiclass approach was developed with CI, FI and BP being the three classes of data, as provided in the succeeding section.

Table 2 SVM based classification—confusion matrix

<p>2a: CI vs FI - Linear</p> <table border="1"> <tr><td></td><td>CI</td><td>FI</td></tr> <tr><td>CI</td><td>38</td><td>6</td></tr> <tr><td>FI</td><td>7</td><td>37</td></tr> </table>		CI	FI	CI	38	6	FI	7	37	<p>2b: CI vs FI - Quadratic</p> <table border="1"> <tr><td></td><td>CI</td><td>FI</td></tr> <tr><td>CI</td><td>39</td><td>5</td></tr> <tr><td>FI</td><td>7</td><td>37</td></tr> </table>		CI	FI	CI	39	5	FI	7	37	<p>2c: CI vs FI - RBF</p> <table border="1"> <tr><td></td><td>CI</td><td>FI</td></tr> <tr><td>CI</td><td>42</td><td>2</td></tr> <tr><td>FI</td><td>3</td><td>41</td></tr> </table>		CI	FI	CI	42	2	FI	3	41
	CI	FI																											
CI	38	6																											
FI	7	37																											
	CI	FI																											
CI	39	5																											
FI	7	37																											
	CI	FI																											
CI	42	2																											
FI	3	41																											
<p>2d: FI vs BP - Linear</p> <table border="1"> <tr><td></td><td>FI</td><td>BP</td></tr> <tr><td>FI</td><td>35</td><td>9</td></tr> <tr><td>BP</td><td>14</td><td>30</td></tr> </table>		FI	BP	FI	35	9	BP	14	30	<p>2e: FI vs BP - Quadratic</p> <table border="1"> <tr><td></td><td>FI</td><td>BP</td></tr> <tr><td>FI</td><td>36</td><td>8</td></tr> <tr><td>BP</td><td>14</td><td>30</td></tr> </table>		FI	BP	FI	36	8	BP	14	30	<p>2f: FI vs BP - RBF</p> <table border="1"> <tr><td></td><td>FI</td><td>BP</td></tr> <tr><td>FI</td><td>42</td><td>2</td></tr> <tr><td>BP</td><td>0</td><td>44</td></tr> </table>		FI	BP	FI	42	2	BP	0	44
	FI	BP																											
FI	35	9																											
BP	14	30																											
	FI	BP																											
FI	36	8																											
BP	14	30																											
	FI	BP																											
FI	42	2																											
BP	0	44																											
<p>2g: CI vs BP - Linear</p> <table border="1"> <tr><td></td><td>CI</td><td>BP</td></tr> <tr><td>CI</td><td>32</td><td>12</td></tr> <tr><td>BP</td><td>14</td><td>30</td></tr> </table>		CI	BP	CI	32	12	BP	14	30	<p>2h: CI vs BP - Quadratic</p> <table border="1"> <tr><td></td><td>CI</td><td>BP</td></tr> <tr><td>CI</td><td>32</td><td>12</td></tr> <tr><td>BP</td><td>6</td><td>38</td></tr> </table>		CI	BP	CI	32	12	BP	6	38	<p>2i: CI vs BP - RBF</p> <table border="1"> <tr><td></td><td>CI</td><td>BP</td></tr> <tr><td>CI</td><td>43</td><td>1</td></tr> <tr><td>BP</td><td>0</td><td>44</td></tr> </table>		CI	BP	CI	43	1	BP	0	44
	CI	BP																											
CI	32	12																											
BP	14	30																											
	CI	BP																											
CI	32	12																											
BP	6	38																											
	CI	BP																											
CI	43	1																											
BP	0	44																											

4.2.2 3-Class classification

The scatter plot for the 3-class data (used for testing) is shown in Fig 4. A 3-class classification was achieved using two approaches namely kNN and SVM techniques with 88 datasets for training and 44 for testing (overall 132 datasets for each of the category).

4.2.2.1 kNN approach (k=3) The confusion matrix obtained using kNN is provided in Table 4, from which it is evident that an accuracy of 95.45% for CI, 86.36% for FI and 61.36% for BP were obtained. An overall accuracy of 81.06% was achieved in kNN approach.

4.2.2.2 Multiclass SVM approach The confusion matrix obtained using multi-class SVM is given in Table 5 which proved an efficiency of 93.18% for CI, 88.63% for FI and 100% for BP. Overall efficiency was observed to be 93.94% in this case.

4.2.2.3 MLP neural network approach In order to assess the possibility of improving the accuracy levels achieved

Table 3 Accuracy values obtained for SVM classification

	Linear (%)	Quadratic (%)	RBF (%)
CI versus FI	85.22	86.36	94.31
FI versus BP	73.80	75.00	97.70
CI versus BP	70.45	79.54	98.86

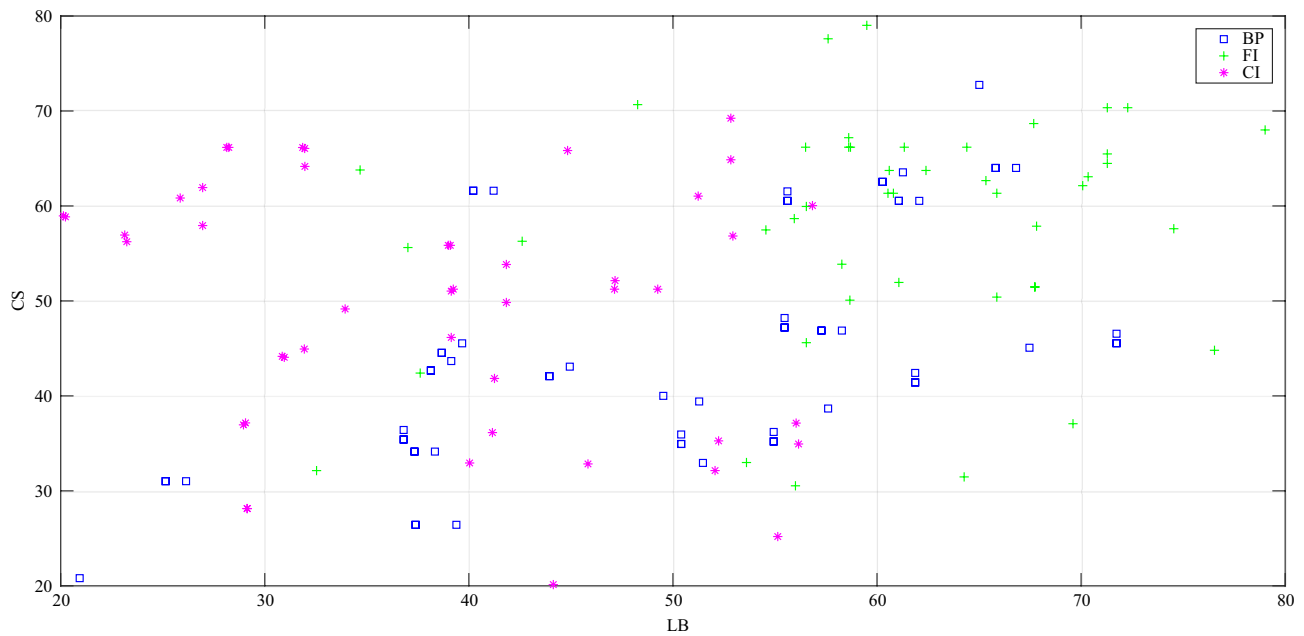


Fig. 4 Scatterplot for 3-class approach (testing data)

Table 4 Multiclass-kNN-confusion matrix

	CI	FI	BP
CI	42	0	2
FI	6	38	0
BP	10	7	27

Table 5 Multiclass-SVM-confusion matrix

	CI	FI	BP
CI	41	1	2
FI	1	39	4
BP	0	0	44

Table 6 Multiclass confusion matrix using MLP neural network

	CI	FI	BP
CI	39	1	4
FI	3	40	1
BP	0	1	43

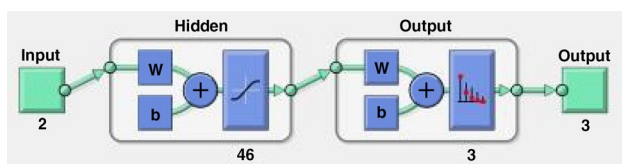


Fig. 5 Neural network architecture designed

using k-means and SVM techniques, an MLP based neural network approach was designed, as shown in Fig. 5 and the confusion matrix was obtained. However, this technique provided an accuracy of 88.63% for CI, 90.9% for FI and 97.7% for BP and an accuracy of 92.4% overall, as shown in Table 6. This was achieved with 46 hidden layers for 3-class outputs with the aid of Levenberg–Marquardt

training function in the hidden layer and Mean Square Error function in the output layer being optimal for the present case.

The error histogram was obtained with 20 bins which depicted a minimal error at 0.04, at which, the final performance was considered to be valid. An overall minimal mean square error was found to be observed at the 46th epoch with a best validation performance of 0.0453. This conclusion was supported by the corresponding ROC curves as well. The present MLP neural network demonstrated an efficiency of 88.64% for CI, 90.91% for FI and 97.73% for BP. Overall efficiency of MLP neural network for Multiclass was found to be 92.42%.

4.3 Inference

The present approach was used to successfully acquire the ATTh of CI, FI and BP classes using ATT paradigm. The results hinted at a definite variation in the classes due to constant exposure to noise, a combined exposure to noise and solvents as well as plausible improvement in auditory perception due to professional training of sports such as that of basketball. While binary classification was

achieved, a multi-class approach, too, was observed to be working fine, for the presently considered classes of subjects. RBF based approach was found to provide the highest accuracy for a binary class SVM based classification. In case of multi-class techniques, surprisingly, SVM based approach surpassed the kNN approach. This could be attributed to the non-linear nature of the datasets. Also SVM surpassed even MLP neural network technique. This could be due to the number of datasets. As the samples increases, neural network could better the results of the SVM in the near future (SVM = 93.94%, MLP-NN = 92.42% and kNN = 81.06%). Further, the response to sound variations with respect to psychological factors such as pitch and melody could be probed into, with refinement of the presently developed protocol. Also, variations could be induced in terms of the subjects to assess the robustness of the presently developed experimental paradigm to assess the psychophysics of auditory temporal resolution in normal healthy human beings. Such tests and paradigms can aid to educate the employees of industries towards a plausible deterioration of hearing perception and help in prevention of such exposures by implementing various techniques so as to avert constant exposure to noise in the workplace. This could also highlight the advantages of sports and regular training in physiological well-being of the individual, which could provide definite cues for trainers and coaches to further encourage individuals to participate in various sports activities and also in professional training and rehabilitation programs. Further, such psychophysical approaches could be embedded into regular sports training to assess whether such approaches could improve the performance of players on-field as well.

Acknowledgements The authors express their sincere gratitude to the Centre for Medical Electronics Research, MSRIT, and Anugraha Chemicals Pvt Ltd, Bangalore, for their support during data acquisition and to MediNXT Technologies, Bangalore for their help during data analytics. Psychoacoustics toolbox was used for data acquisition, with MATLAB 2018 version (MATLAB 2018a-9.4.0.813654) with an academic user license 1113042 used for data analytics. Also, the authors hereby clarify that there was no financial interest associated with this work and also wish to hereby declare that there was no conflict of interest of any kind regarding the publication of this work.

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interests.

Ethical approval This is to declare that the present work has been screened by the ethical committee of MedNXT Innovative Technologies. The proposed work was presented to this committee, chaired by Sri Syed Faisal Ali (Project Manager, MediNXT Innovative Technologies), and the constituent members were Dr. Harshith N (Bangalore Medical College), Dr Bhargavi Sunil (SJC Institute of Technology) and Dr Vijayalakshmi K (BMS College of Engineering). Due approvals have been obtained by this committee prior to data acquisition.

References

- Vogelmeier CF et al (2017) Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017. report GOLD executive summary. *Am J Respir Crit Care Med* 195(5):557–582
- Liberman MC (2017) Noise-induced and age-related hearing loss: new perspectives and potential therapies. *F1000Res* 6:927
- Brown AD et al (2015) Effects of active and passive hearing protection devices on sound source localization, speech recognition, and tone detection. *PLoS ONE* 10(8):e0136568
- Jyothi S et al (2016) Correlation of audio-visual reaction time with body mass index & skin fold thickness between runners and healthy controls. *Indian J Physiol Pharmacol* 60(3):239–246
- Mioni G et al (2016) The impact of a concurrent motor task on auditory and visual temporal discrimination tasks. *Atten Percept Psychophys* 78(3):742–748
- Maniglia M, Grassi M, Ward J (2017) Sounds are perceived as louder when accompanied by visual movement. *Multisensory Research* 30(2):159–177
- Hwang S, Lee KJ, Park JB (2018) Pulmonary function impairment from exposure to mixed organic solvents in male shipyard painters. *J Occup Environ Med* 60(12):1057–1062
- Nakhooda F, Sartorius B, Govender SM (2019) The effects of combined exposure of solvents and noise on auditory function—a systematic review and meta-analysis. *S Afr J Commun Disord* 66(1):568
- Crundwell G, Gomersall P, Baguley DM (2016) Ototoxicity (cochleotoxicity) classifications: a review. *Int J Audiol* 55(2):65–74
- Park SH et al (2015) Differential contribution of visual and auditory information to accurately predict the direction and rotational motion of a visual stimulus. *Appl Physiol Nutr Metabol* 41(3):244–248
- Baraldi A et al. (2018) Earth observation big data analytics in operating mode for GIScience applications? The (GE) OBIA acronym (s) reconsidered. In: *GEOBIA 2018—from pixels to ecosystems and global sustainability*
- Noto M, Nishikawa J, Tateno T (2016) An analysis of nonlinear dynamics underlying neural activity related to auditory induction in the rat auditory cortex. *Neuroscience* 318:58–83
- Plack CJ et al (2016) Toward a diagnostic test for hidden hearing loss. *Trends in hearing* 20:2331216516657466
- Sanjay HS, Bhargavi S, Madhuri S (2018) Auditory psychophysical analysis of healthy individuals based on Audiometry and Absolute threshold tests. *Int Journal Eng Technol* 7(9):99–106
- Manning C et al (2018) Psychophysics with children: investigating the effects of attentional lapses on threshold estimates. *Atten Percept Psychophys* 80(5):1311–1324
- Shen Y, Dai W, Richards VM (2015) A MATLAB toolbox for the efficient estimation of the psychometric function using the updated maximum-likelihood adaptive procedure. *Behav Res Methods* 47(1):13–26
- Soranzo A, Grassi M (2017) Corrigendum: pSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing. *Front Psychol* 8:1185
- Maxwell A et al (2016) Testing auditory sensitivity in the great cormorant (*Phalacrocorax carbo sinensis*): psychophysics versus auditory brainstem response. In: *Proceedings of meetings on acoustics 4ENAL*. vol 27. no 1. ASA
- Levine SM, Schwarzbach J (2017) Decoding of auditory and tactile perceptual decisions in parietal cortex. *NeuroImage* 162:297–305

20. Kell AJE et al (2018) A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron* 98(3):630–644

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.