

RESEARCH

Open Access



Optimizing age-related hearing risk predictions: an advanced machine learning integration with HHIE-S

Tzong-Hann Yang^{1,2,3,4}, Yu-Fu Chen³, Yen-Fu Cheng^{5,6,7,8}, Jue-Ni Huang¹⁰, Chuan-Song Wu^{1,9*} and Yuan-Chia Chu^{10,11,12*}

*Correspondence:
seancswu@seed.net.tw;
xd.yuanchia@vghtpe.gov.tw

⁹ College of Science and Engineering, Fu Jen University, Taipei 243, Taiwan
¹² Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei 112, Taiwan
Full list of author information is available at the end of the article

Abstract

Objectives: The elderly are disproportionately affected by age-related hearing loss (ARHL). Despite being a well-known tool for ARHL evaluation, the Hearing Handicap Inventory for the Elderly Screening version (HHIE-S) has only traditionally been used for direct screening using self-reported outcomes. This work uses a novel integration of machine learning approaches to improve the predicted accuracy of the HHIE-S tool for ARHL in older adults.

Methods: We employed a dataset that was gathered between 2016 and 2018 and included 1,526 senior citizens from several Taipei City Hospital branches. 80% of the data were used for training ($n = 1220$) and 20% were used for testing ($n = 356$). XGBoost, Gradient Boosting, and LightGBM were among the machine learning models that were only used and assessed on the training set. In order to prevent data leakage and overfitting, the Light Gradient Boosting Machine (LGBM) model—which had the greatest AUC of 0.83 (95% CI 0.81–0.85)—was then only used on the holdout testing data.

Results: On the testing set, the LGBM model showed a strong AUC of 0.82 (95% CI 0.79–0.86), far outperforming conventional techniques. Notably, several HHIE-S items and age were found to be significant characteristics. In contrast to traditional HHIE research, which concentrates on the psychological effects of hearing loss, this study combines cutting-edge machine learning techniques—specifically, the LGBM classifier—with the HHIE-S tool. The incorporation of SHAP values enhances the interpretability of the model's predictions and provides a more comprehensive comprehension of the significance of various aspects.

Conclusions: Our methodology highlights the great potential that arises from combining machine learning with validated hearing evaluation instruments such as the HHIE-S. Healthcare practitioners can anticipate ARHL more accurately thanks to this integration, which makes it easier to intervene quickly and precisely.

Keywords: Age-related, Hearing loss, LGBM, Machine learning, HHIE-S, Predictive enhancement, Innovation



Introduction

Age-related hearing loss (ARHL) is a common condition that occurs naturally as individuals age and is characterized by progressive bilateral high-frequency sensorineural hearing loss. As life expectancy increases globally, ARHL has become a significant health concern, particularly for individuals aged 60 and above in both the United States and Taiwan [1, 2]. This age-related condition can lead to various adverse consequences, including difficulties with balance and movement [3–5], social isolation [6], cognitive impairment [6–9], and even an increased risk of mortality [10]. It is crucial to identify and treat ARHL in a timely manner to promote healthy aging and mitigate these negative impacts.

While pure tone audiometry remains the gold standard for detecting hearing loss, its feasibility in large-scale population-based settings may be limited. As a result, there is growing interest in alternative screening techniques that are more accessible and efficient. Recent research has explored various options, such as whisper voice testing [11], telephone-based assessments [12], computer-based evaluations [13, 14], and internet or smartphone applications [15–17], to evaluate hearing sensitivity in older adults. However, questionnaires such as the Traditional Chinese version of the Hearing Handicap Inventory for the Elderly-Screening (TC-HHIE-S) [18] offer a cost-effective and convenient approach to assessing hearing loss, particularly in clinical settings without proper soundproofing. Our previous research has demonstrated that this 10-question questionnaire is a reliable tool for detecting hearing loss and can serve as a suitable alternative to audiometry in large-scale hearing screening initiatives [18].

Declining attention span in older individuals can impact their ability to complete lengthy questionnaires, making it essential to develop a shorter and more efficient screening tool for hearing loss [19]. Additionally, understanding how individual demographics, such as age and gender, influence the accuracy of hearing loss estimation is crucial. Machine learning techniques offer promising solutions to address these challenges. By leveraging data-driven learning without relying on rule-based programming, machine learning algorithms can optimize prediction models using demographic variables and HHIE-S data to achieve accurate hearing loss estimation [20, 21].

The primary objective of this study is to develop a concise version of the HHIE-S questionnaire and evaluate its effectiveness in a Taipei community using a machine learning algorithm [18]. Integrating machine learning into hearing screening protocols can lead to more personalized and effective assessments, improving the overall health outcomes for older individuals with ARHL. It may also help identify high-risk individuals who could benefit from early interventions and support, ultimately contributing to improved quality of life in the aging population.

Methods

Data sources and study population

A government-funded annual geriatric health check-up program was available to Taipei residents 65 years of age and older from January 2016 to December 2018. Notified by the city government, eligible participants usually had examinations at neighborhood community hospitals. Participants from Taipei City Hospital's Heping branch, which serves

the Zhongzheng and Wanhua districts of Taipei, were the study's primary focus. A total of 1,526 adults (706 men and 820 women, $p > 0.05$) were included in the study.

Feature selection

At Taipei City Hospital, information was gathered for the training cohort. The participants were given the Hearing Handicap Inventory for the Elderly—Screening (TC-HHIE-S) questionnaire in Traditional Chinese face-to-face, and an audiologist or an undergraduate student studying audiology under supervision recorded the results in their medical records. There were twelve features in the dataset that were analyzed, including demographic variables like gender and age. The TC-HHIE-S questionnaire, comprising ten questions, was also included, offering a thorough understanding of the participants' hearing health.

Pure-tone audiometry

As the gold standard for detecting hearing loss, pure-tone audiometry was carefully carried out in a sound-treated booth in our study, with ambient noise levels strictly kept below 30 dBA. To perform the audiometry tests, we used the MA30 Audiometer (Maico, Germany) in conjunction with the TDH-39 supra-aural earphones. Four key frequencies—0.5 kHz, 1 kHz, 2 kHz, and 4 kHz—were carefully measured to determine the air conduction pure-tone thresholds. These measurements were primarily taken in the better-hearing ear, with the right ear being used as a default when there was no discernible hearing difference between the two ears. These thresholds were then averaged in the better-hearing ear to determine the pure-tone average (PTA). Narrow band masking was used as needed to guarantee the precision of our measurements, and participants usually indicated their hearing responses with a standard patient response button. An essential component of our approach involved our equipment's yearly calibration, which was conducted with strict adherence to ISO 389–1 and 389–3 guidelines. This step was essential to maintaining the integrity and robustness of our study's findings by ensuring the accuracy and consistency of our audiometric data.

Class definition

To account for varying degrees of hearing loss, we used a straightforward and methodical method to divide the Pure-Tone Average (PTA) results of study participants into two groups. The established hearing threshold cutoff points served as the basis for this classification. More specifically, individuals with a PTA of less than 40 dB HL—a sign of mild to moderate hearing loss—were given a classification value of 1. In contrast, individuals whose PTA matched or exceeded 40 dB HL were assigned a classification value of 0. The latter category represents hearing loss ranging from severe to profound. We were able to classify the study participants' hearing impairment severity with this binary classification system, which gave us a solid foundation for our machine learning analysis later on.

Machine learning model development

For model development, we utilized the graphical tool Auto AI in Watson Studio to analyze the 80% training data ($n = 1220$) and identify the best data transformations, algorithms, and parameter settings for our predictive modeling task [22–24]. The Auto AI

tool presented the results as candidate model pipelines, which were ranked on a leaderboard, enabling the selection of the optimal model [25–27]. We assessed the performance of six diverse machine learning models on the training data, including extreme gradient boosting (XGBoost) [28–30], gradient boosting classifier (GBC) [31–33], snap decision tree classifier (SDTC) [23, 34, 35], light gradient boosting machine (LGBM) [36–38], snap random forest classifier (SRFC) [26, 27, 39], and logistic regression (LR) [40–42] on the training set, and was thus selected for final evaluation on the 20% held-out testing data ($n = 306$). To improve model efficiency, we implemented forward feature selection on the training data to identify the most informative subset of features. By developing and evaluating models solely on the training set, we prevented leakage and overfitting to the test data.

Hyperparameter optimization

To optimize the performance of the machine learning models in predicting ARHL, we employed a grid search combined with fivefold cross-validation for XGBoost, GBC, SDTC, LGBM, SRFC, and LR [43–46]. During the training process, we fine-tuned the hyperparameters for each ensemble model using a grid search to identify the optimal values that yielded the highest F1 score [46, 47]. Table 1 presents the detailed hyperparameter optimization settings for each model. By iteratively exploring a predefined set of hyperparameter values through the grid search, we were able to enhance the accuracy and predictive capability of the models in assessing the risk of ARHL. This iterative approach allowed us to find the best combination of hyperparameter values for each model, ultimately leading to improved performance and more reliable predictions.

Model evaluation

We compared the models’ Area Under the Receiver Operating Characteristic Curve (AUC) values to assess the machine learning models’ discriminative performance. Because it indicates how well a model can distinguish between classes, the AUC metric is important. We used a suite of performance metrics to conduct a thorough assessment of the models’ efficacy on the testing dataset, in addition to AUC. These included the F1 score, which strikes a balance between recall and precision; accuracy, which gauges the model’s overall correctness; precision, which shows the percentage of true positive identifications that were correctly identified; recall, which gauges the percentage of true

Table 1 Hyperparameters of machine learning models

Model	Hyperparameters	Optimal values
Extreme Gradient Boosting (XGBoost)	learning_rate	0.0384
	max_depth	7
	n_estimators	494
Gradient Boosting Classifier (GBC)	min_samples_leaf	5
	n_estimators	33
Stochastic Dual Coordinate Ascent (SDTC)	max_depth	5
	random_state	33
Light Gradient Boosting Machine (LGBM)	learning_rate	0.415
	n_estimators	566

positives that were correctly identified; average precision, which provides a summary of a precision-recall curve; and log loss, which assesses the prediction error of the models. By utilizing this diverse range of metrics, we were able to perform a comprehensive evaluation of the models' efficacy, guaranteeing a detailed and refined comprehension of their potential to forecast age-related hearing loss (ARHL).

Feature importance analysis

We used SHapley Additive exPlanations (SHAP) analysis to clarify the underlying mechanisms guiding the predictions of our model and to provide guidance for bettering hearing screening procedures. A state-of-the-art approach in explainable AI called SHAP offers a detailed perspective of how each feature influences the model's predictions. Our main objective was to analyze the importance of different elements in the HHIE-S survey. We were able to identify the most important questionnaire items and comprehend their individual effects on the risk prediction of age-related hearing loss (ARHL) by utilizing SHAP values. Through the identification of important ARHL risk factors in our dataset, this analysis was able to offer valuable insights for the development of both our predictive model and the field of hearing health assessment as a whole. It enabled us to see the direction and strength of each feature's influence on the model's output in addition to quantifying each feature's importance.

Impact on hearing care management

The conclusions drawn from our SHAP analysis have a major impact on how focused interventions for those at risk of hearing loss are developed. Our machine learning model identifies the major factors that influence the risk of age-related hearing loss (ARHL), and this approach opens the door to more customized and efficient hearing care management plans. Comprehending these crucial elements enables medical practitioners to devise interventions that are tailored to each individual's specific requirements and risk profile, while also being responsive to the broad patterns noted in ARHL. This tailored approach is anticipated to substantially improve the effectiveness of interventions, leading to enhanced outcomes in terms of both health and quality of life for older adults afflicted with ARHL. The application of such data-driven, personalized health-care strategies marks a significant advancement in the field of audiology and geriatric care, promising a more nuanced and impactful approach to managing and mitigating the effects of hearing loss.

Software and model development

The open-source Scikit-learn library was heavily utilized in conjunction with Python, more especially Python Software Foundation version 3.9, as the main platform for all machine learning analyses in our study. A wide variety of tools and algorithms that were essential for our analysis were provided by this library. To effectively test our models on unseen data, we divided the dataset into training and testing sets at random using the `sklearn.model_selection.train_test_split` module. We used the XGBoost Python package, which is well-known for its effectiveness and high performance in gradient boosting, to build the XGBoost model. The ensemble of sklearn. Our Gradient Boosting Decision Tree (GBDT) model was developed with the help of the `GradientBoostingClassifier`

module, which constructed an ensemble of weak learners to enable precise prediction-making. The `sklearn.ensemble` was also used by us. The Extra Trees model, an ensemble approach based on decision trees that provides robustness and feature importance estimation, was built using the `ExtraTreesClassifier`. The `lightgbm.LGBMClassifier` Python package, which is preferred for its effectiveness in large-scale applications, was used to develop the Light Gradient Boosting Machine (LGBM) model. A decision tree classifier and `sklearn.linear_model` were supplied by the `SnapDecisionTreeClassifier` algorithm from the IBM Snap ML library. The logistic regression model, which is frequently applied to binary classification issues, was utilized. Use `sklearn.model_selection` to avoid overfitting and guarantee robustness. For stratified k-fold cross-validation, Stratified K Fold was utilized. To ensure compliance with accepted scientific practices, we maintained a p-value of 0.05 to denote statistical significance throughout our analysis.

Role of the funding source

The study received funding from a specific organization, with the crucial stipulation that the funders remained entirely uninvolved in any aspects of the research process. This encompassed study design, data collection, analysis, interpretation, writing, and paper submission. The clear separation between funding and research activities ensured that the research was conducted independently, without the potential for bias or influence from the funding source. All authors involved in the study had full access to the data, fostering transparency and unbiased research. The corresponding authors assumed final responsibility for paper submission, ensuring that the research adhered to scientific standards and ethical guidelines. They were accountable for upholding the accuracy and integrity of the work presented in the paper.

Results

Patient population overview

In this work, we conducted a thorough analysis of a cohort of 1,526 older adults, which is representative of the larger population affected by age-related hearing loss (ARHL). To ensure a thorough evaluation of our machine learning models, this cohort was carefully split into two subsets: a primary training set of 1,221 participants, which was used to develop and refine the models, and a validation testing set of 305 participants, which was used to independently assess the models' performance and predictive accuracy.

Demographic distribution

The study focused on an elderly population, which is most affected by ARHL, as evidenced by the average age of 74.6 years (± 2.2 years) across the entire cohort in terms of demographic characteristics. Upon closer inspection, we found that the testing set (72.7 years ± 0.6) and the training set (75.2 years ± 2.4) had slightly different average ages. These differences were considered throughout the analysis to guarantee the validity of our results. The cohort's gender distribution was reasonably uniform, with women making up roughly 53.7% of all participants. Table 2 shows that this percentage differed slightly between the training set (53%) and testing set (56.7%). These demographic details are essential because they guarantee that our models are tested on a representative sample of the target population and give context for the findings of our study.

Table 2 Demographics and clinical features of Age-Related Hearing Loss (ARHL) patients

Variables	Full Cohort (n = 1,526)	Training set (n = 1221)	Testing set (n = 305)
Demographic			
Age, years	74.1 ± 7.2	74.5 ± 7.4	72.7 ± 6
Female, n (%)	820 (53.7%)	647 (53%)	173 (56.7%)
HHIE-S Item Score			
HHIE-1	0.6 ± 1.2	0.6 ± 1.2	0.2 ± 0.7
HHIE-2	0.5 ± 1.1	0.6 ± 1.2	0.3 ± 0.8
HHIE-3	1.2 ± 1.5	1.3 ± 1.6	0.8 ± 1.3
HHIE-4	0.6 ± 1.3	0.7 ± 1.3	0.3 ± 0.9
HHIE-5	0.4 ± 1	0.5 ± 1.1	0.2 ± 0.6
HHIE-6	0.3 ± 0.9	0.3 ± 1	0.1 ± 0.5
HHIE-7	0.3 ± 0.8	0.3 ± 0.8	0.2 ± 0.6
HHIE-8	0.8 ± 1.4	0.9 ± 1.4	0.5 ± 1
HHIE-9	0.3 ± 0.9	0.4 ± 1	0.1 ± 0.5
HHIE-10	0.6 ± 1.2	0.6 ± 1.2	0.3 ± 0.9

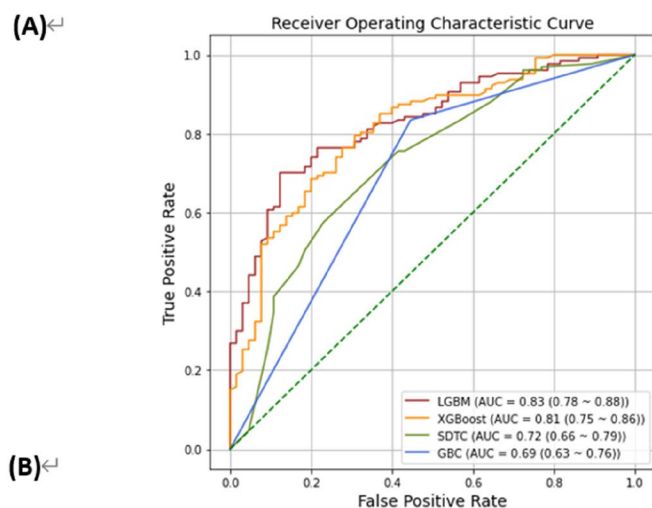
Abbreviations: HHIE-1 "Do hearing problems embarrass you when you meet new people?", HHIE-2 "Does a hearing problem make you feel frustrated when you talk to members of your family?", HHIE-3 "Do you find it difficult hearing when someone speaks in a whisper?", HHIE-4 "Do you feel that you have a disability because of a hearing problem?", HHIE-5 "Does a hearing problem cause you difficulty when visiting friends and relatives or neighbors?", HHIE-6 "Does a hearing problem reduce your attendance at religious ceremonies more than you would like?", HHIE-7 "Does a hearing problem cause disputes between you and a family member?", HHIE-8: "Does a hearing problem cause difficulty when you are listening to the radio or television?" HHIE-9 "Do you feel that any difficulty in hearing limits or hinders your personal or social life?", HHIE-10 "Does a hearing problem cause you difficulty when you are in a restaurant with relatives or friends?"

HHIE-S score insights

A nuanced comparison of HHIE-S scores across the full cohort, training, and testing sets showed subtle differences. For instance, the score for the HHIE-1 item was marginally higher in the full cohort (0.6 ± 1.2) than in the testing set (0.2 ± 0.7), while the training set maintained a score of 0.6 ± 1.2 . Similar patterns were observed for the other HHIE-S items, underscoring the consistent representation of hearing impairment symptoms across subsets.

Model prediction ability

Different performance characteristics were observed in our assessment of the predictive capabilities of different machine learning models, such as XGBoost, Gradient Boosting Classifier (GBC), Snap Decision Tree Classifier (SDTC), Light Gradient Boosting Machine (LGBM), and Snap Random Forest Classifier (SRFC). Remarkably, as shown in Fig. 1A, XGBoost and LGBM had the highest accuracy at 0.75, closely followed by GBC at 0.76. In the context of hearing loss predictions, LGBM proved to be a superior model in terms of specificity, achieving an astounding rate of 0.88 in correctly identifying true negative cases. With a score of 0.87, XGBoost showed remarkable sensitivity, demonstrating its ability to accurately identify true positive cases. Precision, recall, and F1 scores were fairly consistent among the models; LGBM and GBC achieved particularly noteworthy scores in these metrics. Furthermore, with the highest positive predictive value (PPV) of 0.92, LGBM stood out. LGBM exhibited the highest Area Under the Curve (AUC) value of 0.83, as illustrated in Fig. 1B, indicating its superior



Model	Accuracy	Specificity	Sensitivity	Precision	Recall	F1	PPV	AUC
Extreme Gradient Boosting (XGBoost)	0.75	0.63	0.87	0.79	0.86	0.82	0.82	0.81
Gradient Boosting Classifier (GBC)	0.76	0.63	0.88	0.8	0.85	0.83	0.82	0.69
SnapDecisionTreeClassifier (SDTC)	0.73	0.58	0.85	0.80	0.79	0.79	0.8	0.72
Light Gradient Boosting Machine (LGBM)	0.75	0.88	0.77	0.87	0.73	0.79	0.92	0.83 [†]

Fig. 1 Evaluation of Machine Learning Models for Age-Related Hearing Loss Prediction on the Testing Dataset. **A** Receiver operating characteristic (ROC) curves and **B** LGBM demonstrate the highest area under the ROC curve for age-related hearing loss (ARHL) prediction on the testing dataset, outperforming XGBoost and SDTC. Abbreviations: ROC, receiver operating characteristic; AUC, area under the ROC curve; precision, average precision; extreme gradient boosting (XGBoost), gradient boosting classifier (GBC), snap decision tree classifier (SDTC), light gradient boosting machine (LGBM), snap random forest classifier (SRFC)

overall discriminating ability. This thorough evaluation provides important insights for the models’ use in this field by highlighting the various strengths of each model in predicting age-related hearing loss.

Feature importance and SHAP analysis in the LGBM model

In this work, we employed SHapley Additive exPlanations (SHAP) values to determine the significance of each feature in the Light Gradient Boosting Machine (LGBM) model. As seen in Fig. 2A, this required creating feature importance plots that ordered the most significant features in descending order. AGE, HHIE-4, HHIE-1, HHIE-3, HHIE-8, and HHIE-10 were the most important features influencing the model’s predictions, according to our analysis. Higher SHAP values indicated a stronger influence. The SHAP summary plot provided a visual representation of these features’ effects on the prediction probabilities of the LGBM model. Significantly, developing Age-Related Hearing Loss (ARHL) was positively correlated with positive SHAP values (red dots), indicating a higher likelihood, and negatively correlated with negative SHAP values (blue dots). As shown in Fig. 2B, our results showed that while HHIE-10 had a negative effect, characteristics like AGE, HHIE-3, HHIE-1, HHIE-8, and HHIE-4 had a positive influence on the prediction of ARHL. This analysis was essential in emphasizing how each of these

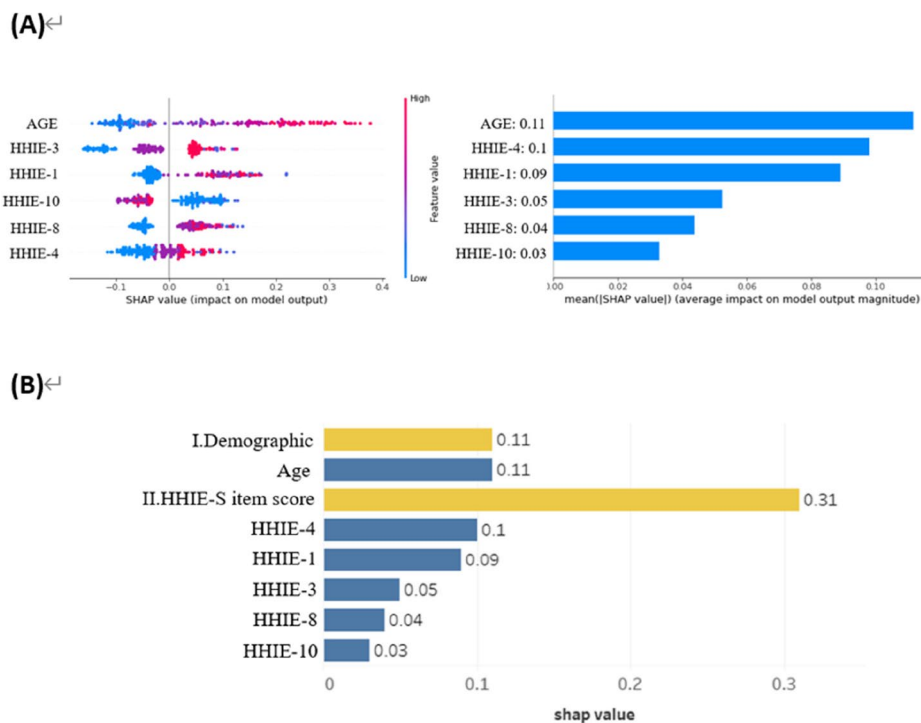


Fig. 2 **A** Feature importance plot and **B** SHAP summary plot revealing the top clinical features for predicting age-related hearing loss (ARHL) in the LGBM model. Abbreviations: HHIE-4: "Do you feel that you have a disability because of a hearing problem? ", Age, HHIE-1: "Do hearing problems embarrass you when you meet new people? ", HHIE-3: "Do you find it difficult to hear someone who speaks in a whisper? ", HHIE-8: "Does a hearing problem cause difficulty when you are listening to the radio or television? ", HHIE-10: "Does a hearing problem cause you difficulty when you are in a restaurant with relatives or friends?"

features individually contributed to the predictive accuracy of the model and in giving users a clear grasp of the critical elements involved in determining an individual's risk of ARHL.

Individual-level explanation of the ML model

For patients at risk of Age-Related Hearing Loss (ARHL), we used SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) to provide comprehensive individual-level explanations of our machine learning (ML) model's predictions. The aforementioned interpretive tools, as illustrated in Fig. 3, efficiently depict the ARHL probability for a pair of representative patients, clarifying the impact of every variable on the probability of hearing impairment. In Case 1, for instance, the LIME plot shows a very low ARHL probability of 0.07, with important contributing factors represented by the red (positive correlation) and blue (negative correlation) variables. These factors include age between 77 and 83, an HHIE-10 score between 0 and 2, and an HHIE-3 score of 1 or less. In contrast, Case 2 shows a 75% chance of ARHL. This higher risk prediction is heavily influenced by similar age parameters combined with scores from HHIE-3, HHIE-8, HHIE-1, and HHIE-10 in particular ranges. These individual-level analyses are essential because they provide a clear and understandable representation of the predictive factors in the model, allowing medical

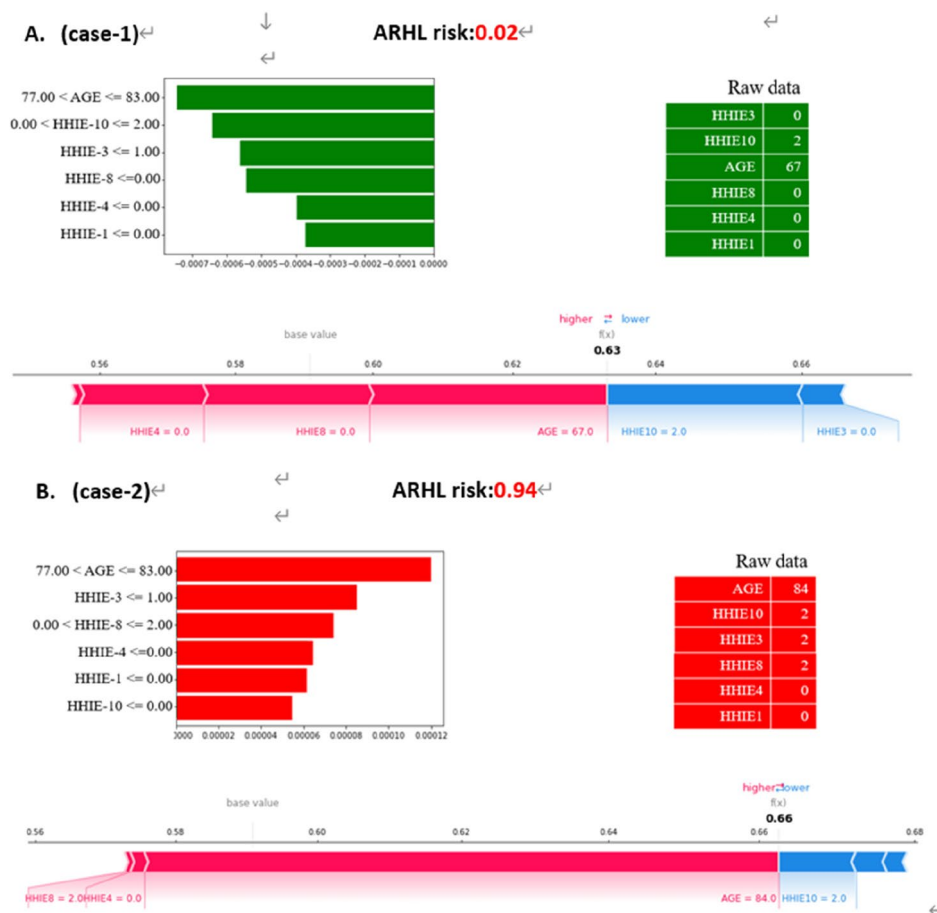


Fig. 3 LIME and SHAP Force Plots Visualizing the Impact of Important Features on the Prediction Model for Individual Patients with Age-Related Hearing Loss (ARHL). The plots use color coding, with red indicating a positive correlation and green indicating a negative correlation between the features and the predicted probability of age-related hearing loss (ARHL) risk assessment. Abbreviations: HHIE-4: "Do you have a disability because of a hearing problem? ", Age, HHIE-1: "Do hearing problems embarrass you when you meet new people? ", HHIE-3: "Do you find it difficult to hear someone who speaks in a whisper? ", HHIE-8: "Does a hearing problem cause difficulty when you are listening to the radio or television? ", HHIE-10: "Does a hearing problem cause you difficulty when you are in a restaurant with relatives or friends?"

practitioners to obtain in-depth understanding of the model's evaluation of patients' ARHL risk.

Discussion

In this study, we utilized machine learning techniques to validate the efficacy of the HHIE-S tool in assessing the risk of ARHL [18, 48, 49]. Our results highlight the potential of machine learning models, especially LGBM, to enhance the predictive capabilities of the HHIE-S tool by incorporating demographic and clinical features [50–52]. Our research is in line with the increasing interest in leveraging machine learning in audiology and hearing healthcare, as evident from recent publications [53–55]. Numerous studies have been dedicated to the development and validation of machine learning models for predicting hearing loss, evaluating interventions, and improving diagnostic accuracy [56–58].

In our study, the LGBM model demonstrated the best specificity (0.88), showcasing its proficiency in accurately identifying true negative cases. Additionally, it exhibited the highest PPV of 0.92, indicating its capability to correctly predict positive cases with high confidence. The AUC analysis revealed that the LGBM model achieved the highest AUC of 0.83, signifying its superior overall discriminative ability. These results align with our previous findings, where PTA > 40 dB HL showed the highest sensitivity (76.9%) in hearing screening, indicating similar discriminatory power [18].

The HHIE-S tool is a cost-effective and easily applicable assessment, making it suitable for diverse settings such as nursing homes, senior centers, primary care facilities, and community hearing screenings with numerous participants within a short timeframe. These locations often lack access to conventional hearing tests. This research emphasizes the significance of the HHIE-S in public health applications, as it can identify moderate and severe hearing loss in a considerable population. To further enhance the practicability and effectiveness of large-scale HHIE-S screenings, future studies should investigate the utilization of internet-based or smartphone-based methods for managing HHIE-S [59, 60].

By employing LGBM models in our study, we observed enhanced predictive performance compared to that of traditional statistical methods. Notably, the LGBM achieved the highest AUC values, indicating its superior ability to predict ARHL risk. This result aligns with recent findings of the effectiveness of LGBM and other ensemble models in various medical applications. Conducting a feature importance analysis using SHAP values provided valuable insights into the most influential features [61]. Age and specific HHIE-S questionnaire items, particularly HHIE-4, emerged as the top predictors [18, 48, 62]. These results are consistent with previous findings that have emphasized the role of age and self-perceived hearing difficulties in predicting hearing loss.

Six features on the importance matrix plot and the SHAP summary plot of LGBM were age and the five HHIE-S results. The most important factor for the model output was "Do you feel that you have a disability because of a hearing problem?": Our model identified the 4th item of the HHIE-S questionnaire as the most influential in determining hearing loss. This question directly addresses the issue of hearing loss in a clear and understandable way. It measures emotional responses to hearing loss and is similar to the commonly used single question (SQ): "Do you have any difficulty with your hearing?", which has shown reasonable effectiveness in detecting hearing loss during large-scale screenings [63–65].

Age

Age is an important variable in detecting hearing loss in our model. Our study aligns with several studies that demonstrated age as a significant variable in questionnaires for detecting hearing loss [66]. The significance may be due, in part, to the fact that the prevalence of ARHL increases with age [67–69]. Although some studies also indicate that older age could lead to overestimation or underestimation in self-reported hearing loss compared to objective hearing loss, it may be attributed to the social acceptability of older individuals experiencing hearing impairment [70–73]. Nonetheless, including age as a variable in a hearing screening session is convenient because age data are easily accessible in any massive screening situation [68, 72, 74, 75].

Do you find it difficult to hear someone who speaks in a whisper?

Our analysis shows that the whisper test is the second most significant factor in measuring hearing loss. This is an interesting finding, as the whisper test is a simple and accurate screening method widely used in hearing screenings for elderly individuals, with reasonable sensitivity and specificity [64, 76–78]. In fact, the World Health Organization recommends using the whisper test in hearing screenings as part of the Integrated Care for Older People (ICOPE) program [79]. During a whisper test, air is pushed through a small opening between the vocal cords without adduction, resulting in a quieter, breathier sound that may be less clear and distinct, making it more difficult to understand the speaker, especially in noisy environments or at a distance. Whispering is typically used to convey secret information or avoid disturbing others in a quiet setting, and failure to hear whispered sounds can cause individuals to perceive hearing loss in these situations.

Does a hearing problem cause difficulty when you are listening to the radio or television?

TV and radio are currently very common forms of media. However, a study has suggested that patients who report increased volume levels while watching TV may have an increased risk of hearing loss [80]. It is possible that individuals who have difficulty hearing the sounds from these devices may assume they have hearing loss when, in fact, they may simply be unable to detect certain frequencies or volumes of sound. Further research is needed to clarify the relationship between increased TV volume and hearing loss.

Does a hearing problem cause you difficulty when you are in a restaurant with relatives or friends?

Restaurants are bustling environments filled with sounds such as background music, conversations, and the clanging of utensils, which can make it challenging for people with hearing loss to discern between sounds or comprehend speech. In fact, speech understanding difficulties in noisy environments are frequently cited as the top reason for consultations in audiology or hearing health services [81]. It is important to conduct further research to better understand this association.

Our study adds to the expanding body of evidence supporting the integration of machine learning models into hearing screening protocols. By accurately identifying individuals at risk of ARHL, healthcare professionals can implement timely interventions, such as hearing aids or rehabilitation programs, to mitigate the impact of hearing loss on quality of life. The use of machine learning models enhances the efficiency and accuracy of screening processes, leading to improved patient outcomes. However, it is essential to acknowledge several limitations in our study. Focusing on a specific population of older adults limits the generalizability of our findings to other age groups or populations [62]. Future research should aim to validate the HHIE-S tool and machine learning models in diverse populations to ensure their applicability across settings. Additionally, the quality and availability of data may influence model performance. Future studies should consider larger sample sizes and more comprehensive datasets to enhance the robustness and generalizability of the models.

Comparison with traditional HHIE approaches

Our study significantly extends the utility of the Hearing Handicap Inventory for the Elderly (HHIE) by integrating it with cutting-edge machine learning models. Traditionally, HHIE has been used primarily to assess the psychosocial impacts of hearing impairment through subjective experiences and self-perceived hearing difficulties. This novel method improves the precision and accuracy of Age-Related Hearing Loss (ARHL) predictions and introduces a more individualized approach to hearing healthcare by transforming HHIE from a simple self-reporting tool into an advanced predictive tool. We can customize interventions based on each patient's unique risk profile by utilizing machine learning, which combines quantitative, data-driven predictions with qualitative insights into patients' actual experiences. This combination marks a significant breakthrough in the field of hearing healthcare and the start of a new phase of proactive, patient-centered, precision-based care. Our approach provides a comprehensive and nuanced view of the management of hearing health by bridging the gap between subjective self-assessment and objective risk prediction. This adds quantitative and data-driven capabilities to the traditional HHIE approach, enabling a more holistic approach to hearing healthcare.

Conclusion

Our research, which combines cutting-edge machine learning methods with the traditional Hearing Handicap Inventory for the Elderly-Screening (HHIE-S) tool, marks a major advancement in the assessment of hearing loss. In the field of audiology and hearing healthcare, the use of the Light Gradient Boosting Machine (LGBM) classifier in particular has shown impressive progress in improving predictive accuracy. In addition to demonstrating the enormous potential of data-driven approaches to supplement conventional diagnostic tools, this study paves the way for more prompt and precisely targeted interventions for age-related hearing loss (ARHL). The significance of adopting such novel approaches is highlighted by our study, as the healthcare sector continues to undergo swift transformation. Ensuring better patient outcomes and leading the way in the transition to a more proactive, precision-driven paradigm in hearing care depend on this. By utilizing machine learning, we create new opportunities for early detection and intervention, which will ultimately improve the quality of life for those who are impacted by hearing loss.

Acknowledgements

We appreciate the help of the Big Data Centre at Taipei Veterans General Hospital for their assessment, and the constructive suggestions from Dr. Shang-Liang Wu.

Authors' contributions

The manuscript was written by TH Yang, YF Chen, JN Huang, and YC Chu, who also revised the final draft. TH Yang, CS Wu, and YC Chu designed the study, analyzed the data, and drafted the initial version of the manuscript. YF Cheng, CS Wu, and YC Chu contributed to the conception and design of the study, the interpretation of the data, and the critical revision of the manuscript. TH Yang, YF Chen, YF Cheng, JN Huang, CS Wu, and YC Chu reviewed and verified the raw data used in the study. All authors had access to all the data collected and analyzed in the study, and they collectively take responsibility for the accuracy and integrity of the data presented in the manuscript. As authors of the study, they collectively agree to submit the manuscript for publication and acknowledge their respective roles in the research.

Funding

The authors acknowledge the financial support from the National Science and Technology Council (NSTC110-2320-B-075-004-MY3) and Taipei Veterans General Hospital (V112C-067, V112E-001-2) for this study. The funding sources had no involvement in the study design, data collection, analysis, interpretation, or manuscript writing.

Availability of data and materials

The data used in this study are not publicly available, as they were obtained from Taipei City Hospital for Department of Otorhinolaryngology, under a license agreement. The data belong to Tzong-Hann Yang, who is one of the authors of this study. If you have a reasonable request for accessing the data, you can contact Tzong-Hann Yang and ask for his permission. He will decide whether to grant you access to the data, based on the terms and conditions of the license agreement with Taipei City Hospital. Potential conflicts of interest to declare.

Declarations

Ethics approval and consent to participate

This study followed the ethical principles of the Declaration of Helsinki and received approval from the Institutional Review Board of Taipei City Hospital (approval no. TCHIRB-10811011-E), which is the authorized body for research data in Taipei City Hospital. The need for informed consent to participate was waived by Institutional Review Board of Taipei City Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Otorhinolaryngology, Taipei City Hospital, Taipei 100, Taiwan. ²General Education Center, University of Taipei, Taipei 10671, Taiwan. ³Department of Speech-Language Pathology and Audiology, National Taipei University of Nursing and Health Sciences, Taipei 112303, Taiwan. ⁴Department of Otolaryngology-Head and Neck Surgery, School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan. ⁵Department of Medical Research, Taipei Veterans General Hospital, Taipei 112, Taiwan. ⁶School of Medicine, National Yang Ming Chiao Tung University, Taipei 112, Taiwan. ⁷Department of Otolaryngology-Head and Neck Surgery, Taipei Veterans General Hospital, Taipei 112, Taiwan. ⁸Institute of Brain Science, National Yang Ming Chiao Tung University, Taipei 112, Taiwan. ⁹College of Science and Engineering, Fu Jen University, Taipei 243, Taiwan. ¹⁰Information Management Office, Taipei Veterans General Hospital, Taipei 112, Taiwan. ¹¹Big Data Center, Taipei Veterans General Hospital, Taipei 112, Taiwan. ¹²Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei 112, Taiwan.

Received: 4 November 2023 Accepted: 28 November 2023

Published online: 14 December 2023

References

- Mahboubi H, Lin HW, Bhattacharyya N. Prevalence, Characteristics, and Treatment Patterns of Hearing Difficulty in the United States. *JAMA Otolaryngol Head Neck Surg.* 2018;144(1):65–70.
- Chang HP, Chou P. Presbycusis among older Chinese people in Taipei, Taiwan: a community-based study. *Int J Audiol.* 2007;46(12):738–45.
- Kowalewski V, Patterson R, Hartos J, Bugnariu N. Hearing Loss Contributes to Balance Difficulties in both Younger and Older Adults. *J Prev Med (Wilmington).* 2018;3(2):12. <https://doi.org/10.21767/2572-5483>. Scitech Biol Biosoft Pvt. Ltd.
- Campos J, Ramkhalawansingh R, Pichora-Fuller MK. Hearing, self-motion perception, mobility, and aging. *Hear Res.* 2018;369:42–55.
- Jiam NT, Li C, Agrawal Y. Hearing loss and falls: A systematic review and meta-analysis. *Laryngoscope.* 2016;126(11):2587–96.
- Maharani A, Pendleton N, Leroi I. Hearing Impairment, Loneliness, Social Isolation, and Cognitive Function: Longitudinal Analysis Using English Longitudinal Study on Ageing. *Am J Geriatr Psychiatry.* 2019;27(12):1348–56.
- Quaranta N, Coppola F, Casulli M, et al. The prevalence of peripheral and central hearing impairment and its relation to cognition in older adults. *Audiol Neurootol.* 2014;19(Suppl 1):10–4.
- Golub JS, Luchsinger JA, Manly JJ, Stern Y, Mayeux R, Schupf N. Observed Hearing Loss and Incident Dementia in a Multiethnic Cohort. *J Am Geriatr Soc.* 2017;65(8):1691–7.
- Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet.* 2017;390(10113):2673–734.
- Liljas AE, Wannamethee SG, Whincup PH, et al. Hearing impairment and incident disability and all-cause mortality in older British community-dwelling men. *Age Ageing.* 2016;45(5):662–7.
- World Health Organization. Integrated care for older people (ICOPE): guidance for person-centred assessment and pathways in primary care. World Health Organization. 2019. Retrieved from <https://apps.who.int/iris/handle/10665/326843>.
- Watson CS, Kidd GR, Miller JD, Smits C, Humes LE. Telephone screening tests for functionally impaired hearing: current use in seven countries and development of a US version. *J Am Acad Audiol.* 2012;23(10):757–67.
- Folmer RL, Vachhani J, McMillan GP, Watson C, Kidd GR, Feeney MP. Validation of a Computer-Administered Version of the Digits-in-Noise Test for Hearing Screening in the United States. *J Am Acad Audiol.* 2017;28(2):161–9.
- Test i, Medical i. 2016. <https://www.iheardmedical.com/ihear-test/>. Accessed 27 Jan 2020.
- de Swanepoel W, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. *Int J Audiol.* 2014;53(12):841–9.

16. Bright T, Pallawela D. Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review. *JMIR Rehabil Assist Technol*. 2016;3(2): e13.
17. Mosley CL, Langley LM, Davis A, McMahon CM, Tremblay KL. Reliability of the Home Hearing Test: Implications for Public Health. *J Am Acad Audiol*. 2019;30(3):208–16.
18. Yang T-H, Chu Y-C, Chen Y-F, Chen M-Y, Cheng Y-F, Wu C-S, Huang H-M. Diagnostic Validity of Self-Reported Hearing Loss in Elderly Taiwanese Individuals: Diagnostic Performance of a Hearing Self-Assessment Questionnaire on Audiometry. *Int J Environ Res Public Health*. 2021;18(24):13215. <https://doi.org/10.3390/ijerph182413215>.
19. Commodari E, Guarnera M. Attention and aging. *Aging Clin Exp Res*. 2008;20(6):578–84.
20. Smith G. Step away from stepwise. *J Big Data*. 2018;5(1):32.
21. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019;380(14):1347–58.
22. Wang D, Ram P, Weidele DK, et al. AutoAI: Automating the End-to-End AI Lifecycle with Humans-in-the-Loop. In: *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion*; 2020. p. 77–78. <https://dblp.org/rec/conf/iui/WangRW0MMWVCTSA20>.
23. Krishna Priya G, Suganthi S T, Vijipriya G, Nirmala Madian. An approach for predicting heart failure rate using IBM Auto AI Service. 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE); 2021. p. 203–07. <https://doi.org/10.1109/ICCIKE51210.2021.9410783>.
24. Cao T, Roy D, Nedelescu T. Optimizing commute time with IBM Watson studio. In: *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*. Association for Computing Machinery (ACM); 2019. p. 388–90. <https://doi.org/10.1145/3339363>.
25. Bhuvaneshwarri I. Deployment of Heart Disease Prediction Model in Cloud Environment. *Iconic Res Eng J*. 2019;3(5):177–80.
26. Tushir K. AutoAI using IBM Watson. Amity School of Engineering & Technology, Amity University, India. <https://doi.org/10.55162/MCET.04.140>, <https://themedicon.com/journals/engineeringthemes/MCET-04-140>. Published May 30, 2023.
27. Jabbar MA, Deekshatulu BL, Chandra P. Knowledge Discovery Using Associative Classification for Heart Disease Prediction. In: *Intelligent Informatics (Advances in Intelligent Systems and Computing)*. Springer Science and Business Media LLC. 2013;182:29–39. https://doi.org/10.1007/978-3-642-32063-7_4.
28. Chen T, He T, Benesty M, et al. Xgboost: extreme gradient boosting. R package version 04–2. 2015; 1(4): 1–4.
29. Sheridan RP, Wang WM, Liaw A, Ma J, Gifford EM. Extreme gradient boosting as a method for quantitative structure–activity relationships. *J Chem Inf Model*. 2016;56(12):2353–60.
30. Babajide Mustapha I, Saeed F. Bioactive molecule prediction using extreme gradient boosting. *Molecules*. 2016;21(8):983.
31. Peter S, Diego F, Hamprecht FA, Nadler B. Cost efficient gradient boosting. In: *Advances in Neural Information Processing Systems*; 2017. p. 30. https://proceedings.neurips.cc/paper_files/paper/2017/hash/4fac9ba115140ac4f1c22da82aa0bc7f-Abstract.html.
32. Chakrabarty N, Kundu T, Dandapat S, Sarkar A, Kole DK. Flight Arrival Delay Prediction Using Gradient Boosting Classifier. In: *Emerging Technologies in Data Mining and Information Security (Advances in Intelligent Systems and Computing)*. Springer Science and Business Media LLC. 2019;813:651–59. https://doi.org/10.1007/978-981-13-1498-8_57.
33. Lusa L. Gradient boosting for high-dimensional prediction of rare events. *Comput Stat Data Anal*. 2017;113:19–37.
34. Swarup Kumar JNV, Hareesh Kumar K, Haleem A, Sivaranjani B, Tirumala Kiran BN, Prameela S. IBM Auto AI Bot: Diabetes Mellitus Prediction Using Machine Learning Algorithms. *International Conference on Applied Artificial Intelligence and Computing (ICAIC)*. 2022;2022:24–9. <https://doi.org/10.1109/ICAIC53929.2022.9792690>.
35. Bertl M, Ross P, Draheim D. Predicting Psychiatric Diseases Using AutoAI: A Performance Analysis Based on Health Insurance Billing Data. In: *Database and Expert Systems Applications (Lecture Notes in Computer Science)*. Springer Science and Business Media LLC. 2021;12923:104–111. https://doi.org/10.1007/978-3-030-86472-9_9.
36. Alzamzami F, Hoda M, El Saddik A. Light gradient boosting machine for general sentiment classification on short texts: a comparative evaluation. *IEEE Access*. 2020;8:101840–58.
37. Kumar S, Mallik A, Panda B. Link prediction in complex networks using node centrality and light gradient boosting machine. *World Wide Web*. 2022;25(6):2487–513.
38. Mishra D, Naik B, Nayak J, Sourji A, Dash PB, Vimal S. Light gradient boosting machine with optimized hyperparameters for identification of malicious access in IoT network. *Digital Commun Netw*. 2023;9(1):125–37.
39. Nirmala M, Saravanan V. Clinical Implication of Machine Learning Based Cardiovascular Disease Prediction Using IBM Auto AI Service. *Int J Res Appl Sci Eng Technol*. 2022;10(8):124–44. <https://doi.org/10.22214/ijraset.2022.46087>.
40. LaValley MP. Logistic regression. *Circulation*. 2008;117(18):2395–9.
41. Wright RE. Logistic regression. In: Grimm LG, Yarnold PR, editors. *Reading and understanding multivariate statistics*. American Psychological Association; 1995. p. 217–44.
42. Nick TG, Campbell KM. Logistic Regression. In *Topics in Biostatistics (Methods in Molecular Biology™)*. Springer Science and Business Media LLC. 2007;404:273–301. https://doi.org/10.1007/978-1-59745-530-5_14.
43. Chadha A, Kaushik B. A hybrid deep learning model using grid search and cross-validation for effective classification and prediction of suicidal ideation from social network data. *N Gener Comput*. 2022;40(4):889–914.
44. Diao X, Huo Y, Zhao S, et al. Automated ICD coding for primary diagnosis via clinically interpretable machine learning. *Int J Med Informatics*. 2021;153: 104543.
45. Jiang X, Xu C. Deep learning and machine learning with grid search to predict later occurrence of breast Cancer metastasis using clinical data. *J Clin Med*. 2022;11(19):5772.
46. Ou S-M, Tsai M-T, Lee K-H, et al. Prediction of the risk of developing end-stage renal diseases in newly diagnosed type 2 diabetes mellitus using artificial intelligence algorithms. *BioData Mining*. 2023;16(1):8.
47. Chen Y-C, Chu Y-C, Huang C-Y, Lee Y-T, Lee W-Y, Hsu C-Y, Yang AC, Liao W-H, Cheng Y-F. Smartphone-based artificial intelligence using a transfer learning algorithm for the detection and diagnosis of middle ear diseases: A retrospective deep learning study. *eClinicalMedicine*. 2022;51:101543. <https://doi.org/10.1016/j.eclinm.2022.101543>. Elsevier BV.

48. Duchêne J, Billiet L, Franco V, Bonnard D. Validation of the French version of HHIE-S (Hearing Handicap Inventory for the Elderly - Screening) questionnaire in French over-60 year-olds. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2022;139(4):198–201. <https://doi.org/10.1016/j.anorl.2021.11.003>. Elsevier BV.
49. Li LY, Wang S-Y, Wu C-J, Tsai C-Y, Wu T-F, Lin Y-S. Screening for Hearing Impairment in Older Adults by Smartphone-Based Audiometry, Self-Perception, HHIE Screening Questionnaire, and Free-Field Voice Test: Comparative Evaluation of the Screening Accuracy With Standard Pure-Tone Audiometry. *JMIR mHealth and uHealth*. 2020;8(10):e17213. <https://doi.org/10.2196/17213>. JMIR Publications Inc.
50. Arnold ML, Schwartz B, Neil H, Chisolm TH, Sanchez VA. Feasibility and Assessment of a Hybrid Audiology Service Delivery Model for Older Adult Hearing Aid Users: A Pilot Study. *Am J Audiol*. 2022;31(35):892–904. https://doi.org/10.1044/2022_AJA-21-00200. American Speech Language Hearing Association.
51. Sulaiman A. Exploring the Potential of Hearing Screening Smartphone Applications to Enhance Access to Hearing Healthcare: A Literature Review. 2023.
52. World Health Organization. WHO ear and hearing: survey handbook. World Health Organization. 2020. Retrieved from <https://apps.who.int/iris/handle/10665/331630>.
53. Wasmann J-W, Pragt L, Eikelboom R, Swanepoel DW. Digital Approaches to Automated and Machine Learning Assessments of Hearing: Scoping Review. *J Med Internet Res*. 2022;24(2):e32581. <https://doi.org/10.2196/32581>. JMIR Publications Inc.
54. Rodrigo H, Beukes EW, Andersson G, Manchaiah V. Exploratory Data Mining Techniques (Decision Tree Models) for Examining the Impact of Internet-Based Cognitive Behavioral Therapy for Tinnitus: Machine Learning Approach. *J Med Int Res*. 2021;23(11):e28999. <https://doi.org/10.2196/28999>. JMIR Publications Inc.
55. Liao W-H, Cheng Y-F, Chen Y-C, Lai Y-H, Lai F, Chu Y-C. Physician decision support system for idiopathic sudden sensorineural hearing loss patients. *J Chin Med Assoc*. 2021;84(1):101–107. <https://doi.org/10.1097/JCMA.0000000000000450>. Ovid Technologies (Wolters Kluwer Health).
56. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care. *JAMA Network Open*. 2022;5(9):e2233946. <https://doi.org/10.1001/jamanetworkopen.2022.33946>. American Medical Association (AMA).
57. Gnanasegaram JJ, Leung R, Beyea JA. Evaluating the effectiveness of learning ear anatomy using holographic models. *J Otolaryngol Head Neck Surg* 2020;49(1). <https://doi.org/10.1186/s40463-020-00458-x>. Springer Science and Business Media LLC.
58. Sammari E, Naceur A. The Effectiveness of an Intervention Program for the Development of Cognitive Capacities of Children with a Hearing Impairment. 2023. <https://www.asjp.cerist.dz/en/article/223060>.
59. Chu Y-C, Cheng Y-F, Lai Y-H, Tsao Y, Tu T-Y, Young ST, Chen T-S, Chung Y-F, Lai F, Liao W-H. A Mobile Phone-Based Approach for Hearing Screening of School-Age Children: Cross-Sectional Validation Study. *JMIR mHealth uHealth*. 2019;7(4):e12033. <https://doi.org/10.2196/12033>. JMIR Publications Inc.
60. Lin H-YH, Chu Y-C, Lai Y-H, Cheng H-L, Lai F, Cheng Y-F, Liao W-H. A Smartphone-Based Approach to Screening for Sudden Sensorineural Hearing Loss: Cross-Sectional Validity Study. *JMIR mHealth uHealth*. 2020;8(11):e23047. <https://doi.org/10.2196/23047>. JMIR Publications Inc.
61. Mangalathu S, Hwang S-H, Jeon J-S. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng Struct*. 2020;219:110927. <https://doi.org/10.1016/j.engstruct.2020.110927>. Elsevier BV.
62. Ye X, Zhu D, Chen S, Shi X, Gong R, Wang J, Zuo H, He P. Effects of providing free hearing aids on multiple health outcomes among middle-aged and older adults with hearing loss in rural China: a randomized controlled trial. *BMC Med*. 2022;20(1). <https://doi.org/10.1186/s12916-022-02323-2>. Springer Science and Business Media LLC.
63. Sindhusake D, Mitchell P, Smith W. Validation of self-reported hearing loss. The Blue Mountains Hearing Study. *Int J Epidemiol*. 2001;30(6):1371–8.
64. Ting HC, Huang YY. Sensitivity and specificity of hearing tests for screening hearing loss in older adults. *J Otol*. 2023;18(1):1–6.
65. Feltner C, Wallace IF, Kistler CE, Coker-Schwimmer M, Jonas DE. Screening for Hearing Loss in Older Adults: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA*. 2021;325(12):1202–15.
66. Oosterloo BC, Homans NC, Baatenburg de Jong RJ, Ikram MA, Nagtegaal AP, Goedegebure A. Assessing hearing loss in older adults with a single question and person characteristics; Comparison with pure tone audiometry in the Rotterdam Study. *PLoS One*. 2020;15(1):e0228349.
67. Agrawal Y, Platz EA, Niparko JK. Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the National Health and Nutrition Examination Survey, 1999–2004. *Arch Intern Med*. 2008;168(14):1522–30.
68. Homans NC, Metselaar RM, Dingemans JG, et al. Prevalence of age-related hearing loss, including sex differences, in older adults in a large cohort study. *Laryngoscope*. 2017;127(3):725–30.
69. Kim S, Lim EJ, Kim HS, Park JH, Jarng SS, Lee SH. Sex Differences in a Cross Sectional Study of Age-related Hearing Loss in Korean. *Clin Exp Otorhinolaryngol*. 2010;3(1):27–31.
70. Kiely KM, Gopinath B, Mitchell P, Browning CJ, Anstey KJ. Evaluating a dichotomized measure of self-reported hearing loss against gold standard audiometry: prevalence estimates and age bias in a pooled national data set. *J Aging Health*. 2012;24(3):439–58.
71. Ferrite S, Santana VS, Marshall SW. Validity of self-reported hearing loss in adults: performance of three single questions. *Rev Saude Publica*. 2011;45(5):824–30.
72. Kamil RJ, Genthner DJ, Lin FR. Factors associated with the accuracy of subjective assessments of hearing impairment. *Ear Hear*. 2015;36(1):164–7.
73. Kim SY, Kim HJ, Kim MS, Park B, Kim JH, Choi HG. Discrepancy between self-assessed hearing status and measured audiometric evaluation. *PLoS ONE*. 2017;12(8): e0182718.
74. Valette-Rosalino CM, Rozenfeld S. Auditory screening in the elderly: comparison between self-report and audiometry. *Braz J Otorhinolaryngol*. 2005;71(2):193–200.

75. Choi JS, Betz J, Deal J, et al. A Comparison of Self-Report and Audiometric Measures of Hearing and Their Associations With Functional Outcomes in Older Adults. *J Aging Health*. 2016;28(5):890–910.
76. Costa-Guarisco LP, Dalpuel D, Labanca L, Chagas MHN. Perception of hearing loss: use of the subjective faces scale to screen hearing among the elderly. *Cien Saude Colet*. 2017;22(11):3579–88.
77. Boatman DF, Miglioretti DL, Eberwein C, Alidoost M, Reich SG. How accurate are bedside hearing tests? *Neurology*. 2007;68(16):1311–4.
78. Pirozzo S, Papinczak T, Glasziou P. Whispered voice test for screening for hearing impairment in adults and children: systematic review. *BMJ*. 2003;327(7421):967.
79. World Health Organization. Integrated care for older people (ICOPE): guidance for person-centred assessment and pathways in primary care. 2023. <https://www.who.int/publications/i/item/WHO-FWC-ALC-19.1>. Accessed 12 Mar 2023).
80. Ranganathan B, Counter P, Johnson I. Validation of self-reported hearing loss using television volume. *J Laryngol Otol*. 2011;125(1):18–21.
81. McArdle RA, Wilson RH, Burks CA. Speech recognition in multitalker babble using digits, words, and sentences. *J Am Acad Audiol*. 2005;16(9):726–39 quiz 63–4.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

