M. Mohanedova *et al*, International Journal of Advances in Agricultural Science & Technology,
Vol.11 Issue.4, April 2024, pg. 10-17

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

An efficient Machine Learning Techniques for Early Detection of Hearing Loss

Dr.Manjula Devarakonda Venkata [1] , Vudathu Venkata Naga Anjana Pravallika[2], Baki Leela Kowshik Reddy[3], Akasam Amulya Sai[4], Marri Sujana[5], Polamarasetti Prem Kumar[6] , [1]Associate Professor , [2,3,4,5, 6]B.tech Students
Department of Computer Science Engineering, Pragati Engineering College, Surampalem , Andhra Pradesh, India
Email: dv.manjula@pragati.ac.in

Abstract:

By 2050, over 700 million people will have severe hearing loss. Audiologists and otolaryngologists are in short supply in underdeveloped and emerging countries, where a considerable part of the population suffers from incapacitating hearing loss. Most hearing impairments are untreated for long periods of time due to a scarcity of specialists. In this study, we present automated hearing impairment diagnosis software based on machine learning to help audiologists and otolaryngologists consistently and effectively identify and classify hearing loss.We discuss the architecture, implementation, and performance evaluation of the two-module automated program for diagnosing hearing impairments: a machine learning model and a module for creating hearing test data. To train and evaluate the machine learning model, the Data Acquisition Module generates a sizable and comprehensive dataset. The kind, degree, and arrangement of hearing loss can be accurately predicted by the model in real time using multiple classes and multi-label classification algorithms that learn from hearing test data.With a log loss reduction rate of 98.48%, a prediction time of 634 ms, and macro and micro precisions of 100%, our proposed machine learning model shows promise and can help audiologists and otolaryngologists quickly and accurately classify the type, degree, and configuration of hearing loss.

Key Words: Audiometry, hearing impairment, machine learning, multiclass classification, multi-label classification

I. Introduction

The World Health Organization (WHO) estimates that by2050[11], approximately 2.5 billion people will suffer from hearingimpairment [1]. Of these individuals, 700 million willexperience disabling hearing loss—defined as an inability toperceive sound lower than 35 decibels (dB) [2]. Incapacitatinghearing loss diminishes the quality of life, learning opportunities,and chances for employment.Approximately 80% of people with disabling hearing impairmentlive in low- or middle-income countries [2]. Thesenations have insufficient hearing care infrastructure as part oftheir national health care systems. According to the WHO,78% of low-income countries possess fewer than one otolaryngologister million, and 93% have fewer than one audiologistper million [1]. This disparity between the hearing-impairedpatients and the health care infrastructure burdens the latter,which is unable to meet the demand of the patients. By facilitating the precise, quick, and effective detection of hearing impairment, automatic hearing diagnosis software that is being developed will help the overworked health care system. The second author's earlier open-source Audiometry application [3]–[8] is enhanced by the suggested software. The application Audiometry has the ability to save,process, and visualize data for tuning fork tests includingWeber, Teal, Schwabach, Rinne, Gelle, and absolute boneconduction; speech audiometry; biothermal caloric test, pure-tone audiometry (PTA), impedance audiometry, and advanced tests such as Stenger, tone decay, short increment sensitive index, and alternate binaural loudness balancing [9]–[11]. It does not, however, carry out automated tests to support the diagnosis of hearing impairment. This functionality will be added to the

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

application via the proposed software. The design, deployment, and outcome analysis of an automatic hearing loss diagnosis software are covered in this study.The machine learning model and data generation are the two elements that make up the software. The Machine Learning Model is trained and tested using[12-14] PTA hearing test data generated by the Data Generation Module. The suggested software will help audiologists and otolaryngologists diagnose hearing impairment by improving the diagnosis's timeliness, accuracy, and efficiency. The structure of the paper is as follows: The specifications for the suggested software are listed in Section II. The Data Generation Module for the hearing test, the Machine Learning Model, and the overall software architecture are covered in Section III. Section IV presents the findings of the suggested software's performance analysis. Section V provides instructions for accessing,executing, and extending the proposed software. Finally, weconclude the paper in Section VI with possible future researchdirections.

II. LITERATURESURVEY

"A model-view-view model (MVVM) application framework for hearing impairment diagnosis - Design and features," Around 466 million people worldwide (more than 5% of the population) suffer from debilitating hearing loss, 34 million of whom are children. According to predictions, by 2050, Globally, in excess of 900 million individuals would experience severe hearing loss. Untreated hearing loss costs the world economy $750 billion annually. Early identification of hearing loss can save a lot of money and lessen the burden on an individual's life. The design and features of a free software framework for diagnosing hearing impairments are covered in this article, which is the first of three in the series. The Models-View-View Model (MVVM) architecture, upon which the framework is based, divides the creation of a graphical user interface (GUI) from the back-end and business logic. Among the benefits of the MVVM paradigm are its reusable components, independent construction of GUI[13] and business or back-end logic, the flexibility to alter GUI without affecting business or back-end logic, ease of unit testing, and decreased maintenance overhead. The proposed architecture, together with the open-source code, makes it simple to extend the application's capabilities, allowing other academics and practitioners to develop their own hearing loss diagnosis tools. An otolaryngologist evaluated the proposed software and discovered it to be very effective in assisting a clinician in making a more methodical, rapid, and accurate diagnosis of hearing impairment.

"Classification of hearing loss," A common issue that everyone encounters occasionally is hearing loss. The temptation to pop them open to improve hearing usually arises when a full sensation develops in the ears, which usually happens while flying or climbing a mountain. Your hearing may also be affected by an ear infection. Most of these hearing loss causes are temporary. On the other hand, most people have lifelong sensorineural hearing loss as a result of aging to varying degrees.

"Multiclass-multilabel classification with more classes than examples," We examine multiclass-multilabel classification scenarios when there are an enormous number of possible labels. The majority of current multiclass-multilabel learning algorithms break down if they only receive a small number of samples with a particular label and instead assume a big sample size from each class. We suggest and assess the subsequent two-phase method: First, build an initial classifier using an arbitrary (perhaps heuristic) classification algorithm. Next, use a straightforward but ethical technique to improve this classifier by eliminating undesirable labels from its output. A thorough theoretical analysis enables us to validate our method in a set of reasonable circumstances. even if the training set does not offer a statistically accurate representation of the majority of classes (such as label sparsity and a power-law distribution of label frequencies). Interestingly, our theoretical method still works when there are more classes than there are in the sample. We use the 1.5 million categories defined on Wikipedia to categorize the entire web to demonstrate the effectiveness of our approach.

"Based on the maximum entropy principle, queueing theory estimates the performance distribution. It is common practice in related research on queuing systems to make the assumptions that the client arrival and service ratio distributions are well characterized and that the system is stable. This study treats the queuing system as a black box, assuming simply that the queueing system is stable and making no assumptions about the distribution of arrival and service ratios. Applying the maximum entropy principle to various commonly

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

obtainable metrics, such as the system's capacity, the average number of clients, and the average server utilization, yields the performance distribution of queuing systems. Performance distributions are calculated and some uncommon instances are simulated. The goodness of fit test using chi-square.

III.SYSTEM ANALYSIS

A. EXISTING SYSTEM

Existing approaches for identifying hearing loss in newborns, infants, and toddlers mostly include screening tests, diagnostic evaluations, and clinical assessments conducted by healthcare professionals. The otoacoustic emissions (OAE) test is a popular screening method that analyses the sound waves generated by the inner ear in response to auditory stimulation. The Auditory Brainstem Response (ABR) test measures electrical activity in the auditory nerves and brainstem in response to sound. These screening tests are usually done shortly after delivery or during routine paediatriccheckups. If a screening test indicates a probable hearing loss, further diagnostic procedures, such as comprehensive audiologic exams and medical examinations, may be required to confirm the diagnosis and assess the severity and type of hearing loss. Speech-language pathologists and paediatric audiologists also conduct clinical assessments to assess infants' and toddlers' auditory and communication abilities. Even if the methods used now to detect hearing loss are successful, they may have drawbacks, such as the need for specialized equipment and trained workers, restricted accessibility in some areas, and difficulties in identifying mild or progressive hearing loss at an early stage. The detection of hearing loss in newborns, infants, and toddlers may be made more accurate, efficient, and accessible by incorporating machine learning techniques into current systems.

DISADVANTAGES OF THE EXISTING SYSTEM

1. **Data Quality and Diversity**: The system's ability to generalize to a broad range of hearing impairments may be limited by the absence of comprehensive and diverse datasets for training the machine learning model.
2. **Hardware and Software Constraints:** The hardware and software resources of the system can have an impact on how well it performs. The system's capabilities can be limited by the hardware that can be deployed if it requires a significant amount of processing power.
3. **Interpretability of Results:** Otolaryngologists and audiologists may find it difficult to comprehend and totally accept the diagnosis if the machine learning model's decision-making process is not comprehensible.
4. **Real-world Variability:** The system's accuracy may be influenced by external variables and changes in the physical environment. This includes differences in testing conditions, background noise, and patient compliance.
5. **Ethical and legal issues:** The use of automated diagnosis systems raises ethical and legal concerns, such as those related to patient privacy, data security, and the role of medical personnel in decision making.
6. **Restricted Focus or Scope:** The system may be designed to target specific types or intensities of hearing loss, potentially limiting its application to a broader range of situations.
7. **Updating and Maintaining Challenges:** Consistent updates and maintenance are required to maintain the system current and proper throughout time. If this is not done, performance may degrade as more data becomes available.
8. **User Acceptance:** Patients and healthcare providers must accept automated technologies. Opposition to or concern about the system's trustworthiness may have an impact on its acceptance and efficacy.

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

B. PROPOSED SYSTEM

The Hearing Test Data Generation Module substantially facilitates the development of a comprehensive and diverse dataset required for the machine learning model's training and assessment. This module ensures that the model is exposed to a wide range of scenarios and circumstances, hence increasing its potential for successful generalization.

The Machine Learning Model analyses the hearing test data with powerful multiclass and multilabel classification techniques. This enables the model to reliably predict the presence of hearing loss and identify its kind, severity, and configuration.The recommended model produces promising results, with a log-loss reduction rate of 98.48%, macro and micro precision of 100%, and a forecast time of 634 ms.

The proposed system aims to aid otolaryngologists and audiologists in clinical decision-making by providing precise and effective automated diagnosis capabilities. The system's ability to swiftly and effectively define the characteristics of hearing loss may result in a significant reduction in the time necessary to diagnose impairments, particularly in locations where access to specialist healthcare providers is limited. This addresses a large gap in healthcare services and improves overall public health outcomes related to hearing impairments. Furthermore, the system's modular architecture enables scalability, adaptation, and interaction with emerging technologies, ensuring its continued efficacy in the ever-changing healthcare business.

IV.SYSTEM DESIGN

SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture.

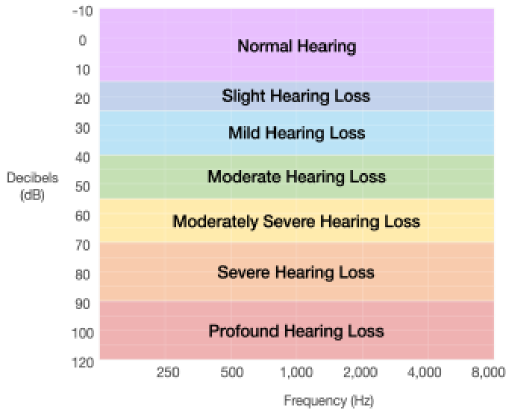

Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

MODULES

Module for Generating Hearing Test Data: This module is in charge of creating a large dataset that will be used to evaluate and train machine learning models. It ensures that the model is exposed to a wide range of settings and circumstances, allowing it to generalize effectively. The obtained data serves as a solid foundation for the machine learning process by embracing a wide range of types and intensities of hearing impairments.

The Machine Learning Model: which employs complex multiclass and multi-label classification algorithms, is the system's primary component. After being trained on a range of datasets generated by the Data Generation

M. Mohanedova *et al*, International Journal of Advances in Agricultural Science & Technology,
Vol.11 Issue.4, April 2024, pg. 10-17

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

Module, this model can quickly and precisely predict the presence of hearing loss, as well as the kind, degree, and configuration of the condition. To offer reliable diagnostic information, the model must be able to understand and interpret complex patterns in the data.

User Interface Module:

This module serves as a liaison between the automated system and otolaryngologists, audiologists, and other medical professionals. It provides an easy-to-use interface for interfacing with the system, entering patient data, and receiving diagnostic results. A user-friendly interface that offers information increases the system's overall acceptability and usability among healthcare professionals.

Feedback and Improvement Module:

A feedback and improvement module might be used to guarantee ongoing system performance improvement. Healthcare practitioners' opinions on the system's overall efficacy and diagnosis accuracy are gathered in this module. Iterative improvements are made possible by this feedback loop, which helps the system adjust to new obstacles and changing healthcare needs.

Security and Privacy Module:

Given the sensitivity of healthcare data, a separate security and privacy module is required. This module implements strong security methods to protect patient data while ensuring adherence to ethical norms and data protection regulations. It includes auditing tools, access controls, and encryption mechanisms to ensure the security and integrity of the diagnostic data that the system processes.

VI .RESULTS AND DISCUSSION

The experiment's setup consists of a Python notebook with an HDD greater than 100 GB and 12 GB of RAM that runs on a web-based Google Collab site. This experiment does not use an active GPU. For this experiment, a pre-trained Google word2vec model with 300 dimensions and a vocabulary of around 100 billion words is used. The corpus was divided into test and training data segments at a ratio of 8:2. The machine learning model was trained on train data, which was also used to calculate the starting scores for the score prediction modules. The machine learning model was then updated by feeding system testing data one at a time. Cosine similarity and word mover's distance are utilized in tandem with a Multinomial Naive Bayes model to produce the findings. Both the approaches with and without the model produced results in under a minute at Google Collab. The results are as follows.

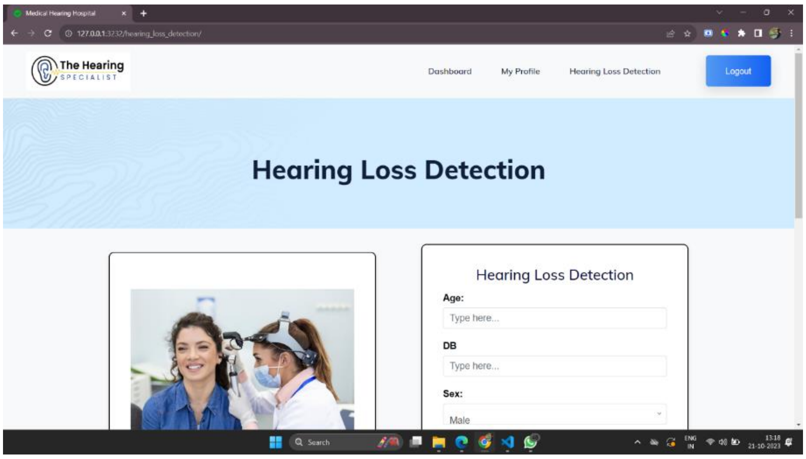

Fig 2. Input Data For Finding Hearing Loss

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

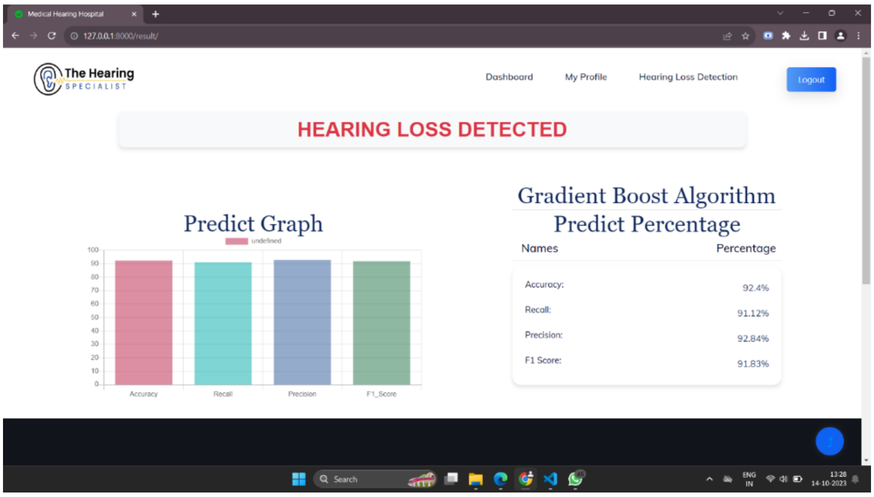

Fig 3. Predecting Hearing Loss

VII. CONCLUSION AD FUTURE WORK

There are two contributions to this project. First, we created a hearing loss data generation module. Using an 8GB RAM laptop with a 2.2 GHz CPU, our module generates 17,500 synthetic PTA hearing test data instances in less than a second. It makes sure the information satisfies the requirements for categorizing hearing loss. The source code for the Data Generation Module can be used by other researchers to provide data for their own automated models of hearing diagnosis. Second, in order to create a machine learning model, we used multiclass and multi-label classification techniques. The data used by the machine learning model for testing and training is produced by the data generation module.The normalization technique used by the Data Generate Module allowed the Machine Learning Model to achieve 100% macro and micro precision. Our trained model is 98.48% more likely to make the right forecast than a random guess, with a log-loss reduction rate of 98.48%.

REFERENCES :

[1] WHO. (2021) WHO: 1 in 4 people projected to have hearing problems by 2050. [Online]. Available: https://www.who.int/news/item/02-03- 2021-who-1-in-4-people-projected-to-have-hearing-problems-by-2050

[2]Deafness and hearing loss. [Online]. Available: https: //www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[3] W. Sheikh and N. Sheikh, "A model-view-view model (MVVM) application framework for hearing impairment diagnosis - Design and features," in 2020 Intermountain Engineering, Technology and Computing (IETC), 2020, pp. 1–6. [Online]. Available: 10.1109/ IETC47856.2020.9249124

[4]"A model-view-view model (MVVM) application framework for hearing impairment diagnosis - Type dependency architecture," in 2020 Intermountain Engineering, Technology and Computing (IETC), 2020, pp. 1–6. [Online]. Available: 10.1109/IETC47856.2020.9249181

[5] "A model-view-view model (MVVM) application framework for hearing impairment diagnosis - Class inheritance architecture," in 2020 Intermountain Engineering, Technology and Computing (IETC), 2020, pp. 1–6. [Online]. Available: 10.1109/IETC47856.2020.9249215

[6]"A model-view-view model (MVVM) application framework for hearing impairment diagnosis," Journal of Open Source Software, vol. 5, no. 51, 2016. [Online]. Available: https://doi.org/10.21105/joss.02016

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

[7] W. Sheikh. Audiometry application.[Online]. Available: https: //sourceforge.net/projects/audiometry-hearing-loss/

[8] drwaseemsheikh Audiometry GitHub repository. [Online]. Available: https://github.com/ drwaseemsheikh/ audiometry

[9] F. H. Bess and L. E. Humes, Audiology: The Fundamentals, 4th ed. Lippincott Williams & Wilkins, 2020.

[10] Manjula Devarakonda Venkata**,** Sumalatha Lingamgunta," Triple-Modality Breast Cancer Diagnosis and Analysis in middle aged women by Logistic Regression", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-4, February 2019

[11] Manjula Devarakonda Venkata**,** Sumalatha Lingamgunta, "Statistical and Texture Features based Breast Cancer Classification from Multi image Modalities using Support Vector Machines with Gaussian Radial Basis Function", International Journal of Control and Automation Vol. 13, No.1, (2020), pp. 384 - 396

[12] Shaik Mohammad Rafee, Manjula Devarakonda Venkata," An Ensemble Neuro Fuzzy Algorithm For Breast Cancer Detection And Classification", Ictact Journal On Soft Computing, January 2024, Volume: 14, Issue: 03, DOI: 10.21917/ijsc.2024.0460

[13] Manjula Devarakonda Venkata**,** Sumalatha Lingamgunta Health Care Automation in Compliance to Industry 4.0 Standards :A Case Study of Liver disease Prediction", Journal of Scientific & Industrial Research Vol. 82, February 2023, pp. 263-268 DOI: 10.56042/jsir.v82i2.70215

[14] S. Kramer and D. K. Brown, Audiology: Science to Practice, 3rd ed. Plural Publishing, 2019.

[15] S. I. Nikolenko, "Synthetic data for deep learning," p. 3, 2019. [Online]. Available: https://arxiv.org /abs/1909.11512

[16] D. Rankin, M. Black, R. Bond, J. Wallace, M. Mulvenna, and G. Epelde, "Reliability of supervised machine learning using synthetic data in health care: Model to preserve privacy for data sharing," JMIR Medical Informatics, vol. 8, no. 7, p. e18910, 2020. [Online]. Available: https://doi.org/10.2196/18910

[17] J. J. Walker, L. M. Cleveland, J. L. Davis, and J. S. Seales, "Audiometry screening and interpretation," American Family Physician, vol. 87, no. 1, pp. 41–47, 2013. [Online]. Available: https: //www.aafp.org/afp/2013/0101/p41.html

[18] W. B. Alshuaib, J. M. Al-Kandari, and S. M. Hasan, "Classification of hearing loss," in Update on Hearing Loss, F. Bahmad, Ed. IntechOpen, 2015. [Online]. Available: https://www.intechopen.com/books/updateon-hearing-loss/classification-of-hearing-loss

[19] O. Dekel and O. Shamir, "Multiclass-multilabel classification with more classes than examples," in AISTATS, 2010.

[20] Microsoft. What is ML.NET? [Online]. Available: https://dotnet. microsoft.com/learn/ml-dotnet/what-is-mldotnet

[21] A. Dubey, O. Gupta, R. Raskar, and N. Naik, "Maximum entropy fine-grained classification," CoRR, vol. abs/1809.05934, 2018. [Online]. Available: http://arxiv.org/abs/1809.05934

ISSN: 2348-1358

Impact Factor: 6.901

NAAS Rating: 3.77

[22] D. He, R. Li, Q. Huang, and P. Lei, "Maximum entropy principle based estimation of performance distribution in queueing theory," PLoS ONE, vol. 9, no. 9, p. e106965, 2014.[Online]. Available: https://doi.org/10.1371/journal.pone.0106965

[23] A. Bhandari. (2020, Apr.) Everything you should know about confusion matrix for machine learning. [Online]. Available: https://www. analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/