e-ISSN:2582-7219

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 6, June 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

# Gender Recognition Using Voice

**Chidrewar Adhithi, Namsani Chandana, Biradar Nikita, Dr.D.Shravani**

B.E, Department of CME, SCETW, OU, Hyderabad, Telangana, India

B.E, Department of CME, SCETW, OU, Hyderabad, Telangana, India

B.E, Department of CME, SCETW, OU, Hyderabad, Telangana, India

Associate Professor, ADCE, SCETW, OU, Hyderabad, Telangana, India

**ABSTRACT:** The field of voice-based demographic analysis, particularly in gender and age classification, is increasingly significant in the realms of digital interaction and personalized services. Despite substantial advancements in machine learning and deep learning, the performance disparity among these methodologies remains a pressing issue. This research scrutinizes the efficacy and efficiency of various approaches, primarily focusing on the comparative analysis of gender and age classification through the application of XGBoost and neural networks. The challenges highlighted involve computational demands, scalability, and practical deployment hurdles, which are critical in determining the feasibility of these methods in real-world applications. The study leverages a comprehensive evaluation framework to assess the performance of these classifiers in terms of accuracy and computational efficiency, using a well-curated voice dataset to ensure robust analysis.

The primary objective of this study is to compare two distinct machine learning approaches—XGBoost and neural networks—in their ability to classify gender and age from speech samples accurately and efficiently. By analyzing a range of metrics, including accuracy, precision, recall, and F1-scores, this research aims to unearth which method performs better under varying conditions and why. Particular attention is paid to the adaptability of these methods in handling large and diverse datasets typical of real-world settings. The analysis extends to evaluating the computational requirements and scalability of each approach, offering insights into their practical applications and limitations. The results indicate that while XGBoost provides a robust method for individual classification tasks, neural networks excel in handling combined gender and age detection, showcasing superior performance in terms of both accuracy and efficiency.

This study contributes significantly to the understanding of machine learning applications in voice-based demographic analysis, highlighting the strengths and weaknesses of XGBoost and neural networks. These insights not only foster a deeper understanding of each methodology's technical and practical aspects but also guide future advancements in the field. The study's findings are poised to influence subsequent research directions, encouraging the development of more advanced, efficient, and scalable solutions for gender and age classification. This exploration ultimately serves to bridge the gap between theoretical research and practical implementation, providing valuable recommendations for enhancing existing technologies and developing new approaches in the rapidly evolving field of voice-based demographic analysis.

## I.INTRODUCTION

Gender and Age Detection Using Deep Learning embarks on an exploration of the cutting-edge methodologies that leverage deep learning techniques to accurately identify a person's gender and age from various data sources such as images, videos, and audio recordings. This endeavor is rooted in the interdisciplinary fields of computer vision, speech analysis, and artificial intelligence, showcasing the remarkable capabilities of deep learning algorithms in interpreting complex patterns that are often imperceptible to the human eye or ear.

The need for gender and age detection applications stems from their vast potential to enhance personalized user experiences and improve demographic analytics. In a world where digital interaction is becoming increasingly personalized, the ability to automatically determine a user's gender and age can significantly tailor content, recommendations, and advertisements to suit individual preferences. Moreover, these applications serve crucial roles in enhancing security systems, where they can be used to strengthen access control measures based on age and gender-specific criteria.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

Real-time applications of gender and age detection span a multitude of sectors, including but not limited to, retail, entertainment, healthcare, and security. In retail, for instance, digital signage and kiosks can dynamically alter displayed content to match the perceived demographic characteristics of the viewer, thereby increasing engagement and potential sales. In the realm of entertainment, streaming services can offer more accurate content recommendations, enriching user experience and satisfaction.

In healthcare, gender and age detection technologies can streamline patient triage by automating the initial assessment processes, thereby allowing for a more efficient allocation of medical resources. Furthermore, these technologies can contribute to elderly care, offering non-invasive monitoring solutions to ensure the safety and well-being of seniors, especially those living independently.

Security systems are also beneficiaries of gender and age detection technologies. These systems can be augmented to include demographic-specific access controls, enhance surveillance operations by providing detailed demographic analytics in real-time, and improve the accuracy of crowd monitoring tools used in public safety initiatives.

The societal benefits of gender and age detection applications are multifaceted. By enabling more personalized interactions, these technologies can improve user engagement across digital platforms, leading to enhanced satisfaction and loyalty. Moreover, by contributing to more accurate demographic analytics, they aid businesses and organizations in making data-driven decisions that are more closely aligned with the needs and preferences of their target audiences.

However, the development and deployment of these applications are not without challenges. Concerns regarding privacy, ethics, and the potential for bias necessitate rigorous standards and transparent practices. Ensuring that gender and age detection technologies are developed with a commitment to fairness, accuracy, and respect for individual privacy is paramount.

Deep learning, as the backbone of these applications, offers sophisticated models that can learn from vast amounts of data, continuously improving their accuracy and reliability. These models, including convolutional neural networks (CNNs) for image-based detection and recurrent neural networks (RNNs) for audio analysis, are at the forefront of technological advancements in the field.

Despite the potential for bias and privacy concerns, the benefits provided by gender and age detection technologies, when ethically implemented, can significantly outweigh the challenges. They offer a path toward more secure, personalized, and efficient systems that can adapt to the needs of diverse user groups.

The application on gender and age detection using deep learning marks a significant step forward in harnessing the power of AI to understand and interact with the world in more nuanced and sophisticated ways. As technology continues to evolve, the potential applications and benefits of these systems are bound to expand, paving the way for innovations that could redefine user experiences and demographic analytics across various sectors.

## II. LITERATURE REVIEW

In recent studies, various approaches have been explored for age and gender classification from speech data using deep learning techniques. Qawaqneh et al. proposed a joint age and gender classification model leveraging deep neural networks, achieving high accuracy by incorporating multimodal information. Sánchez-Hevia et al. introduced temporal convolutional neural networks to enhance temporal feature processing, leading to improved demographic predictions, albeit with increased model complexity. Kumbhar et al. developed a versatile deep learning-based gender and age detection system capable of capturing subtle patterns, offering adaptability to diverse datasets but requiring substantial data and careful tuning to avoid overfitting. Yasmin et al. combined rough set theory with deep learning, integrating traditional computational techniques with modern neural architectures to potentially enhance prediction accuracy, though balancing between the two approaches may necessitate intricate tuning. These studies collectively illustrate the evolving landscape of age and gender classification research, highlighting the strengths, challenges, and trade-offs associated with different methodologies.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

## III.METHODOLOGY

The systematic approach to gender and age detection using voice data, emphasising the need for integrated methods to leverage correlations between these attributes. Traditional approaches treat gender and age detection separately, employing machine learning algorithms and feature extraction techniques. However, these methods suffer from limitations such as computational inefficiencies and biased predictions due to limited dataset diversity. To address these challenges, there is a growing interest in developing integrated models capable of detecting both gender and age simultaneously. Such models would enhance prediction accuracy and applicability across various domains, including interactive technologies, security, telemedicine, and education. By overcoming current limitations, integrated approaches offer promising opportunities to improve demographic detection in real-world scenarios.

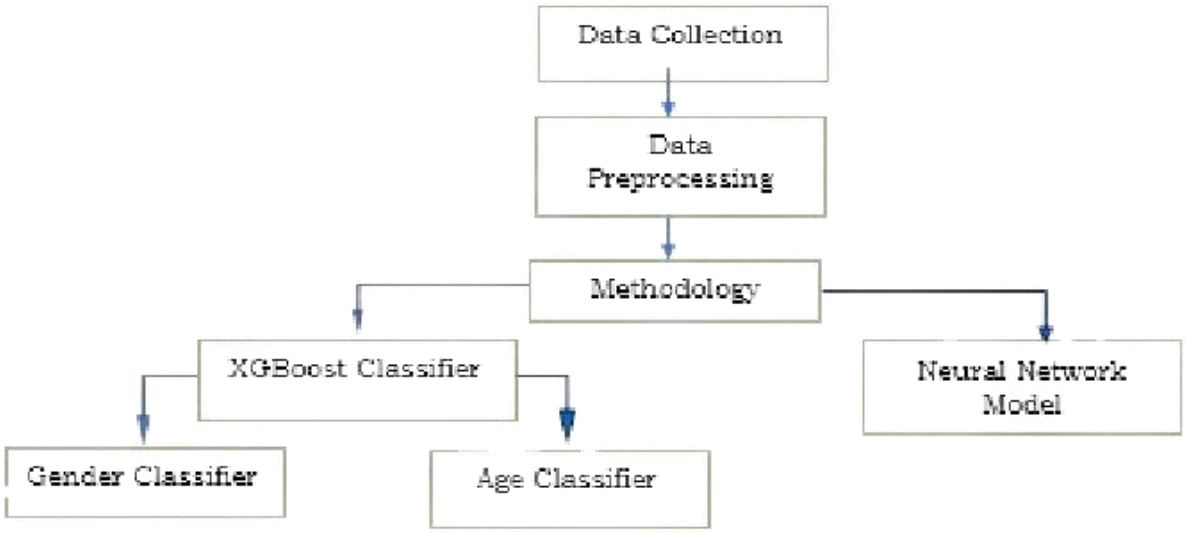


**Figure 3.1: Flowchart of the Proposed Algorithm**

**PROPOSED SYSTEM:**

**1. XGBoost Algorithm:**

The XGBoost model provides a robust solution for detecting both gender and age (gender) and regression (age), XGBoost efficiently captures subtle patterns associated with each attribute. Leveraging its capabilities in classification and regression tasks, XGBoost enables accurate predictions by utilizing features such as pitch, tone, speech patterns, and speech rate. This approach offers flexibility in dataset integration and parameter optimization, ensuring optimal performance across diverse demographic groups and scenarios. Overall, XGBoost presents a versatile and efficient solution with wide-ranging applications in various fields**.**

**Algorithm for Gender Detection Using XGBoost:**

For gender detection using XGBoost, the algorithm involves several key steps. Initially, audio features such as pitch, frequency, and energy levels are extracted from voice samples. These features form the input variables for the model. The process includes:

- **Data Preparation:** Voice recordings are pre-processed to extract meaningful acoustic features which may include MFCCs (Mel Frequency Cepstral Coefficients), pitch, and tone.
- **Feature Selection:** Relevant features that significantly contribute to gender differences in voice are selected to reduce model complexity and improve performance.
- **Model Training:** The selected features are fed into the XGBoost classifier to train the model on a labeled dataset, where labels denote the gender associated with each voice sample.
- **Model Optimization:** Parameters of the XGBoost model, such as learning rate, number of trees, and depth of trees, are tuned using techniques like cross-validation to enhance the model's accuracy.
- **Prediction and Validation:** The trained model is then used to predict gender on unseen data, and its performance is validated using metrics such as accuracy, precision, and recall.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

**Algorithm for Age Detection Using XGBoost:**
The algorithm for age detection is similar in structure to gender detection but tailored to address the regression nature of predicting age, which is a continuous variable:

- **Data Preparation:** Extract relevant features from voice data that may include intonation, speed, and speech rhythm, alongside standard acoustic features.
- **Feature Engineering:** Develop new features or transform existing features to better capture age-related changes in voice, which might involve statistical summaries of tone or dynamic changes in pitch across a speech sample.
- **Model Training:** Use these features to train the XGBoost model, configuring the algorithm to optimize for regression tasks, targeting the actual ages as regression targets.
- **Hyperparameter Tuning:** Adjust the model's complexity and learning capacity by tuning hyperparameters to minimize prediction errors and avoid overfitting.
- **Evaluation:** Assess the model's effectiveness in predicting age using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), which provide insights into the accuracy of age estimations

**2. Neural Network:**
The neural network model is designed to simultaneously detect gender and age from voice data using a multi-task learning framework. By incorporating shared layers for initial feature extraction, followed by separate branches for gender classification (utilizing softmax activation) and age prediction (employing linear activation), the model efficiently captures relevant patterns for both attributes. Techniques like batch processing, dropout, and regularization are integrated to prevent overfitting and enhance generalization, ensuring robust performance across diverse datasets and scenarios. This approach not only streamlines the learning process but also allows the network to leverage shared information between tasks, improving overall prediction accuracy. With its adaptable architecture and advanced training techniques, the neural network model offers a promising solution for gender and age detection in various applications, including healthcare, marketing, and security.

**Algorithm for Age and Gender Detection Using Neural Networks:**
A neural network designed to detect both gender and age simultaneously typically uses a multi-task learning framework, where the network learns to perform both tasks together, benefiting from shared feature representations. The architecture might involve:

- **Input Layer:** Audio data is pre-processed, and relevant features are fed into the neural network.
- **Hidden Layers:** Multiple layers, potentially including convolutional and recurrent layers, extract and learn intricate patterns in the data, with shared representations that are useful for both gender and age detection.
- **Output Layer:** The network splits into two branches at the final stages—one for classifying gender (using softmax activation) and another for predicting age (using linear activation).
- **Loss Function:** The loss function is combined, typically adding the losses from the gender classification and age regression, allowing the model to update its weights to minimize both prediction errors simultaneously.
- **Training and Validation:** The model is trained on a diverse dataset and validated using a separate validation set to ensure that it generalizes well to new, unseen data.

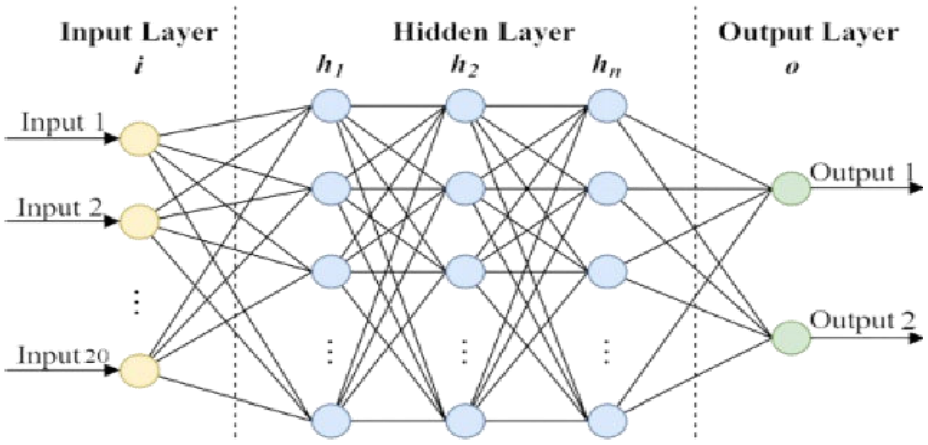


**Figure 3.2: Neural Network Architecture**

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

## IV.EXPERIMENTAL RESULTS AND DISCUSSION

In the evaluation of machine learning models for gender and age detection from speech samples using the Mozilla dataset, findings indicate that while XGBoost demonstrates solid performance, a neural network model outperforms it in both gender and age detection tasks. This neural network architecture seamlessly integrates gender and age detection, leveraging deep learning techniques for enhanced feature extraction. Notably, the dual-output system of neural networks significantly boosts precision and recall rates for both gender and age predictions. This comparative analysis underscores the superiority of neural networks over traditional models, signaling a promising direction for advancements in speech analysis. With their ability to effectively handle complex patterns and interdependencies between gender and age attributes, neural networks offer a compelling solution for accurate and efficient demographic detection from voice data as shown in Figure 4.1 and Figure 4.2.

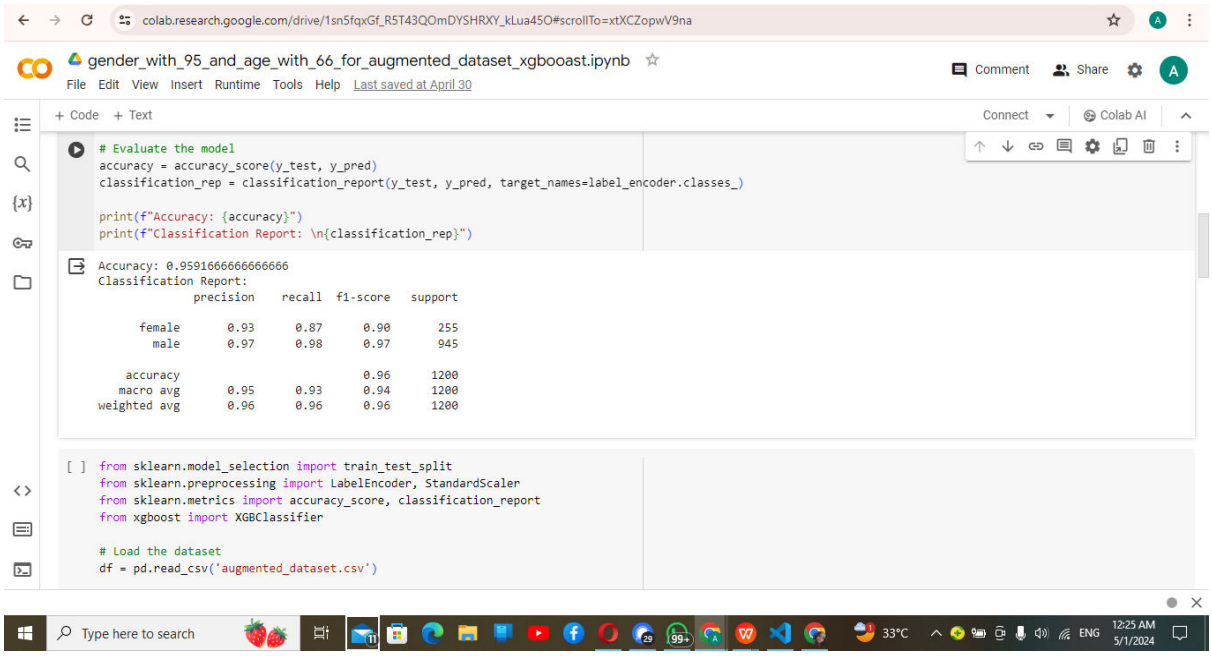


**Figure 4.1**

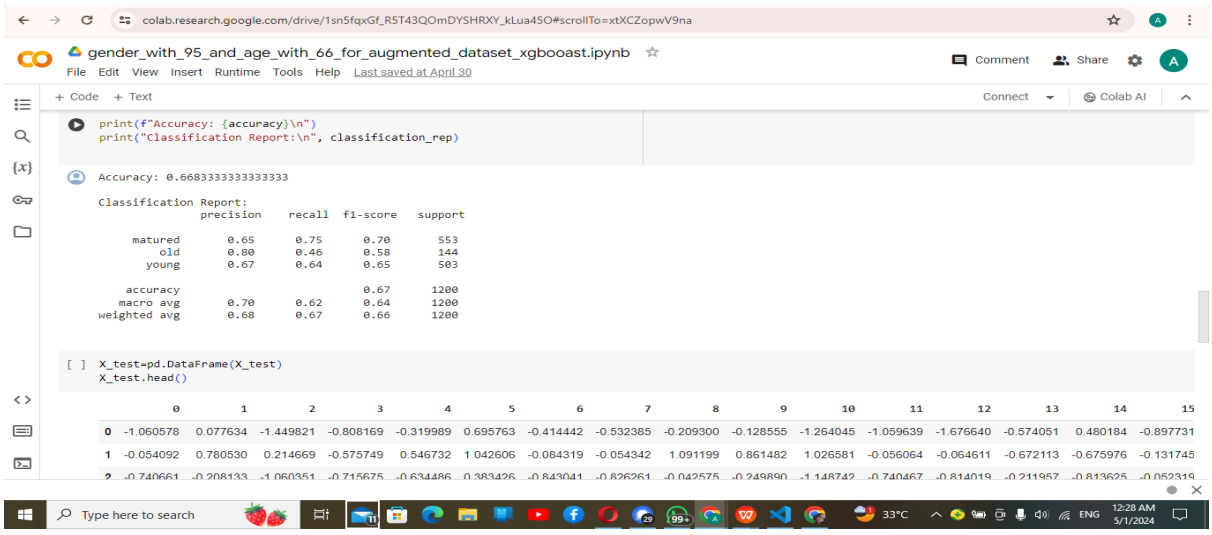


**Figure 4.2**

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 6, June 2024 |

| DOI:10.15680/IJMRSET.2024.0706009 |

## V.CONCLUSION

In conclusion, this study explores the innovative application of deep learning in analyzing speech samples to detect demographic characteristics. Leveraging the Mozilla audio dataset, the research focuses on comparing the effectiveness of XGBoost and neural networks in gender and age detection. While XGBoost serves as a robust baseline, neural networks exhibit superior capabilities in extracting intricate speech patterns. The model developed demonstrates high accuracy in predicting gender and age, offering promising implications for user experience enhancement in technology interfaces, personalized marketing, and security applications.

## REFERENCES

1. Uddin, Mohammad Amaz, "Gender and Region Detection from Human Voice Using the Three-Layer Feature Extraction Method with 1D CNN," Journal of Information and Telecommunication, Vol. 6, Issue 1, 2022.
2. Jaleel, Hanan Qassim, Jane Jaleel Stephan, and Sinan A. Naji, "Gender Identification from Speech Recognition Using Machine Learning Techniques and Convolutional Neural Networks," Webology, Vol. 19, Issue 1, pp. 1666-1688, 2022.
3. Gupta, Yashi, "Gender and Age Recognition Using Audio Data—Artificial Neural Networks," Soft Computing for Security Applications: Proceedings of ICSCS, Springer Singapore, 2022.
4. [Shagi, G. U., and S. Aji, "A Machine Learning Approach for Gender Identification Using Statistical Features of Pitch in Speeches," Applied Acoustics, Vol. 185, 2022.
5. Sánchez-Hevia, Héctor A., "Age Group Classification and Gender Recognition from Speech with Temporal Convolutional Neural Networks," Multimedia Tools and Applications, Vol. 81, Issue 3, pp. 3535-3552, 2022.
6. Alnuaim, Abeer Ali, "Speaker Gender Recognition Based on Deep Neural Networks and ResNet50," Wireless Communications and Mobile Computing, 2022, pp. 1-13.
7. Zaman, Syed Rohit, "One Source to Detect Them All: Gender, Age, and Emotion Detection from Voice," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE, 2021.
8. Kwasny, Damian, and Daria Hemmerling, "Gender and Age Estimation Methods Based on Speech Using Deep Neural Networks," Sensors, Vol. 21, Issue 14, pp. 4785, 2021.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com