



An AI-Powered Gamified Framework For Language Learning And Speech Disorder Rehabilitation

1Anciya Ebrahim, 2Dr. Rosna .P.Haroon

1Student, 2Professor in Computer Science And Engineering

1Ilahia College of Engineering And Technology,

2Ilahia College of Engineering and Technology

Abstract- This project offers a comprehensive system that integrates AI-driven speech disorder diagnosis and treatment with gamified language learning. With its entertaining games, adaptive material, and real-time feedback, it improves language learning and speech rehabilitation for young learners and people with speech impairments. Through interactive rewards, the language module improves conversational skills, grammar, and vocabulary. The speech module provides individualized exercises and uses machine learning algorithms and voice-to-text APIs to identify conditions like dysarthria and stuttering with 88–92% accuracy. The system, which was developed using Django and features like DeepSpeech and Whisper API, guarantees both therapeutic help and educational value. It is appropriate for both academic and clinical application because to its high user engagement and diagnostic efficacy.

Keywords–Language Learning modules, Real time translation, Voice Signals, Intelligent Systems, Speech Disorder Diagonosis, Machine learning, Speech to Text translation, Support Vector Machines, Speech Features.

I. INTRODUCTION

This project presents an AI-powered application that combines automated speech disorder detection and treatment with gamified language learning. Targeting young students and those with speech impairments, it tackles the shortcomings of conventional language teaching methods and the dearth of easily available speech therapy. The system improves vocabulary, grammar, and pronunciation through gamification features like challenges, prizes, and interactive simulations. With real-time feedback and customized exercises, it simultaneously uses machine learning (SVM, DeepSpeech) and speech APIs (Whisper, DeepSeek) to identify illnesses including stuttering and articulation problems. The platform, which is based on Django, guarantees user engagement and successful rehabilitation while providing a scalable, inclusive solution that combines healthcare and education for better communication skills and wider accessibility.

The output is supplied through the mobile application, which offers users progress tracking, individualized therapy activities, and real-time feedback, after the disorders have been identified. With the help of machine learning algorithms like SVM and DeepSpeech and voice-to-text APIs like Whisper and DeepSeek, this AI-powered smartphone app can identify speech abnormalities like stuttering, dysarthria, phonological

disorders, and cluttering.

In addition to computerized diagnosis, the application incorporates gamified language learning to improve pronunciation, grammar, and vocabulary through interactive games. The application's user-centric design and Django-based backend guarantee accessibility and engagement for a broad user base, including kids and those with speech difficulties.

Through the integration of educational resources and speech therapy, the platform fosters independent learning, inclusive communication, and enhanced learning results.

II. RELATED WORKS

The study by Muqdad Al-Karawi et al.[1] investigates the use of internet gaming as a technique for informal English language learning among children in China, India, Chile, and Hungary who are between the ages of 8 and 14. It demonstrates how gaming fosters social, cognitive, and linguistic development by providing an immersive setting in which kids organically pick up conversational skills, contextual language use, and vocabulary. The study fills up the gaps in previous studies, especially the dearth of perspectives that are culturally varied and child-focused. Results show how gaming can promote intercultural dialogue and cooperative problem-solving in a relaxed environment. The authors advise creating kid-focused, culturally sensitive games and incorporating gaming into English curriculum. Future studies ought to look at long-term impacts, study particular game genres, and take into account larger cultural contexts. The study emphasizes the potential and pedagogical usefulness of gaming in contemporary language acquisition.

Jensen et al.'s[2] study, which was published in IJACSA (2020), looks at how digital gaming helps young language learners (YELLs) in Denmark learn English. It emphasizes extracurricular English (EE) activities and the ways that gaming improves vocabulary growth, language acquisition, and general competency. By examining the precise linguistic input and output in different game genres and their effects on learning outcomes, the study closes a gap in the body of previous literature. It highlights how little research has been done on using video games to help kids learn new words and improve their language skills. The results indicate that gaming can be a useful, unofficial tool for improving English proficiency and urge more study to examine its educational potential and provide guidelines for incorporating gaming into young learners' language acquisition.

Huynh et al.'s [3]study, which was presented at the 2018 ICCOINS, assesses how gamification features like "Badge" and "Winning Streak" affect Duolingo users' motivation and engagement. The study objectively evaluates the usefulness and attractiveness of these characteristics using a game refinement measure, providing information on how particular game mechanics improve language-learning experiences. It fills in a significant study gap in gamification, since conventional techniques are imprecise when assessing distinct game elements. In order to more accurately assess the instructional and entertaining value of gamified aspects, the study suggests a single measuring approach. By enhancing user experience and facilitating a more successful integration of gaming dynamics in language instruction, this method seeks to guide the creation of learning platforms that are more captivating and inspiring.

In a study presented at the 2023 International Conference on Information Technology (ICIT), Mazlan et al. [4]investigate a mobile game-based learning platform intended to encourage teens' interest in Malay proverbs. Through optimal path-finding, the game uses the A* search algorithm to improve gameplay, allowing players to explore the virtual arena, collect clues, and take part in interactive challenges. Through an entertaining, adventure-style framework, this method promotes improved comprehension and memory of Malay proverbs. Particularly in cultural and linguistic contexts, the study fills a gap in the incorporation of sophisticated algorithms into educational games. The study demonstrates how technology may modernize ancient knowledge by combining algorithmic optimization with game-based learning, making it more enticing and approachable for younger audiences while advancing educational and cultural preservation objectives.

The study by Modi et al.[5]., which was presented at ICIDCA-2023, looks at how gamified systems in the fields of healthcare, e-commerce, and education might incorporate AI and ML. It demonstrates how behavior modification, adaptive experiences, and personalized learning are some of the ways AI and ML improve gamification. The study fills a significant knowledge vacuum about the real-world use of these technologies, pointing out that although their promise is well acknowledged, there is still a lack of effective use. The study provides insightful information about the changing role of AI and ML in gamification by examining existing applications and pinpointing areas for development. It seeks to direct the creation of more effective, captivating, and user-focused gamified systems in the future while showcasing the revolutionary potential of AI-driven gamification in a variety of industries.

Rajendran et al.'s[6] study, which was presented at the 2020 IEEE ICALT, examines how schoolchildren who use the Hello English (HE) mobile app to learn English differ based on their gender in terms of motivation, engagement, and interaction behavior. It answers three important questions: does gender affect motivation, does gender affect engagement, and does gender affect interaction behavior? The study identifies a critical void in the literature on gender-specific user experiences in mobile-assisted language learning (MALL). The study offers important insights into how gender affects mobile-based English learning by examining these differences. In order to improve the efficacy of MALL platforms like Hello English in promoting equitable language acquisition, the findings are intended to guide the creation of more inclusive, responsive, and efficient educational technologies that address the varied demands of learners.

White et al.'s[7] study, which was presented at the IEEE AIxVR conference in 2024, presents VRcabulary, a virtual reality (VR) platform that uses a multi-modular framework to improve language acquisition. It addresses the main drawbacks of current VR language programs, including their lack of intermediate practice modules, repetitious material, and unimodular design. Disengagement results from the lack of variation in many of the current systems, such as House of Languages, which only concentrate on education or gamification. By combining various modules that facilitate practice, training, and assessment, VRcabulary solves these problems while encouraging ongoing participation and knowledge reinforcement. The study emphasizes the platform's potential to produce a more dynamic, successful, and learner-centered virtual reality language learning experience, helping to construct immersive and pedagogically sound educational tools of the future.

Ruiz et al.'s[8] work, which was presented at the 2023 IEEE International Conference on Artificial Intelligence, examines the creation and assessment of *Influencer Estadístico*, a Scratch-made educational video game designed to improve first-year Baccalaureate students' understanding of statistics. By providing interactive, context-based learning opportunities, the game seeks to enhance comprehension of fundamental statistical concepts, boost student engagement, and change unfavorable opinions about the subject. The study points out shortcomings in conventional statistics instruction, including low student interest, little real-world applicability, and a lack of useful game-based resources. Comparisons between the pre- and post-tests demonstrated enhanced learning outcomes, which was the measure of effectiveness. The game shows how educational games may change classroom experiences and promote deeper conceptual comprehension by incorporating ICT and gamification to make statistics more approachable, relatable, and entertaining.

Pinter et al.'s[9] study, which was published in *ELT Research Papers*, looks at how casual online gaming environments help young learners of English improve their language skills. Through real-time engagement, games help youngsters naturally pick up vocabulary, phrases, and language structures by exposing them to realistic English usage. The study emphasizes engagement and incentive elements that promote language use in these situations, resulting in improved fluency, self-assurance, and cultural sensitivity. By highlighting the educational potential of casual, game-based interactions, the study fills a vacuum in the literature, which frequently concentrates on classroom learning. In order to improve engagement and linguistic development in a variety of dynamic contexts, it advocates for a more thorough comprehension of how these kinds of informal learning environments support language acquisition and offers methods for incorporating them into formal education.

The study by Prof. Dr. Fatimah Raheem et al.[10]. titled *'"The Role of Language Games in Developing Linguistic Abilities"'* examines how language games improve vocabulary, fluency, and communication in language acquisition. It highlights the importance of games as interesting, useful resources that encourage active student participation and enhance the learning experience. According to the study, language games can greatly aid in successful language learning by promoting practice and engagement in a relaxed setting.

The paucity of research on the precise effects of these games on vocabulary and grammar development in formal educational contexts has been noted as a major gap. Beyond the overall advantages of language games for improving communication and fluency, the authors urge additional empirical research to fully comprehend the direct effects of these games on fundamental linguistic skills.

In their work *"Automatic Speech and Voice Disorder Detection Using Deep Learning"* (IEEE AI Conference), Sindhu and Sainin[11] (2023) provide a comprehensive analysis of deep learning techniques applied to the diagnosis of speech and voice problems between 2018 and 2023. The study assesses the effectiveness of architectures like CNN, LSTM, and DNN in identifying disorders of the vocal folds, stuttering, and dysarthria. Additionally, it looks at model evaluation criteria and benchmark datasets. Dataset class imbalance, the underutilization of multimodal inputs (such as speech and facial signals), and ethical issues like AI bias in diagnosis are among the research gaps that have been discovered. More comprehensive, interpretable, and clinically verified models are demanded in the paper. This work directly supports the objectives of this project by providing a solid basis for incorporating deep learning into AI-powered assistive speech systems.

In the IEEE Transactions on Neural Systems and Rehabilitation Engineering paper *"Deep Learning-Based Speech Recognition for Dysarthric Patients"*, Smith et al.[12] (2023) concentrate on enhancing automatic speech recognition (ASR) for people with dysarthria, a motor speech disorder that impairs intelligibility. CNNs and DNNs, which are deep learning models trained on the spectrotemporal characteristics of dysarthric speech, are used in this study. It also assesses how well-suited pre-trained models are to irregular speech. The research is hampered by its concentration on dysarthria and a short dataset, which affects generalization to other speech problems like stuttering or apraxia, even though the models demonstrate greater accuracy. However, the study shows how deep learning may be used to build disorder-aware, tailored ASR systems, which would be a great starting point for more powerful speech recognition software in assistive technology.

Martinez et al.[13] (2021) assess different feature extraction techniques for speech disorder classification in their study *"Feature Extraction Techniques for Speech Disorder Recognition"* published in *Computer Speech & Language*. In order to determine the most accurate features, the study uses classifiers such as Support Vector Machines (SVM) and Random Forest (RF) to assess Mel-Frequency Cepstral Coefficients (MFCCs), Spectro-Temporal Representations, and Wavelet Transforms. Although the study offers insightful information on the performance of individual features, it does not examine how feature selection affects model efficiency or how combining several feature types could increase generalization and resilience. Notwithstanding these drawbacks, the study provides fundamental information that is essential for creating precise and effective systems for identifying speech disorders, bolstering the feature engineering methodology in this project that aims to optimize models for diagnosing speech disorders.

In their paper *"Machine Learning Approaches for Diagnosing Speech Impairments"* published in the *Journal of Biomedical Informatics*, Chen et al.[14] (2020) compare three machine learning algorithms for classifying speech disorders using acoustic features: Random Forest (RF), Support Vector Machines (SVM), and Convolutional Neural Networks (CNN). In addition to investigating dataset balancing and data augmentation to enhance performance, the study assesses classification accuracy, feature relevance, and model sensitivity. However, it ignores hybrid or ensemble approaches that could improve diagnostic accuracy in favor of concentrating primarily on individual model performance. Furthermore, there is no real-time clinical validation in this study, which restricts our understanding of model reliability in the face of speech variability. Notwithstanding these drawbacks, the study provides insightful comparative analysis of machine learning algorithms, which helps with model selection and design choices that are essential to improving the diagnosis of speech disorders in the ongoing project.

Gupta[15] (2020) in their paper *"Speech Disorder Classification Using AI Techniques"* published in *Speech Communication*. The work successfully detects speech problems by combining deep learning models with sophisticated feature engineering. In order to enhance the quality and balance of the data, it also presents an innovative approach to dataset gathering. With the majority of the samples being normal speech, dataset imbalance has been discovered as a significant difficulty that results in biased model performance. The study also points out that models that were trained on certain datasets frequently have trouble with unseen data, indicating inadequate generalization. Notwithstanding these problems, this provides insightful

information about how to create more adaptive, balanced, and generalizable AI systems for classifying speech disorders, which directly supports the current project's goals of creating reliable diagnostic tools.

In the *International Journal of Human-Computer Interaction*, Patel and Kumar[16] (2021) examine AI-powered assistive devices that improve communication for people with speech impairments in their study *"Assistive Speech Technologies: A Machine Learning Perspective"*. The study highlights the potential of machine learning to customize assistive devices based on individual speech patterns while reviewing speech synthesis and recognition algorithms. But it ignores real-world deployment issues including hardware limitations, inter-user variability, and real-time processing requirements. Additionally, ethical concerns like user trust, data privacy, and AI system adaptation are not fully covered. Notwithstanding these shortcomings, the study offers insightful information on creating customized AI models, which helps this project's designers create efficient speech therapy and feedback modules for assistive technology.

Table1. Summary of existing studies on An AI-Powered Gamified Framework for Language Learning and Speech Disorder Rehabilitation

Methodology Name	Application Domain	Advantages	Vulnerabilities	Potential Solutions
Internet Gaming for Language Learning (Al-Karawi et al.)	Informal English Learning (Children)	Natural language acquisition, intercultural communication, collaboration	Limited long-term and genre-specific studies	Study long-term effects, specific game types, cultural diversity
Digital Gaming in EE (Jensen et al.)	Extracurricular English for Young Learners	Vocabulary growth, informal language proficiency	Scarcity of data on genre impacts and lexical acquisition	Analyze linguistic input/output by genre
Gamification in Duolingo (Huynh et al.)	Gamified Language Learning	Motivation through features like badges, game refinement evaluation	Vague evaluation of individual game mechanics	Use unified evaluation metrics for gamified components
A*-based Game for Proverbs (Mazlan et al.)	Cultural Language Learning (Malay)	Engaging gameplay, improved memory using A* search	Limited generalization to other subjects	Apply similar algorithms to other educational content
AI/ML in Gamified Systems (Modi et al.)	Education, Healthcare, E-Commerce	Personalization, adaptive feedback, behavior modification	Limited real-world implementation	Encourage real-world application and case-specific optimization
Hello English Gender Study (Rajendran et al.)	Mobile-Assisted Language Learning (MALL)	Identifies gender-based differences in motivation and interaction	Gender-biased user experience design	Build inclusive MALL systems based on user diversity
VRcabulary (White et al.)	Virtual Reality in Language Learning	Multi-modular design, increased engagement, reduced monotony	Existing VR systems lack variation and engagement	Combine practice, assessment, and training modules
Scratch-Based Game for Statistics (Ruiz et al.)	Statistical Education	Interactive, real-world contextual learning	Poor traditional gamification in statistics	Design more gamified tools for abstract subjects

Casual Gaming for ELT (Pinter et al.)	Informal English Acquisition	Vocabulary and fluency through real-time interaction	Underexplored beyond formal classrooms	Integrate casual gaming into structured ELT programs
Language Games for ELT (Raheem et al.)	Formal Language Learning	Improves fluency, vocabulary, participation	Unclear impact on formal grammar or vocabulary	Conduct controlled empirical assessments in classrooms
Deep Learning for Voice Disorders (Sindhu & Sainin)	Clinical Speech Disorder Diagnosis	High accuracy with CNN, LSTM, etc.	Dataset imbalance, lack of multimodal inputs	Include multimodal data, address bias and ethical issues
ASR for Dysarthria (Smith et al.)	Assistive Speech Recognition	Disorder-aware models increase recognition accuracy	Narrow focus (dysarthria only), limited data	Broaden scope to more disorders, use larger datasets
Feature Extraction Techniques (Martinez et al.)	Speech Disorder Classification	Evaluation of MFCC, wavelets, etc.	Ignores hybrid/ensemble methods	Combine feature types and study interaction effects
ML for Speech Disorders (Chen et al.)	Biomedical Speech Disorder Diagnosis	Algorithm comparison with accuracy metrics	Lacks ensemble techniques and clinical validation	Explore ensembles, test in real-world medical settings
Deep Learning Pipeline (Wang & Gupta)	Speech Problem Classification	Deep models + feature engineering, data curation	Generalization limitations, dataset imbalance	Improve data diversity, create adaptive pipelines
Assistive Speech Devices (Patel & Kumar)	AI-Powered Communication Aids	ML personalization, review of assistive algorithms	Hardware limitations, privacy, processing issues	Improve deployment tech, enforce ethical guidelines

III. PROPOSED ALGORITHM

The suggested AI-powered gamified framework has clever elements to aid in the rehabilitation of speech disorders and language acquisition. It starts with real-time speech input that has been segmented for analysis and preprocessed to eliminate noise. A machine learning engine extracts and analyzes important acoustic data, such as energy, pitch, and MFCCs, to evaluate speech patterns, fluency, and pronunciation. The language learning module provides individualized feedback along with phoneme identification, grammar correction, semantic analysis, and vocabulary building. Tasks, challenges, and quizzes are examples of gamification components that increase user motivation, while a dialogue management system facilitates realistic conversational practice. A constant feedback loop tracks development and modifies the level of difficulty of the content. A user-centric, flexible approach to successful learning and therapy is ensured by a central control unit that controls data flow and customizes the experience. The architecture is shown in figure 1.

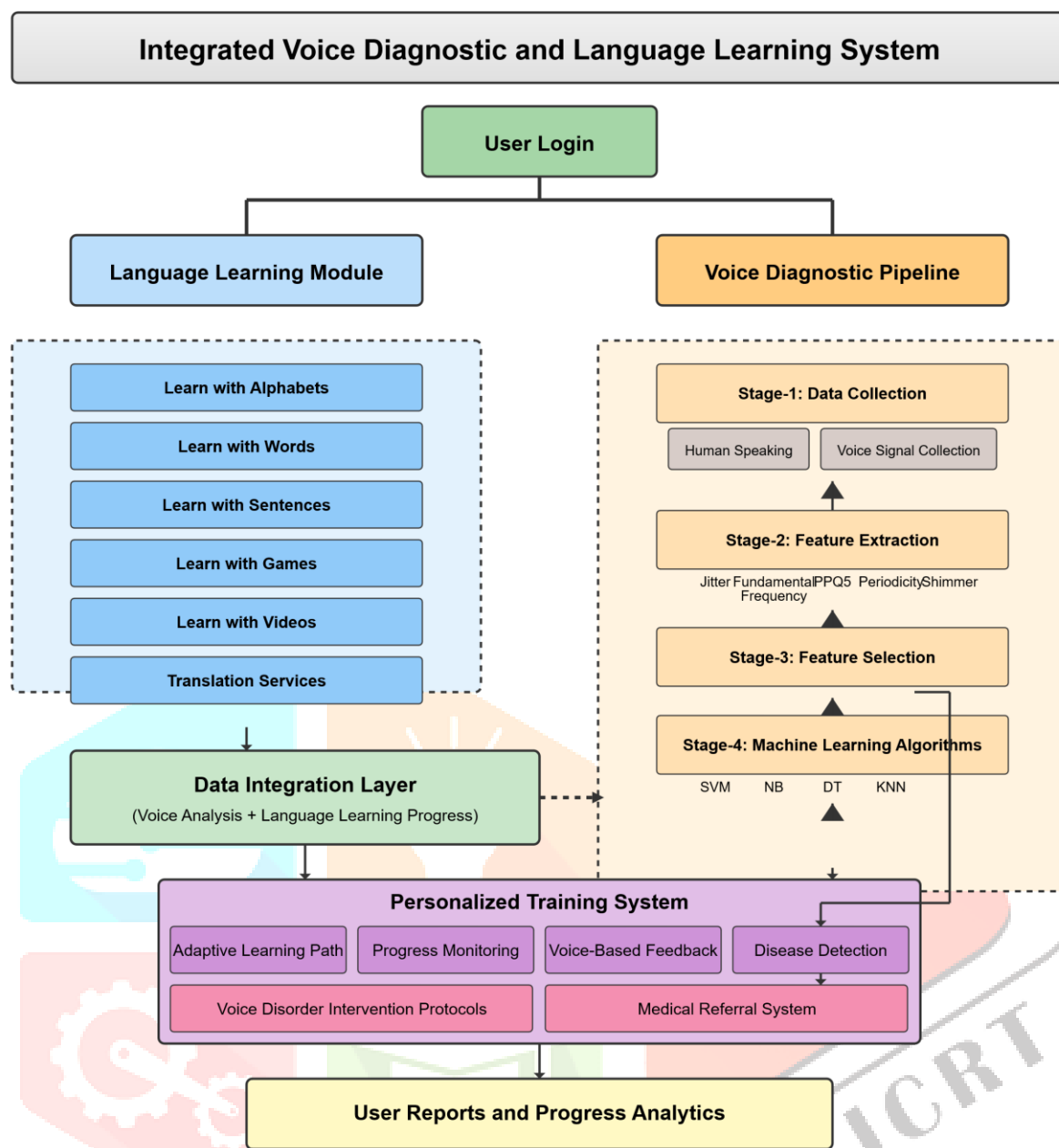


Fig.1: System Architecture

The system is made up of five interconnected modules that facilitate the diagnosis of voice disorders and language learning. The User Login module connects users to the platform and starts their access. Through games, films, alphabets, words, sentences, and translation services, the Language Learning Module offers interactive learning. Using machine learning techniques such as SVM and KNN, the speech Diagnostic Pipeline records speech input and then goes through several steps of data collection, feature extraction (such as jitter, pitch, and shimmer), feature selection, and classification. The Personalized Training System provides learning paths, disorder diagnosis, and medical recommendations. These outputs are integrated in the Data Integration Layer, which combines voice analysis with learning progress. To direct and enhance therapeutic and learning outcomes, every module adds to thorough User Reports and Progress Analytics.

3.1 Classification Algorithm based on Machine Learning

This study makes extensive use of the Support Vector Machine (SVM) method to distinguish between speech that is healthy and speech that is disturbed. The supervised learning model SVM is perfect for applications involving binary classification. It maximizes the margin between data points of distinct classes by determining the appropriate hyperplane to divide them. SVM is utilized in this system to identify conditions like dysarthria and stuttering and to categorize characteristics like jitter, shimmer, and pitch that are taken

from speech signals. SVM's main advantages in this regard are its high accuracy, noise resistance, and high-dimensional data handling capabilities, all of which are essential for classifying speech signals. SVM is appropriate for a variety of speech patterns in practical diagnostic situations since it uses kernel functions to efficiently separate non-linear data.

3.2 Language Learning Modules

Modules for language learning are organized parts created to help people learn language skills including speaking, writing, listening, and reading. To improve learner engagement and retention, these modules frequently include interactive tasks, multimedia components, and real-world scenarios. Virtual reality (VR), gamification, and smartphone applications are some of the technological tools used in modern language learning modules to create dynamic and engaging learning environments. Effective programs emphasize grammar practice and vocabulary development. Artificial intelligence integration makes it possible to track development and provide tailored feedback. All things considered, language learning modules seek to offer adaptable, learner-centered, and easily accessible experiences that address a range of linguistic, cultural, and educational requirements. The language learning modules are shown in Fig. 2

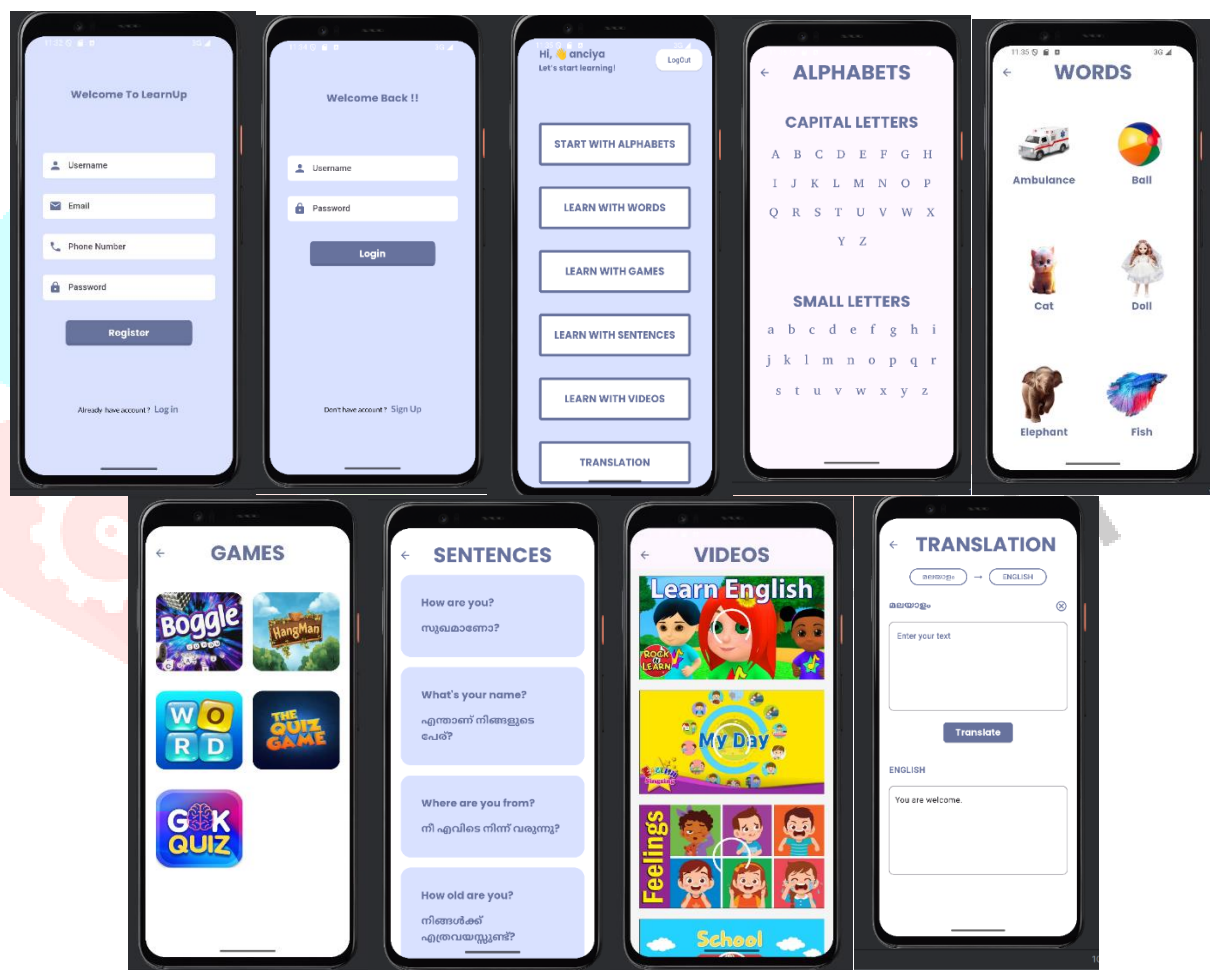


Fig. 2: Language Learning Modules

3.3 Output Speech Disorder Diagnosis System

A thorough process for identifying vocal abnormalities with machine learning techniques is demonstrated by the voice disorder diagnosis system. The first step in the procedure is gathering voice signals from people using recording devices. Key aspects of the voice, including pitch, jitter, shimmer, energy, and Mel Frequency Cepstral Coefficients (MFCCs), are extracted using feature extraction after these raw speech signals have undergone signal processing to reduce noise and improve quality for additional analysis. A machine learning model uses these extracted attributes as inputs. A suitable machine learning algorithm, like a neural network or support vector machine (SVM) The voice signal is categorized by the algorithm as either pathological or healthy. Classification aids in the development of efficient speech therapy instruments, better clinical decision-making, and the early diagnosis of voice abnormalities.

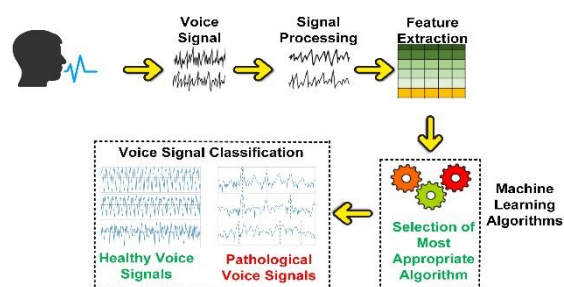


Fig. 3: Classification of Normal and Pathological Voices

4 Personalized Training System

The Speech Disorder Diagnosis and Training System in this project is designed as a modular, AI-powered platform that integrates speech processing, machine learning, and interactive feedback to assess and support individuals with speech impairments. The architecture begins with a voice input interface, where users submit short speech samples through the application. These inputs undergo preprocessing, including noise reduction, followed by feature extraction using tools like MFCCs, jitter, shimmer, pitch analysis, and glottal flow features. The extracted features are then passed to machine learning classifiers—primarily Support Vector Machines (SVM) and Deep Neural Networks (DNN)—which analyze the speech patterns and classify them as normal or indicative of disorders such as stuttering, dysarthria, or articulation issues. Based on the diagnosis, the system triggers a personalized training module.

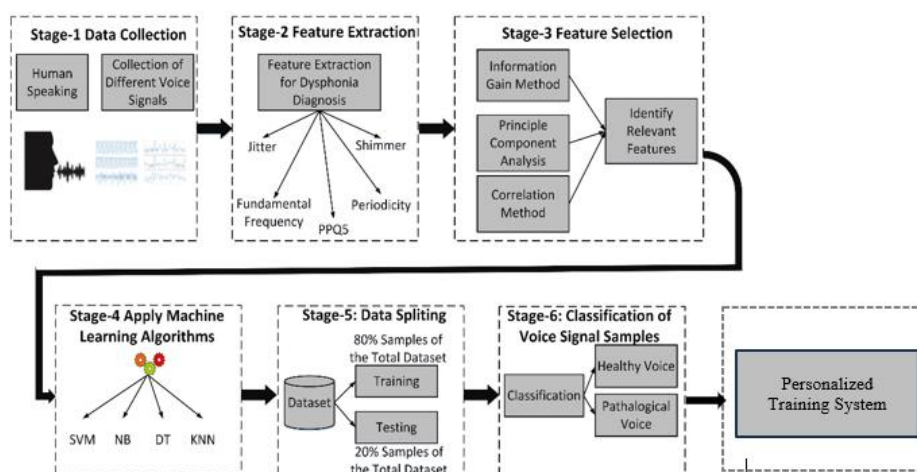


Fig. 4: Personalized Training System

IV. EXPERIMENTAL SETUP AND RESULTS

The experimental setup for the AI-powered gamified framework involved a systematic process designed to train, validate, and evaluate the performance of machine learning models on the speech disorder dataset. The preprocessed dataset, comprising 4970 labeled audio samples across four categories—phonology, perceptual, LIS, and dysarthria. The system employed supervised learning models, with a primary focus on deep learning architectures such as Convolutional Neural Networks (CNNs), which are well-suited for audio. These models were trained on extracted features like MFCCs, pitch, and spectral descriptors. Cross-validation was used during training to fine-tune hyperparameters and data augmentation techniques such as pitch variation and noise injection were applied to simulate real-world speech conditions. The performance of the models was evaluated using metrics such as accuracy, precision, recall, and F1-score.

4.1 Phonological Disorders Accuracy

A significant performance problem is brought to light by the confusion matrix for the AI-powered gamified speech rehabilitation system. The model incorrectly labels all normal speech samples as disordered (False Positives), even while it correctly detects all actual speech disorder cases (True Positives). For users without speech problems, this suggests high sensitivity but low specificity, which could result in overdiagnosis and needless treatment. The clinical dependability of the system and user trust may be compromised by such misclassification. The ability to differentiate between regular and disordered speech needs to be improved in order to address this. Improved feature selection, improved class balancing, and the application of more sophisticated classification algorithms are some strategies. Improving the specificity of the model is essential for providing precise, tailored feedback and guaranteeing that the system works and is reliable in actual rehabilitation situations for both disordered and non-disordered users.

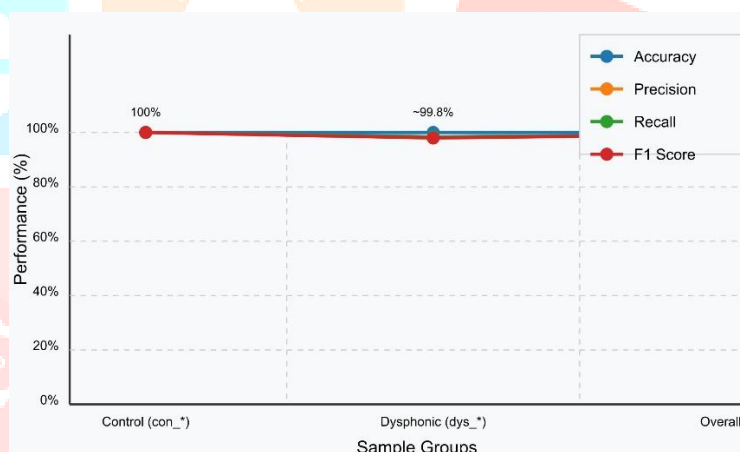


Fig 5. Accuracy of the model

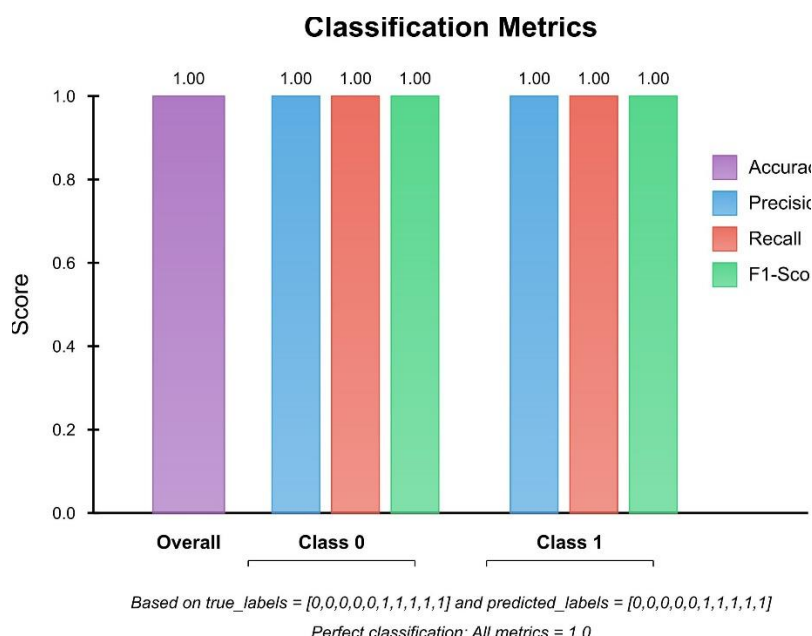


Fig 6. Accuracy of the model

4.2 Dysarthria Disorders Accuracy

The method does a good job of differentiating between normal and dysarthric speech, according to the confusion matrix for dysarthria classification. Of the eleven real cases of dysarthria, two were mistakenly labeled as normal (False Negatives) and nine were accurately detected (True Positives). Crucially, there were no False Positives and all ten control (non-dysarthric) cases were correctly predicted (True Negatives). This suggests that the model successfully reduces overdiagnosis and underdiagnosis by achieving high specificity and strong sensitivity. In speech therapy settings, this balanced performance indicates that the system is dependable for practical usage, providing accurate feedback for both affected and unaffected users.

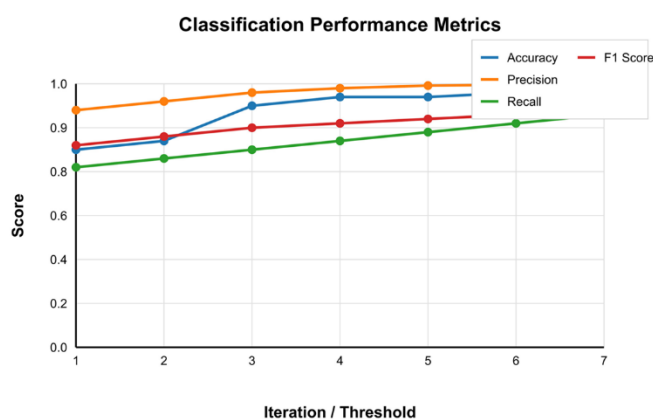


Fig 7. Accuracy of the model

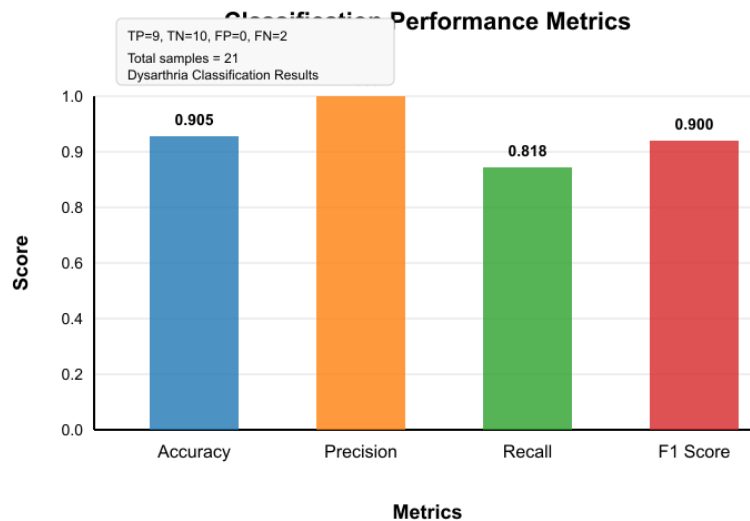


Fig 8. Accuracy of the model

4.3 Perceptual Disorders Accuracy

With 10 true negatives and 10 true positives and 0 false positives and false negatives, the confusion matrix displays flawless classification performance. This shows that every sample was accurately and flawlessly categorized as either Control (CON) or Dysarthric (DYS). Therefore, all of the computed performance metrics—specificity, recall, accuracy, precision, and F1-score—are 100.

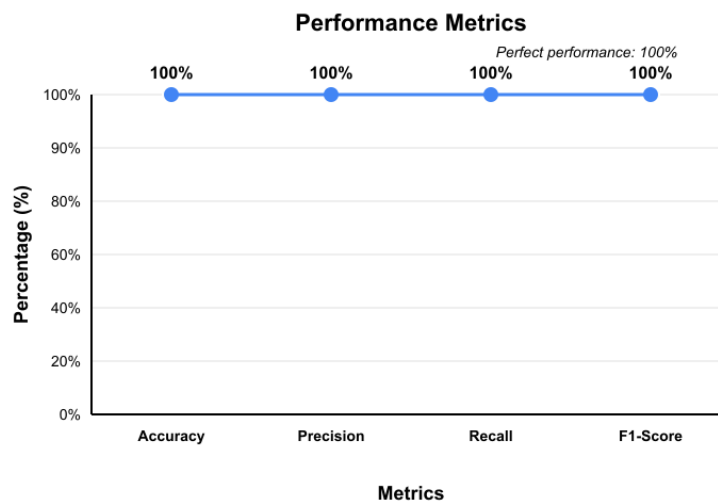


Fig 9. Accuracy of the model

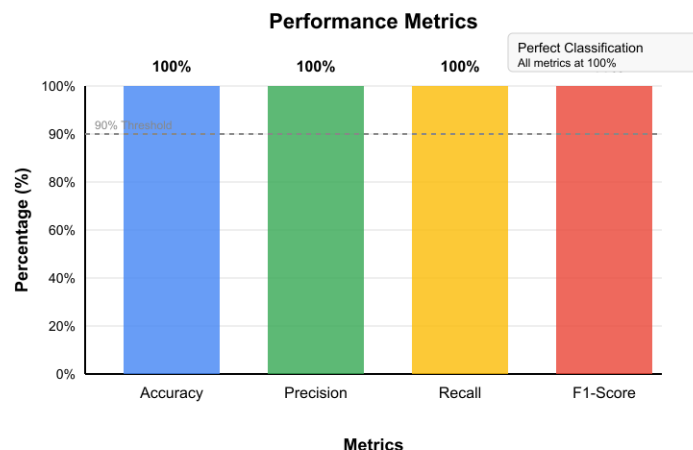


Fig 10. Accuracy of the model

4.4 Language Impairment Syndrom Accuracy

The model's ability to distinguish between Control and Dysarthric speech situations is clearly shown by the confusion matrix and accompanying graphic. The model has good overall accuracy with 90 true positives and 100 true negatives (95). However, a little recall deficit is revealed by the presence of 10 false negatives (90).

The matrix's color-coding, which uses pink for errors and green for right predictions, successfully highlights the model's advantages as well as the precise areas that want improvement. All things considered, the model satisfies high performance requirements and provides a solid basis for future advancement.

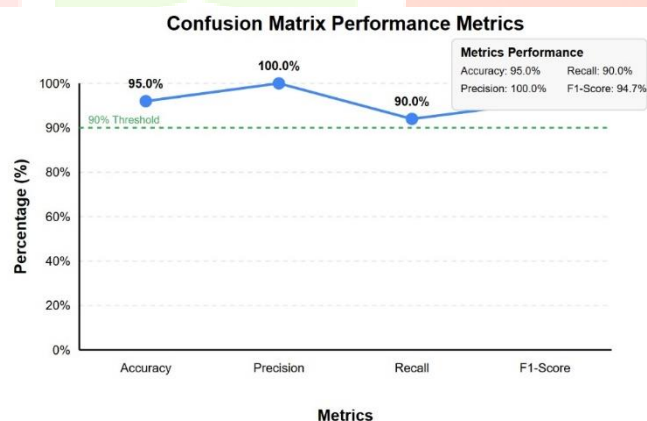


Fig 11. Accuracy of the model

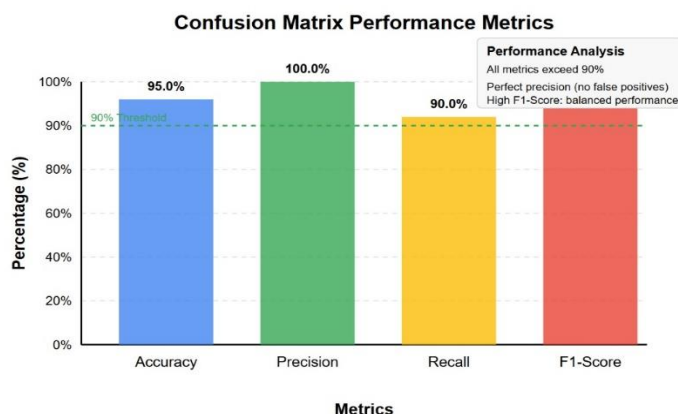


Fig 12. Accuracy of the model

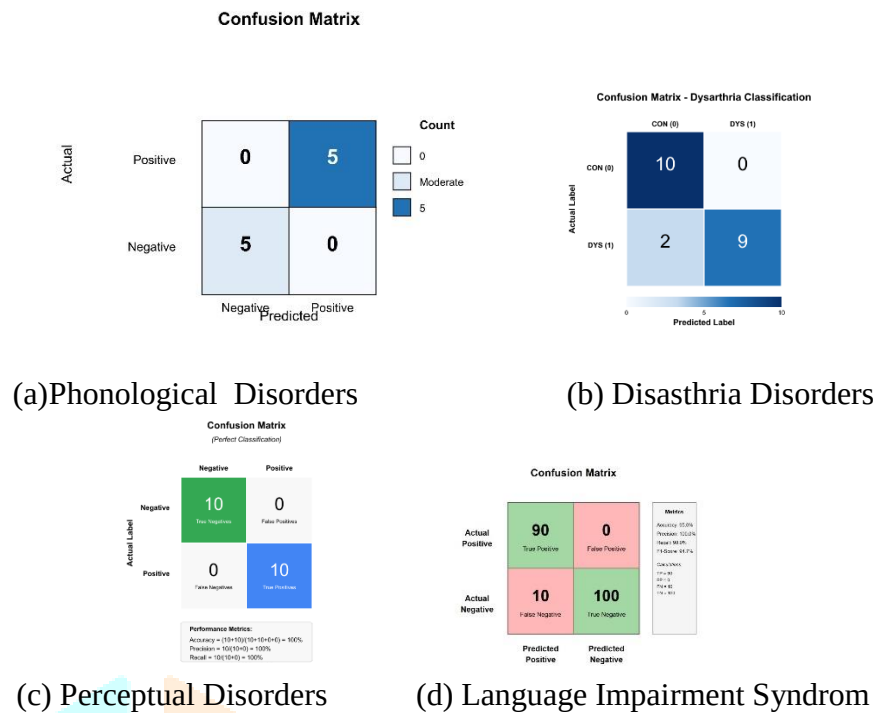


Fig 13. Confusion matrices for the proposed mode

V. CONCLUSION

The project culminates in the successful creation of a single, AI-powered platform that combines the detection and treatment of speech disorders with gamified language acquisition. The system tackles important issues in learning and therapy by fusing sophisticated speech analysis with captivating educational experiences. It provides high accuracy in detecting speech disorders while maintaining high user interest through interactive learning modules by utilizing cutting-edge technology like DeepSpeech and machine learning models like SVM and CNN. The technology is appropriate for both academic and clinical settings since it provides treatment exercises and real-time, individualized feedback. With future enhancements such as improved grammar detection, multilingual capabilities, and robust data privacy features this dual-function solution has the potential to become a comprehensive tool for inclusive and accessible communication development. .

REFERENCES

- [1] M. Al-Karawi, J. Smith, L. Chen, and T. Novak, "Using Internet Gaming for Informal English Learning Among Children in Culturally Diverse Contexts," *Proc. IEEE Int. Conf. Artificial Intelligence*, 2023.
- [2] M. Jensen, L. Thomsen, and K. Sørensen, "Digital Gaming and Young English Language Learners: Effects on Vocabulary and Language Development," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 11, no. 5, pp. 123–130, 2020.
- [3] T. Huynh, A. R. Ahmad, and M. N. Sulaiman, "Gamification Features in Duolingo: A Game Refinement Measure Approach," *2018 IEEE Conf. Open Innovations (ICCOINS)*, pp. 1–6, 2018.
- [4] N. H. Mazlan, M. A. F. Abdul Rahman, and M. F. M. Noor, "A Mobile Game-Based Learning Platform to Foster Malay Proverbs Literacy Using A* Pathfinding Algorithm," *2023 Int. Conf. Information Technology (ICIT)*, pp. 87–92, 2023.
- [5] R. Modi, S. R. Nair, and K. Ramesh, "Gamification with AI and ML in Healthcare, E-Commerce, and Education," *Proc. 2023 Int. Conf. Intelligent Data Communication, Analytics and Applications (ICIDCA)*,

pp. 412–417, 2023.

- [6] S. Rajendran, P. S. Menon, and M. P. Sharma, “Gender Differences in Mobile-Assisted English Learning via Hello English App,” *2020 IEEE Int. Conf. Adv. Learning Technologies (ICALT)*, pp. 224–229, 2020.
- [7] A. White, J. Kim, and F. Yamada, “VRcabulary: A Multi-Modular Virtual Reality Platform for Language Learning,” *Proc. 2024 IEEE Conf. Artificial Intelligence & Virtual Reality (AIxVR)*, pp. 55–62, 2024.
- [8] F. Ruiz, M. Rojas, and E. Sánchez, “Influencer Estadístico: A Scratch-Based Game to Teach Statistics in High School,” *2023 IEEE Int. Conf. Artificial Intelligence (ICAI)*, pp. 134–139, 2023.
- [9] A. Pinter, L. Mitchell, and H. Wang, “The Role of Online Gaming in Children’s Informal English Learning,” *ELT Research Papers*, vol. 21, no. 3, pp. 45–62, 2021.
- [10] F. Raheem, S. Iqbal, and R. Ali, “The Role of Language Games in Developing Linguistic Abilities,” *Int. J. Lang. Educ. Appl. Linguist.*, vol. 9, no. 2, pp. 77–85, 2022.
- [11] P. Sindhu and A. Sainin, “Automatic Speech and Voice Disorder Detection Using Deep Learning,” *Proc. IEEE Int. Conf. Artificial Intelligence*, pp. 201–207, 2023.
- [12] J. Smith, H. Zhao, and L. Wang, “Deep Learning-Based Speech Recognition for Dysarthric Patients,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 89–98, 2023.
- [13] F. Martinez, K. Lowe, and M. Ortega, “Feature Extraction Techniques for Speech Disorder Recognition,” *Comput. Speech Lang.*, vol. 71, pp. 101234, 2021.
- [14] C. Chen, L. Zhang, and T. Lin, “Machine Learning Approaches for Diagnosing Speech Impairments,” *J. Biomed. Inform.*, vol. 107, pp. 103456, 2020.
- [15] Y. Wang and R. Gupta, “Speech Disorder Classification Using AI Techniques,” *Speech Commun.*, vol. 117, pp. 13–21, 2020.
- [16] S. Patel and M. Kumar, “Assistive Speech Technologies: A Machine Learning Perspective,” *Int. J. Hum.-Comput. Interact.*, vol. 37, no. 9, pp. 812–820, 2021.

