# IDENTIFICATION OF TYPE AND PERCENTAGE OF HEARING LOSS

Robotics And Automation, Symbiosis Institute Of Technology

***Abstract*-**  Hearing loss detection is a crucial aspect of healthcare, with implications for an individual's quality of life, health outcomes, and public health activities. It may hinder their ability to communicate, socialize, and participate in activities, resulting in feelings of depression and isolation. Early detection and intervention may help in minimizing these impacts and improve an individual's overall well-being.Machine learning models have the potential of improving the accuracy and efficiency of hearing loss identification, allowing for early treatment and improved results for those with hearing loss.In this study we focus on developing a machine learning algorithm for identification of hearing loss and percentage of disability using classification and regression models respectively. The results showed that the stacking model with final estimator as Random Forest Classifier and regression model Random Forest Regressor were the most effective compared to other explainability techniques.

***Index Terms***- Conductive Hearing Loss, Mixed Hearing Loss, Hearing Disability,Pure Tone Average, Percentage of disability.

1. Introduction

Hearing loss is the inability to hear sounds in one or both ears.It can be temporary or permanent, and can affect people of all ages. A variety of factors can contribute to it, including age, genetics, loud noise exposure, infections, and certain medical conditions. Hearing loss is a relatively common condition, particularly among the elderly. According to the World Health Organization (WHO), approximately 466 million people worldwide,and this number is expected to increase to over 900 million by 2050. In India the prevalence of hearing loss is higher in rural areas than in urban areas, and it is also more common among older adults. In addition, noise pollution is a growing problem in urban areas, which can contribute to hearing loss. According to WHO estimates, around 63 million people in India suffer from Severe Auditory Impairment, placing the estimated prevalence at 6.3% of the Indian population. Based on the NSSO survey, there are currently 291 persons per lakh who are suffering from severe to profound hearing loss. A substantial percentage of these are youngsters aged 0 to 14 years. With such a big proportion of hearing-impaired young Indians, it amounts to a significant loss of physical and economic productivity. Milder degrees of hearing loss and unilateral (one-sided) hearing loss impact an even greater percentage of our population.Unfortunately, many people with hearing loss in India do not have access to adequate healthcare services or hearing aids, which can make it difficult for them to communicate effectively and participate fully in society. There is a need for increased awareness, education, and resources to address this issue.

Having disabling hearing loss, is defined as hearing loss in the better hearing ear greater than 40 decibels in adults and greater than 30 decibels in children. Hearing loss can occur gradually or suddenly, and can affect one or both ears.The degree of hearing loss can range from mild to profound and can affect different frequencies of sound differently. Some common signs of hearing loss include :

1. Difficulty understanding speech, particularly in noisy environments
2. Frequently asking people to repeat themselves or speak more loudly
3. Turning up the volume on the TV or radio to a level that others find too loud
4. Missing parts of conversations or misunderstandings
5. Difficulty hearing high-pitched sounds, such as a doorbell or phone ringing
6. Feeling like people are mumbling or not speaking clearly
7. Tinnitus (ringing, buzzing, or hissing sound in the ears)
8. Feeling fatigued or stressed from the effort of trying to hear
9. If you are experiencing any of these symptoms, it is important to consult with a hearing healthcare professional for an evaluation and proper diagnosis.

Hearing loss is classified into three types: Conductive hearing loss which occurs when there is a problem with the transmission of sound waves through the outer or middle ear. This can be caused by a blockage in the ear canal, damage to the eardrum or the small bones in the middle ear, or fluid accumulation in the middle ear. Secondly, sensorineural hearing loss occurs when there is damage to the inner ear (cochlea) or to the auditory nerve that carries sound signals from the inner ear to the brain. This type of hearing loss is the most common and can be caused by a variety of factors such as aging, exposure to loud noise, genetics, certain medications, and some medical conditions like Meniere's disease or acoustic neuroma and mixed hearing loss is a combination of conductive hearing loss and sensorineural hearing loss. This means that there is both a problem with the transmission of sound through the outer or middle ear (conductive hearing loss) and damage to the inner ear or auditory nerve (sensorineural hearing loss).Causes of mixed hearing loss can include a variety of factors such as chronic ear infections, head trauma, genetic conditions, exposure to loud noise, and certain medical conditions that affect the inner ear or the nerves that transmit sound.

Hearing tests are important for identifying hearing loss, determining the type and degree of hearing loss, and guiding the selection of appropriate treatment options, such as hearing aids, cochlear implants, or other interventions. Diagnosis of hearing loss typically involves a hearing evaluation performed by a healthcare professional, such as an audiologist or an ear, nose, and throat (ENT) doctor. The evaluation may include several tests to assess the type, severity, and possible causes of hearing loss .

Some common hearing tests used for diagnosis include:

1. Pure tone audiometry (PTA) - This test measures a person's hearing sensitivity across a range of frequencies and is the most common type of hearing test.
2. Speech audiometry - This test assesses a person's ability to hear and understand speech, often in noisy environments.
3. Tympanometry - This test measures the movement of the eardrum in response to changes in air pressure, which can help identify problems with the middle ear.
4. Otoacoustic emissions (OAE) - This test measures sounds that are produced by the inner ear in response to sounds played into the ear.
5. Auditory brainstem response (ABR) - This test measures the electrical activity of the auditory nerve and brainstem in response to sounds
6. Based on the results of these tests, the healthcare professional can determine the type and severity of hearing loss, and recommend appropriate treatment options, such as hearing aids, cochlear implants, or other assistive listening devices.

Treatment options for hearing loss depend on the cause and degree of hearing loss and may include hearing aids, cochlear implants, and other assistive listening devices. To treat the underlying cause of hearing loss, medication or surgery may be required in some cases. If you suspect you have hearing loss, it is critical to seek medical attention because early diagnosis and treatment can improve outcomes.

1. Previous work

1. Countless studies have produced hearing loss techniques and approaches that can enhance or facilitate the tasks of otolaryngology practitioners. A decision support system for evaluating hearing loss symptoms using machine learning techniques was developed to aid physicians with diagnosis of hearing loss. This research employs a secondary type of data.The data was gathered from Malaysia's National Medical Research Register (NMRR), which is considered the official data bank in the healthcare profession. For this research, the researchers received approval for both NMRR registration and sample data collection. The FP-Growth and association rule algorithms were tested on the two datasets, a small sample dataset and a large sample dataset, in order to determine the relationship between hearing thresholds and hearing loss symptoms with a very low error rate.The classifier is a multivariate Bernoulli naive Bayes classification model, and the feature transformation method is the Frequent Pattern Growth (FP-Growth) algorithm. The findings demonstrated a correlation between hearing thresholds and hearing loss symptoms,and a mix of the FP-Growth and naive Bayes algorithms was found to be effective at identifying hearing-loss symptoms with a very low error rate.With five training sets, the average accuracy and average error rates and average error rate for the multivariate Bernoulli model with FP-Growth feature transformation are 98.25% and 1.73%, respectively.This implies that in the majority of cases, the model was successful in accurately identifying the symptoms of hearing loss.

2. An alternative approach for hearing loss detection involved gathering data from 207 adults of various native languages at hearing loss prevention and awareness events. The participants had to communicate with the platform via a mouse. The age requirement was 18 years, and the exclusion factors were hearing aids or implantable devices.The article examines the performance of various machine learning models in detecting hearing loss from a speech-in-noise screening test. The findings demonstrated that machine learning models such as LR, SVM, KNN, or RF are capable of determining the presence of hearing loss with a high degree of accuracy. These models were found to be more effective than a univariate classifier based on the Speech Reception Threshold (SRT) in identifying mild to severe hearing loss.The logistic regression model had a precision of 0.87, while the support vector machine (SVM) model had a precision of 0.86. The random forest (RF) model obtained 0.85 accuracy, while the k-nearest neighbors (KNN) model achieved 0.83 accuracy. In addition, the paper describes the accuracy of each single feature across the entire dataset in terms of area under the receiver operating curve (AUC), accuracy, and cutoff value.

3. A study also reflects on using machine learning algorithms to predict the success of hearing preservation surgery in patients with vestibular schwannoma. The research compared four different algorithms: support vector machine (SVM), gradient boosting machine (GBM), deep neural network (DNN), and diffuse random forest. (DRF). The model's variables were categorized as either continuous or categorical, with the surgical approach being the only categorical variable. Supervised machine learning was used, with boosting and bagging being the most common methods among tabular datasets. The estimated accuracy of hearing preservation varied from 62% (SVM) to 90%. (DNN).

4. Approaches involving examination of the miRNA expression profiles of people with various degrees of sensorineural hearing loss (SNHL) in order to find miRNAs that might be essential in developing a diagnostic algorithm for SNHL. The study's findings demonstrated that cochlear implant patients with and without residual hearing could be distinguished using miRNA expression profiles with 100% precision. The miRNAs which influence the model were given an overall score by the researchers using a permutation feature of significance. Following this, they mapped out various regulatory pathways using the programme Ingenuity Pathway Analysis to identify known and highly predicted miRNA cochlear mRNA interactions. As a result, they were successful to identify miRNA profiles distinctive to ear diseases and utilize those profiles to determine the severity of SNHL.

5. Based on a cognitively inspired feature extraction and speech recognition approach, this study provides an innovative approach that automatically identifies any hearing issue. For feature extraction, the proposed system employs an adaptive filter bank with weighted Mel-frequency cepstral coefficients.The adaptive filter bank implementation is based on the notion of spectrum sensing in cognitive radio, that is aware of its surroundings and adapts to statistical differences in input stimuli by learning from those. The comparative performance study shows that our automated hearing test approach has the potential to produce results that are comparable to the clinical actualities established by an expert audiologist testing.A hidden Markov model achieves an overall accuracy of 96.67%. (HMM). The proposed method could offer audiologists a second opinion and act as a cost-effective pre-screening test to anticipate hearing loss at an early stage.

1. Material and Method

Hearing tests assess your hearing capacity. In order to hear normally, sound waves must enter your ear and cause the eardrum to vibrate. The waves are propagated further into the ear by the vibration, which causes nerve cells in the ear to activate and send sound information to the brain. The noises you hear are a translation of this information.

There are several techniques available for evaluating hearing. Hearing tests can be broadly divided into two categories: Screening tests and Diagnostic tests. A screening test simply reveals if someone has hearing loss or not, it does not reveal the type or severity of the hearing loss. An audiologist will conduct diagnostic tests to determine the degree and type of hearing loss.

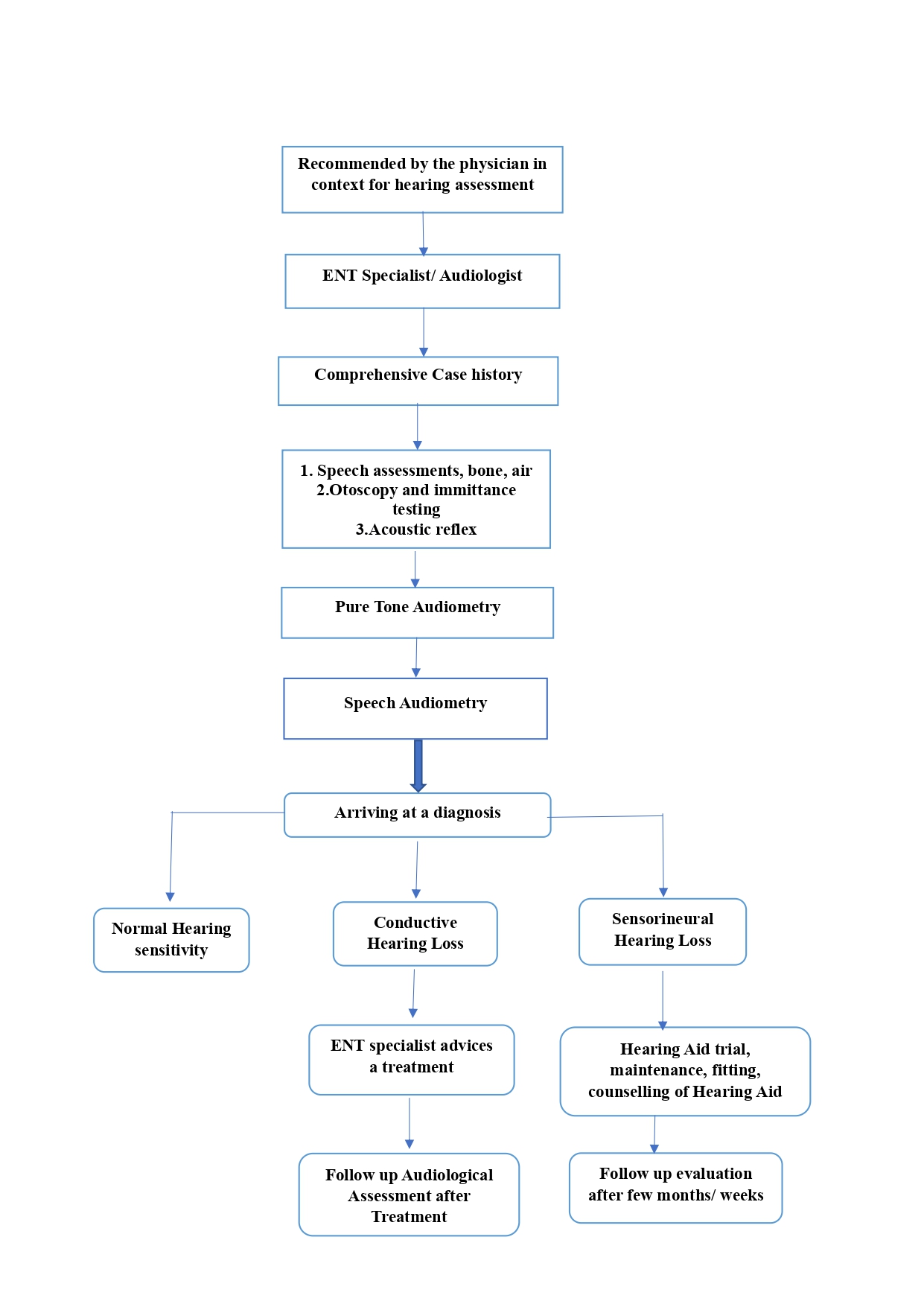


Figure 3.1: Procedure of hearing test

*A. Pure-Tone Audiometry Test*

Pure-tone audiometry is used to measure the auditory threshold of an individual. The instrument used in the measurement of auditory threshold is known as the audiometer.It is the most common type of hearing test and is usually performed in a soundproof booth by an audiologist or hearing healthcare professional. It includes a signal generator that generates a pure tone. A pure-tone signal has energy at only one frequency. This pure-tone signal's intensity and frequency can be adjusted. It should always be performed in a room free of background noise. Even faint or soft sounds can have an impact on the test results. The rooms should be acoustically treated so that the noise inside the room does not interfere with testing.

This test includes two parts: 1. Air conduction test 2. Bone conduction test. Air conduction testing is a type of hearing test that measures the ability of the inner ear (cochlea) and the nerve pathways that lead from the inner ear to the brain to detect and process sounds. During air conduction testing, sounds are presented through headphones or speakers, and the patient is asked to respond when they hear the sound.The test typically involves a range of tones and frequencies , with the volume of the sounds increasing and decreasing to determine the patient's hearing thresholds.**.**Whereas, Bone conduction testing is a type of hearing test that measures the ability of the inner ear (cochlea) to detect sounds that are transmitted through the bones of the skull, bypassing the outer and middle ear using a small device called a bone conductor is placed on the bone behind the ear, typically the mastoid bone.The patient is asked to indicate when they hear the sound, and the volume of the sound is gradually increased and decreased to determine the patient's hearing thresholds. This type of testing is used to assess the integrity of the inner ear and to determine the type and severity of hearing loss.The results of the bone conduction test and air conduction test are plotted on an audiogram, which shows the softest sounds the patient can hear at different frequencies.Together, these tests can help diagnose the type and severity of hearing loss, and determine the most appropriate course of treatment. Bone conduction testing may also be used to evaluate the efficacy of hearing aids or other devices designed to improve hearing.

1. **Part 1: Air Conduction Testing**

1. Place Headphone – Blue on left; red on right.

2. First test better ear (from case history).

3. Select 1000 Hz and 40/70/90dB.

4. If the client responds, decrease the intensity by 10, if not respond, increase the intensity by 5. and present the sound.

5. Repeat this till a lowest threshold at which 2/3 times responses are present.

6. Change frequency and repeat the whole procedure 2000, 4000, 8000.

7. Recheck at 1000 Hz.

8. If less than 5 dB – then the whole procedure 500 Hz and 250 Hz.

9. If more than 5 dB – 1000, 2000, 4000, 8000, 1000, 500, 250 Hz.

10. Repeat the procedure for 500 Hz, 250 Hz.

11. Repeat the procedure for the other ear.

1. **Part 2: Bone Conduction Test**

1. This is an important measurement of hearing threshold using a bone vibrator.

2. This helps to differentiate conductive from sensorineural hearing loss.

3. The equipment necessary is just a bone vibrator connected to the audiometer.

4. The bone vibrator is placed over the mastoid process of the side to be tested.

5. The auditory threshold is assessed as described for air conduction assessment.

The PTA test can also be used to assess the person's ability to understand speech in different listening environments. This is usually done by testing the person's speech recognition thresholds (SRT) and word recognition scores (WRS) in quiet and noisy environments.Overall, the PTA test is a valuable tool for diagnosing and monitoring hearing loss and can help guide the selection of appropriate hearing aids or other interventions to improve communication and quality of life.

## B. Identification Of Total % Of Monaural Hearing Disability

The calculation of the percentage of hearing disability can vary depending on the specific method used for measuring hearing loss and the criteria used for defining disability. Here's an example of how the percentage of hearing disability can be calculated:

1. An audiogram is conducted to measure the person's hearing ability in both ears. The results are plotted on a graph called an audiogram, which shows the person's hearing thresholds at different frequencies.
2. The average hearing threshold in decibels (dB) is calculated for the frequencies of 500, 1000, and 2000 Hz, which are important for speech understanding.
3. The hearing threshold is compared to the average hearing threshold of a person with normal hearing, which is 0-25 dB. For example, if the person's average hearing threshold is 60 dB, they have a moderate hearing loss.
4. The % of disability of each ear is examined to determine the hearing capability of both the ears. The ear with less % of disability is considered as the Better Ear (BE) and the ear with greater % of disability is considered as the Poorer Ear (PE).

Using the formula :

[(5 X BE % OF HD) + PE % OF HD] /6

Overall % of disability is calculated for the individual. If the % of disability is 40% or above then he/she falls in the criteria of disability.

Additionally, hearing disability can also be measured and classified in other ways, such as by using speech recognition tests or functional assessments.

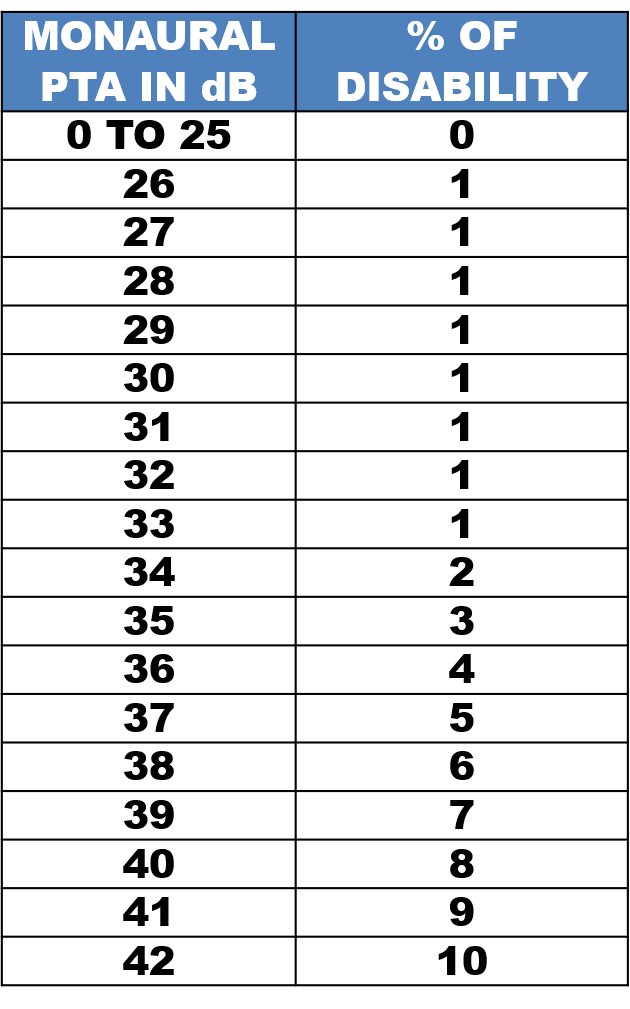
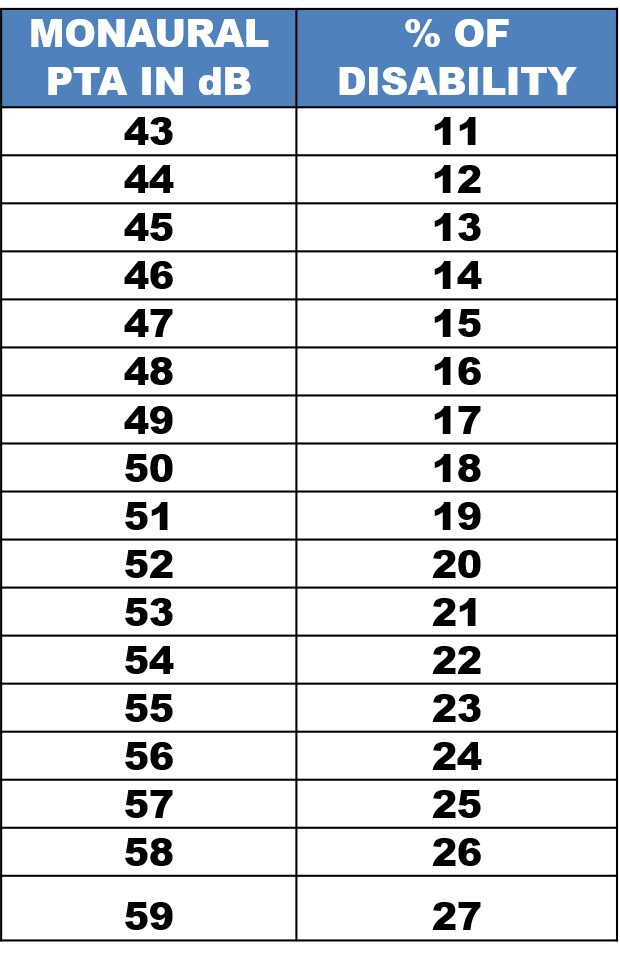
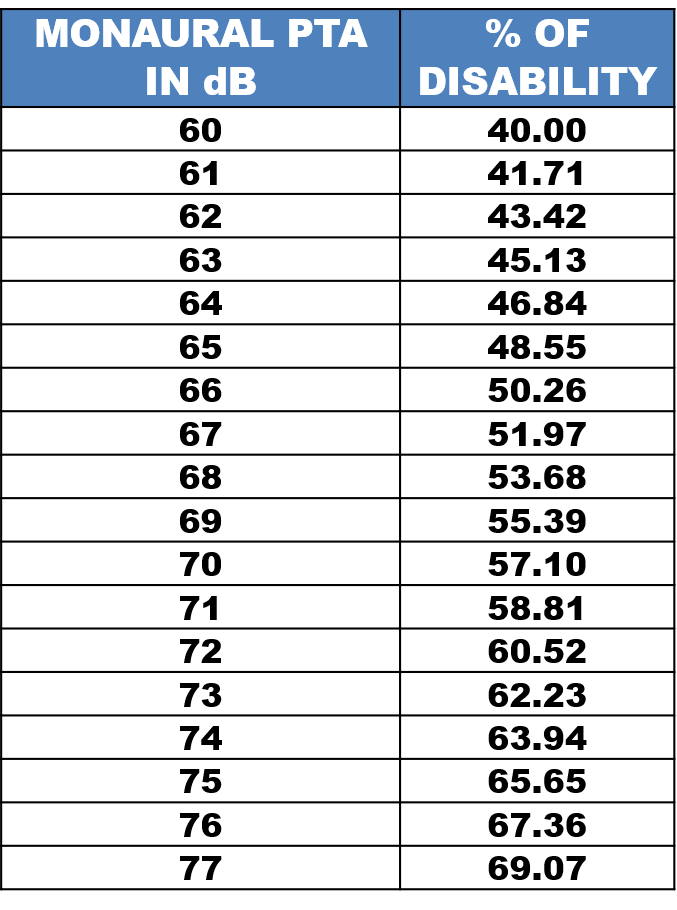
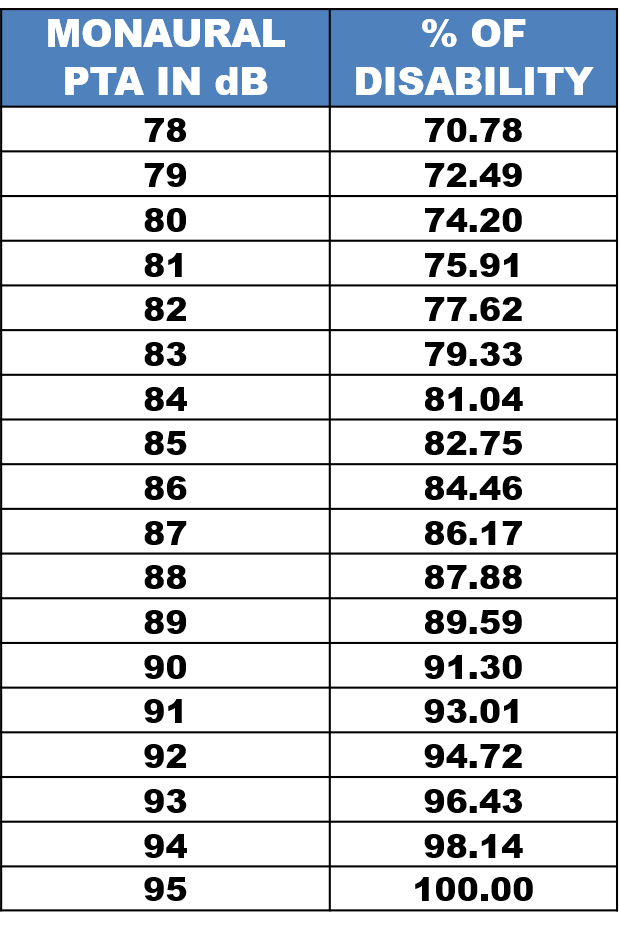
   

Table 3.2: Mapping of PTA with % of disability

*C. Dataset Description*

The dataset used in this research has 5000 values for each ear (left and right ear). It is divided into two sections: air conduction frequency and bone conduction frequency, with frequencies ranging from 250 Hz to 500 Hz, 1000 Hz to 2000 Hz, and 4000 Hz. These frequencies are used to obtain air conduction and bone conduction test results, which are used to diagnose the type and severity of hearing loss.For assessing the degree of hearing loss, the air conduction test data is examined and the four frequencies from the given column of frequencies (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz) of the sound obtained in dB are then taken into consideration,and an average is taken .This average is known as the "Pure Tone Average (PTA)," which is used to define the severity of the hearing loss and the percentage hearing disability of that particular ear.Following this, bone conduction and air conduction frequencies are compared to determine the type of hearing loss.If the values are less than 15 dB, it is considered conductive; else, each bone conduction frequency value is considered and compared with the corresponding air conduction frequency values. If the difference is less than or equal to 10 dB, the hearing loss is sensorineural. It is a mixed type of hearing loss if the difference is more than 10 dB. There are six main categories of hearing loss in the dataset: minimum hearing loss, profound hearing loss, mild hearing loss, moderate hearing loss, severe hearing loss, and moderately severe hearing loss.Considering the various types and severity of hearing loss, there are approximately 17 different types of hearing disability combinations.Each hearing disability is anticipated by applying predefined sets of rules to characterize the patient's hearing impairment.The overall percentage of hearing loss of an individual is calculated by calculating the percentage of disability of each ear using a formula that takes the severity of the hearing loss into account.

| **Type of impairment** | **Range (dB)** | **Probable communication problems** |
| --- | --- | --- |
| Normal | -10 to 15 | No problem in hearing sounds |
| Minimal | 16 to 25 | Minimal hearing loss |
| Mild | 26 to 40 | Has difficulty hearing faint or distant speech |
| Moderate | 41 to 55 | Understands conversational speech only from a distance of 3-5 feet |
| Moderately Severe | 56 to 70 | Conversation must be loud to be understood. There is great difficulty in group conversation |
| Severe | 71 to 90 | May hear a loud voice about one foot from the ear, may identify environmental noises; may distinguish vowels but not consonants |
| Profound (Total deafness) | > 90 | May hear loud sounds; does not rely on hearing as a primary channel for communication |

## 

## D. Proposed Model

The main objective of this research was to evaluate the effectiveness of different types of machine learning models used in hearing screening tests and to look into the significance of measured features in the detection of hearing loss using explainability techniques. Random Forest, Decision Tree, Logistic Regression, Support Vector Machine (SVM), KNN Classifier, and Naive Bayes were among the six machine learning models used, with Random Forest outperforming other models in terms of accuracy and precision.The next step in this study is to compute the percentage of disability using regression models. The regression models used are Decision Tree Regressor, Linear Regressor, Support Vector Regressor, Random Forest Regressor. The results reflected that Random Forest Regressor is more accurate in calculating the percentage of disability.

1. Random Forest is a machine learning algorithm that combines the concepts of ensemble learning and decision trees to improve the accuracy and robustness of predictions.The algorithm works by creating multiple decision trees using a random subset of the available features and a random sample of the training data. Each decision tree is trained on a different subset of the data and features, resulting in a collection of independent models.To make a prediction, the random forest algorithm combines the predictions of all the individual decision trees, either through averaging or voting. The ensemble of trees helps to reduce overfitting and improves the accuracy of the predictions.

2. K-Nearest Neighbors (KNN) is a simple and effective machine learning algorithm used for classification and regression.The algorithm works by storing all the available cases and classifying new cases based on the similarity measures (such as distance) between the features of the new data and the existing data points. The "K" in KNN refers to the number of nearest neighbors to consider when making a prediction.For example, in a classification task, the KNN algorithm finds the K nearest neighbors of a new data point in the feature space, and the class of the new data point is determined by a majority vote of the class labels of its K nearest neighbors.

3. Support Vector Machines (SVMs) are a powerful and widely used machine learning algorithm for classification, regression, and outlier detection. SVMs work by finding the hyperplane that best separates the data into different classes.In a binary classification problem, SVMs find the hyperplane that maximizes the margin between the two classes. The margin is defined as the distance between the hyperplane and the closest data points of each class. The data points that lie on the margin are called support vectors and play a critical role in determining the hyperplane. SVMs aim to find the hyperplane that maximizes the margin while minimizing the classification error.

4.A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is a supervised learning algorithm used in the field of data mining and machine learning for making decisions.In a decision tree, each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label or a probability distribution over class labels. The tree is constructed by recursively splitting the data into subsets based on the attribute that best separates the classes or reduces the impurity of the dataset.They can handle both categorical and numerical data, are easy to understand and interpret, and can be used for both classification and regression tasks. They can also handle missing values and outliers in the data.

Stacking is an ensemble learning technique that involves combining the predictions of multiple base models using a meta-model.In stacking, the final estimator is typically a simple machine learning algorithm, such as linear regression, logistic regression, or a neural network, that takes the predictions of the base models as input features and learns to make the final predictions.The choice of the final estimator is an important consideration in building an ensemble model, as it can have a significant impact on the overall performance of the model.Once the final estimator is trained, it can be used to make predictions on new, unseen data by first passing the data through the base models to create a new feature matrix, and then passing the feature matrix through the final estimator to make the final predictions.The advantage of stacking is that it can combine the strengths of different machine learning models to produce a more accurate and robust prediction. It can also handle complex relationships between the features and the target variable, and it can help to reduce overfitting.

Steps involved in training the machine learning model to predict the type of hearing loss and percentage of disability are as follows :

1. Data Collection: The first step in building a machine learning model is to collect the data. The quality and quantity of the data collected determines the performance of the model. The collected data is described in the previous section.
2. Data Preprocessing: Once the data is collected, the next step is preprocessing . The unwanted columns with empty values and those which are not contributing anything towards the final predictions are eliminated.
3. Data Splitting: Following preprocessing, the data is divided into two parts: one for identifying the type of hearing loss and the other for predicting the percentage of hearing disability. Both parts are further divided for training and testing. The training set is used to train the machine learning model, whereas the testing set is used to assess the model's performance. The input features for the model are air conduction and bone conduction frequency, which assist in predicting the type of hearing disability and the percentage disability of each ear, which will be utilized in estimating the total hearing disability.
4. Model Selection and Training: There are various machine learning algorithms available,and six of those were examined for predicting the type of hearing loss: Random Forest, Decision Tree, Logistic Regressor, Support Vector Machine (SVM), KNN Classifier, and Naive Bayes.. Different Regressor models, including Linear Regressor, Support Vector Regressor, Decision Tree Regressor, and Random Forest Regressor, are used to predict the total percentage of hearing disability; the Random Forest Regressor achieved the best R2 score of 0.999.
5. Model Training: Once the algorithm is selected, the next step is to train the model using the training set. During training, the model learns to make predictions by adjusting its parameters based on the input data. To further increase the accuracy of the model, stacking technique is used, which takes the predicted value of base models as input for the final estimator. We have KNN, Decision Tree and SVM as our base estimators which are trained on the training data and then the final estimator i.e. Random Forest Classifier is used for making the final prediction based on the predictions of the base estimator. We will be directly training the Random Forest Regressor model as it provides the best score.
6. Model Evaluation: After the model has been trained, one must assess its effectiveness using the testing set. The stacking model had the best accuracy of 93% as we evaluated the classification model using the classification metric. As we examined the regression model using the R2 score, the Random Forest Regressor model achieved an R2 score of 0.999.
7. RESULTS

Random Forest (90.5%), Decision Tree (80.7%), Logistic Regression (34.9%), Support Vector Machine (SVM) (84.6%), KNN Classifier (87.3%), and Naive Bayes (65.3%) outperformed other models when employed for hearing loss detection. For determining the percentage of disability, Random Forest Regressor (99.91%) had the highest accuracy compared to other models like Decision Tree (99.72%), Linear Regression (81.15%), and Support Vector Regressor (99.47%).

1. CONCLUSION

We have demonstrated the successful training of a model which will predict the type and percentage of hearing loss. The model was trained on a large data set, consisting of different hearing frequencies and percentage of disability of both ears.

The results of our evaluations indicate that the model is able to predict the type of hearing loss and the percentage of hearing loss quite accurately.

1. REFERENCES
2. G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
3. W.-K. Chen, *Linear Networks and Systems* (Book style)*.* Belmont, CA: Wadsworth, 1993, pp. 123–135.
4. H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.
5. B. Smith, “An approach to graphs of linear forms (Unpublished work style),” unpublished.
6. E. H. Miller, “A note on reflector arrays (Periodical style—Accepted for publication),” *IEEE Trans. Antennas Propagat.*, to be published.
7. J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.
8. <https://www.mdpi.com/2071-1050/13/10/5406#sec4dot1-sustainability-13-05406>
9. <https://pubs.asha.org/doi/full/10.1044/2022_AJA-21-00194#d1e586>
10. <https://www.nature.com/articles/s41598-020-64175-1>
11. <https://www.nature.com/articles/s41598-019-40192-7>
12. https://link.springer.com/article/10.1007/s12559-018-9607-4