LEXAYUDHA : PERSONALIZED AI-DRIVEN REHABILITATION FOR ADOLESCENTS WITH DYSLEXIA AND DYSCALCULIA

Thathsara Pramodya Thalangama

(IT21223594)

BSc (Hons) degree in Information Technology Specializing in Software Engineering

Department of Computer Science & Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

April 2025

LEXAYUDHA : PERSONALIZED AI-DRIVEN REHABILITATION FOR ADOLESCENTS WITH DYSLEXIA AND DYSCALCULIA

Thathsara Pramodya Thalangama

(IT21223594)

Dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Hons) in Information Technology Specializing in Software Engineering

Department of Computer Science & Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

April 2025

DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 10/04/2025

The above candidate has carried out research for the bachelor's degree Dissertation under my supervision.

Signature of the supervisor: Date: 2025-04-11

Abstract

Dyslexia and dyscalculia are prevalent learning disorders that significantly impair an individual's ability to comprehend written and spoken language. Adolescents with these conditions often struggle to process spoken language at typical conversational speeds, leading to difficulties in understanding verbal instructions and engaging effectively in learning environments. This limitation not only hinders academic progress but also impacts their social and emotional development. Despite the availability of various assistive tools, there remains a critical gap in addressing the unique needs of dyslexic and dyscalculic learners, particularly in personalizing speech pace delivery to enhance comprehension.

This study introduces LexAyudha, an AI-powered personalized learning aid tool designed to improve speech comprehension for dyslexic and dyscalculic adolescents aged 8–12 years. The system leverages a hybrid Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) model to analyze mel spectrogram images derived from speech audio data. By extracting spatial and temporal features, the model predicts optimal speech pace tailored to individual needs. The predicted speech pace is then integrated into a Google Text-to-Speech (TTS) system, delivering customized audio output that enhances the learning experience. The methodology involves capturing speech samples through calibration activities, converting audio into mel spectrogram images using libraries like Librosa, and analyzing these images with the hybrid CNN-RNN model.

Testing and validation demonstrate LexAyudha's effectiveness in predicting and delivering personalized speech pace with high accuracy, achieving a low error margin. Results indicate that the system significantly improves speech comprehension, reduces cognitive load, and fosters inclusivity for dyslexic and dyscalculic learners. Beyond its technical contributions, this study highlights the potential of AI-driven solutions to address educational challenges faced by neurodiverse populations. By integrating advanced deep learning techniques with user-centric design, LexAyudha exemplifies how technology can transform learning experiences for individuals with learning disabilities.

Keywords*: Dyslexia, Deep Learning, CNN, RNN, Mel Spectrogram, TTS*

Acknowledgement

I would like to express my heartfelt gratitude to everyone who contributed to the successful completion of this project. First and foremost, I am deeply grateful to my supervisor Prof. Samantha Thelijjagoda , co-supervisor Dr. Junius Anjana , and our external supervisor Dr. Asiri Hewamalage , for their invaluable guidance, constructive feedback, and unwavering support throughout this research journey. Their expertise was instrumental in shaping the direction and quality of this project.

I am also immensely thankful to Mrs. G. Nadeeka Priyangani Rupasinghe , Principal of R/Raththurugala Vidyalaya, Ratnapura , for her generous support in facilitating access to students for collecting data and performing usability testing. Her cooperation ensured rigorous evaluation of the system in real-world scenarios, which significantly enhanced its credibility and effectiveness.

Special acknowledgment goes to my fellow team members, whose dedication, collaboration, and innovative ideas were pivotal in bringing this project to fruition. I also extend my thanks to my peers and colleagues for their contributions to data collection, testing, and troubleshooting, as well as the students who participated in testing.

On a personal note, I am incredibly grateful to my family for their unwavering encouragement and emotional support throughout this journey.

This project would not have been possible without the collective efforts of these individuals and organizations, and I am truly grateful for their support.

TABLE OF CONTENTS

Declaration i
Abstract ii
Acknowledgement iii
List of Tables vi
List of abbreviations vi
1. Introduction 1
 1.1 Background Literature 2
 1.2 Research Gap 4
 1.3 Research Problem 5
 1.4 Research Objectives 6
  1.4.1 Main Objective 7
  1.4.2 Sub Objectives 8
2. Methodology 13
 2.1 Methodology 13
 2.2 Commercialization Aspects of the Product 20
  2.2.1 Target Market 20
  2.2.2 Pricing Strategy 21
  2.2.3 Marketing Strategy 22
  2.2.4 Global Expansion and Localization 23
  2.2.5 Revenue Projections and Long-Term Vision 23
 2.3 Testing and Implementation 23
  2.3.1 Testing Methodologies 24
  2.3.2 Implementation Strategy 25
3. Results and discussion 29
 3.1 Results 29
  3.1.1 Speech Pace Prediction Accuracy 29

3.1.2 System Performance 31
3.2.3 User Feedback 32
3.2.4 Visualizing Results 33
3.2 Research Findings 35
3.2.1 Importance of Personalized Speech Pace 36
3.2.2 Effectiveness of Hybrid CNN-RNN Models 37
3.2.3 Role of Deep Learning in Dyslexia Support 38
3.2.4 Impact on Learning Outcomes 39
3.3 Discussion 41
3.3.1 Significance of the result 41
3.3.2 Comparison with Existing Solutions 43
3.3.3 Limitations 44
3.3.4 Opportunities for Future Research 45
3.4 Contribution 47
4. conclusion 49
5. References 52

LIST OF FIGURES

Figure 1: System Overview Diagram 2
Figure 2: Hybrid CNN-RNN Model Architecture 8
Figure 3: Wav2Vec 2.0 model initialization 10
Figure 4: Branch reduction layers 10
Figure 5: Feature combined layer 10
Figure 6: Speech Calibration UI 14
Figure 7: Creating spectrogram image using audio file 17
Figure 8: Gradual unfreezing of layers 18
Figure 9: Training and Validation Loss over Epochs chart 34
Figure 10: Model performance metrics 35
Figure 11: Appending batch normalization layer after each conv layer 38

LIST OF TABLES

LIST OF ABBREVIATIONS

AI – Artificial Intelligence

RNN - Recurrent Neural Network

CNN - Convolutional Neural Networks

TTS – Text to Speech

1. INTRODUCTION

The personalized speech pace module in LexAyudha is a cutting-edge solution designed to address the speech comprehension challenges faced by dyslexic adolescents. By leveraging advanced deep learning techniques—including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN)—this module offers a tailored speech pace that adapts to individual needs.

The process begins with the display of predefined sentences, ensuring multiple attempts to minimize user error. The adolescent's spoken responses are recorded via the device's microphone and sent as an array of audio files to a backend server. Each file is processed individually by the "FlaskAIServices" microservice, which handles AI-related tasks within the LexAyudha application.

Within FlaskAiServices, every audio file is segmented into 10-second intervals and converted into both an audio waveform and its mel spectrogram representation. These dual inputs are then processed by a specialized Hybrid CNN-RNN model, which analyzes temporal and spatial details to accurately determine the adolescent's speech pace. The predicted pace is subsequently relayed to the SpeechRateService microservice, where an average speech pace is calculated and stored in the database.

Finally, this personalized speech pace is utilized by the Google Text-to-Speech service to deliver customized audio output, enhancing the speech comprehension experience for dyslexic adolescents and the overall system diagram is shown in Fig. 1. This report provides an in-depth analysis of the module's design, Methodology, outlining the transformative role it plays in personalized learning environments such as LexAyudha.

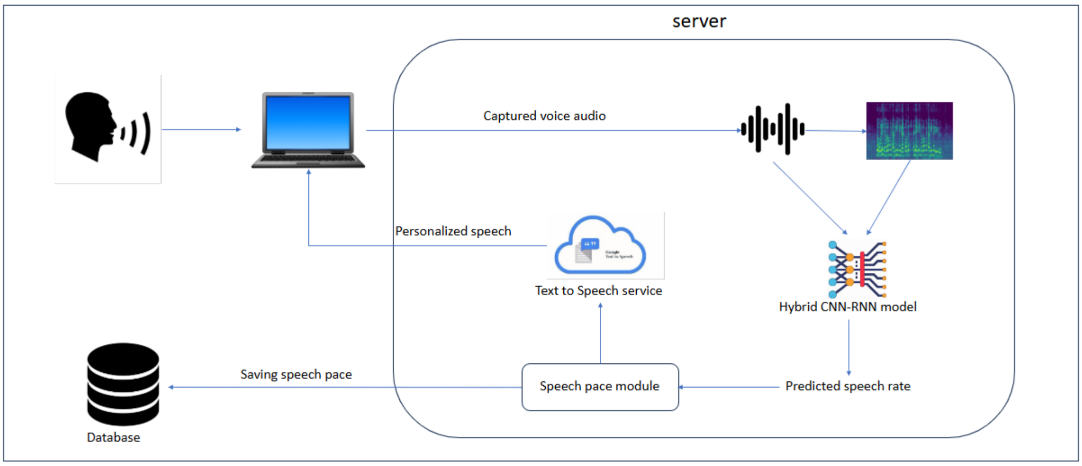

Figure 1: System Overview Diagram

1.1 Background Literature

In today's rapidly evolving educational landscape, the integration of technology into learning environments has become a cornerstone for addressing diverse needs. Among these needs, learning disorders such as dyslexia have garnered significant attention due to their profound impact on individuals' ability to comprehend written and spoken language. Dyslexia, one of the most prevalent learning disabilities, affects approximately 20% of the global population [1]. Adolescents with dyslexia often face challenges in processing spoken language at typical conversational speeds, leading to difficulties in understanding verbal instructions and engaging effectively in learning activities. This challenge underscores the critical importance of speech comprehension as a foundational skill for effective communication and academic success.

The field of neurodiversity has gained traction in recent years, emphasizing that neurological differences such as dyslexia, ADHD, and autism spectrum disorders are natural variations of the human brain rather than deficits. Despite this growing awareness, traditional educational systems and tools often fail to provide personalized support tailored to the unique needs of dyslexic learners. Many technological solutions have emerged to address these challenges, ranging from text-to-speech applications like Speechify and TextAid to advanced tools like Speech Rate Meter (SRM+).

However, while these tools offer features such as adjustable reading speed and customizable speech rates, they lack the sophistication required to deliver truly personalized speech pace adjustments based on individual characteristics.

Recent advancements in artificial intelligence (AI) and deep learning have opened new avenues for addressing the specific needs of dyslexic adolescents. Techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have demonstrated exceptional capabilities in analyzing complex data patterns, including audio signals and mel spectrogram images. These technologies have been successfully applied in fields such as emotion detection in speech [2] and classification of dyslexia among school students using deep learning models [3]. However, the application of AI-driven methodologies to detect and personalize speech pace for dyslexic individuals remains underexplored.

Existing literature highlights the importance of temporal and spatial features in speech analysis. For instance, spectrograms—visual representations of audio signals—have proven effective in capturing phonetic features, intonation patterns, and temporal dynamics [4]. These features are particularly relevant for dyslexic individuals, whose speech patterns may differ significantly from those of non-dyslexic individuals. By leveraging advanced AI models, it is possible to extract meaningful insights from spectrogram images and raw audio data, enabling accurate prediction of speech pace.

Despite these promising developments, the integration of AI technologies into practical tools for dyslexic learners has been limited. Most existing solutions focus on general accessibility features rather than personalized interventions. For example, while tools like Speechify allow users to adjust speech rate manually, they do not automatically adapt to an individual's unique speech comprehension needs. Similarly, SRM+ provides metrics such as articulation rate and pause scores but lacks the capability to tailor speech pace dynamically based on real-time user input. This gap in the current technological landscape underscores the need for innovative solutions that combine cutting-edge AI techniques with user-centered design principles.

In summary, the background literature reveals a growing recognition of the challenges faced by dyslexic adolescents in processing spoken language. While existing tools offer some level of support, they fall short of delivering personalized speech pace adjustments that align with individual needs. The integration of AI-driven methodologies, particularly hybrid CNN-RNN models, presents a promising opportunity to address this limitation and enhance the learning experience for dyslexic adolescents.

1.2 Research Gap

Despite significant advancements in technology and increasing awareness of dyslexia, many challenges faced by dyslexic adolescents remain underexplored. One such challenge is the personalized delivery of speech pace, which plays a crucial role in enhancing speech comprehension and overall learning outcomes. While various tools and methods exist to determine speech pace, or speech rate, significantly fewer studies have explored the relationship between speech rate and dyslexic conditions. Specifically, temporal and spatial patterns common to dyslexic individuals have been largely neglected in existing methods and tools.

The detection of speech features has gained attention in recent years, resulting in promising developments in areas such as hate speech detection [5] and emotion detection in speech [2]. However, the application of modern technology to accurately determine speech pace for dyslexic adolescents remains lacking. Existing tools often rely on manual adjustments or generalized algorithms that fail to account for the unique characteristics of dyslexic speech. This limitation highlights a critical research gap: the absence of AI-driven solutions capable of detecting and adapting to the specific speech patterns of dyslexic individuals.

Furthermore, the integration of spectrogram-based analysis with deep learning models has not been fully explored in the context of personalized speech pace detection. Spectrograms, which capture both frequency and temporal information, offer a rich source of data for analyzing speech patterns. However, most existing studies focus on either visual feature extraction using CNNs or temporal pattern recognition using RNNs, without combining these approaches into a unified framework. This fragmented approach limits the accuracy and effectiveness of speech pace prediction, particularly for individuals with learning disabilities.

Another notable gap lies in the application of hybrid AI models, such as CNN-RNN architectures, to address the specific needs of dyslexic learners. While these models have demonstrated success in other domains, such as detecting brain abnormalities [6], their potential for personalized speech pace detection remains untapped. By leveraging the strengths of both CNNs and RNNs, it is possible to develop a robust system capable of extracting spatial and temporal features from speech data, thereby enabling accurate and personalized predictions.

Finally, the integration of predicted speech pace with Text-to-Speech (TTS) systems represents another area of opportunity. While TTS technologies have advanced significantly, their ability to adapt dynamically to individual speech comprehension needs is limited. By combining AI-driven speech pace prediction with TTS functionality, it is possible to create a seamless and personalized learning experience for dyslexic adolescents.

In conclusion, the research gap identified in this study encompasses the lack of AI-driven solutions for personalized speech pace detection, the underutilization of spectrogram-based analysis, and the limited integration of hybrid AI models and TTS systems. Addressing these gaps has the potential to significantly enhance the learning experience for dyslexic adolescents and contribute to more inclusive educational environments.

1.3 Research Problem

Dyslexia poses a significant challenge to personal development, often hindering academic and social growth. Studies have shown that early diagnosis and intervention can mitigate the adverse effects of dyslexia, potentially enhancing the learning process later in life [7]. However, despite these findings, many dyslexic adolescents continue to struggle with speech comprehension, which impairs their ability to learn efficiently. Effective speech comprehension is essential for better learning and overall comprehension, yet individuals with dyslexia often face barriers in processing spoken language at typical conversational speeds.

The primary research problem addressed in this study is the lack of personalized speech pace delivery systems tailored to the unique needs of dyslexic adolescents. Current tools and methodologies fail to account for the temporal and spatial features of speech that are particularly relevant for dyslexic individuals. This limitation results in suboptimal learning experiences, as dyslexic adolescents are unable to engage effectively with spoken content delivered at standard speech rates.

Furthermore, the absence of AI-driven solutions capable of dynamically adapting to individual speech comprehension needs exacerbates the problem. While existing tools offer some level of customization, they rely on manual adjustments or generalized algorithms that do not adequately address the specific challenges faced by dyslexic learners. This gap in technological support underscores the need for an innovative solution that leverages advanced AI techniques to detect and personalize speech pace. By addressing this research problem, the proposed study aims to develop a system that serves as both a training and aiding tool, designed to improve speech comprehension skills in dyslexic adolescents. The ultimate goal is to enhance the learning experience and provide meaningful support for those affected by dyslexia, thereby fostering more inclusive educational environments.

1.4 Research Objectives

The research objectives outlined in this study are meticulously designed to address the challenges faced by dyslexic adolescents in processing spoken language at typical conversational speeds. By leveraging advanced AI technologies, the study aims to develop a personalized learning aid tool that enhances speech comprehension and provides tailored support for dyslexic learners. The objectives are structured to ensure a systematic approach to achieving the overarching goal of improving learning outcomes for dyslexic adolescents aged 8–12 years.

1.4.1 Main Objective

The primary objective of this study is to design, develop, and implement an AI-powered learning aid tool LexAyudha that delivers personalized learning experiences specifically tailored for dyslexic adolescents. Dyslexia is one of the most prevalent learning disorders, affecting approximately 20% of the global population [1]. Adolescents with dyslexia often face significant challenges in traditional educational settings due to the lack of personalization in instructional methods. These challenges include difficulty in processing spoken language at standard speeds, fatigue caused by cognitive overload, and reduced engagement with learning materials.

To address these issues, LexAyudha incorporates cutting-edge AI technologies such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to analyze and predict individual speech pace. The system also integrates a Text-to-Speech (TTS) model to deliver customized audio output that aligns with the predicted pace. By providing a personalized learning experience, the tool aims to reduce cognitive fatigue, enhance engagement, and improve overall learning outcomes for dyslexic adolescents.

This main objective will be achieved through the realization of several key milestones:

- Development and Training of Hybrid CNN-RNN Models : The hybrid model architecture as shown in Fig. 2 combines the strengths of CNNs and RNNs to extract spatial and temporal features from spectrogram images and raw audio data. This ensures accurate prediction of speech pace.
- Integration of System Components : The project involves seamless integration of various components, including audio capture, spectrogram conversion, speech pace prediction, and TTS functionality.
- Comprehensive User Testing : Rigorous testing will be conducted to validate the effectiveness and reliability of the system, ensuring that it meets its objectives and provides measurable improvements in learning outcomes for dyslexic adolescents.

The project is designed to be completed within a one-year timeframe, ensuring timely delivery of a solution that aligns with both the project's goals and broader societal needs.

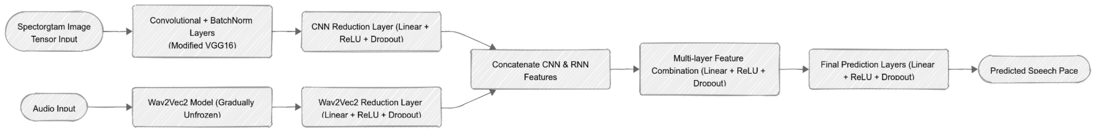

Figure 2: Hybrid CNN-RNN Model Architecture

1.4.2 Sub Objectives

To achieve the main objective, the following sub-objectives have been identified. Each sub-objective addresses a specific aspect of the research problem and contributes to the overall success of the project.

I. Convert Audio Files to Spectrogram Images

This initial sub-objective involves transforming raw audio inputs into visual representations known as spectrogram images. Spectrograms provide a detailed view of the frequency and intensity of sounds over time, capturing essential vocal characteristics that are critical for analyzing speech patterns. In the LexAyudha system, this step is crucial for preparing the data for advanced analysis and ensuring that subsequent processing accurately reflects the unique speech patterns of dyslexic individuals.

<u>Detailed Process</u>

1. Audio Capture : The process begins with the display of predefined sentences during a series of calibration tests. Adolescents' spoken responses are recorded via the device's microphone using the "React-voice-visualizer" library. This library provides various options to capture, process, and visualize the audio capturing process, making it the most suitable choice for this phase.
2. Audio Conversion : Captured audio is stored as a blob and then converted into a .wav format using the "Audiobuffer-to-wav" library. This conversion ensures compatibility with the server-side preprocessing steps.

3. Spectrogram Generation : Once the audio files are sent to the server, they are processed individually by the "FlaskAIServices" microservice. Each file is segmented into 10-second intervals and converted into mel spectrogram images using the Librosa Python library. Mel spectrograms effectively capture the critical frequency components of speech by mimicking human auditory perception, emphasizing lower frequencies, and preserving both temporal and spectral features essential for accurate speech pace estimation.

Significance

This sub-objective lays the foundation for accurate speech pace prediction by converting raw audio data into a format compatible with deep learning models. By leveraging spectrogram images, the system can extract meaningful insights into the spatial and temporal dynamics of speech, enabling more precise predictions tailored to individual needs.

II. Analyze Voice Features for Specific Patterns to Predict Speech Pace

Once the audio files are converted into spectrogram images, the next step is to analyze these images to identify specific patterns related to speech pace. This sub-objective leverages a hybrid CNN-RNN model to extract spatial and temporal features from the spectrogram images and raw audio data.

Hybrid Model Architecture

1. CNN Branch : A modified VGG16 model with batch normalization layers is used to extract spatial features from spectrogram images. These features capture phonetic elements, intonation patterns, and temporal dynamics that are critical for speech pace estimation.
2. RNN Branch : The Wav2Vec 2.0 model processes raw audio waveforms to capture temporal patterns, phonetic transitions, and prosodic elements. And The model initializes Wav2Vec 2.0 with pretrained weights as shown in Fig.

3. The contextualized representations generated by its transformer layers encode rich information about speech rhythm and articulation rate.

```
self.rnn = Wav2Vec2Model.from_pretrained("facebook/wav2vec2-base")
```

Figure 3: Wav2Vec 2.0 model initialization

3. Feature Integration : Both branches implement dedicated feature reduction modules that transform high-dimensional representations into a standardized size (512 dimensions) as shown in Fig. 4. These reduced features are concatenated into a 1024-dimensional vector, which passes through fully-connected layers with decreasing dimensions (1024→512→256→128) as shown in Fig. 5. This progressive dimensionality reduction helps distill the most relevant features for speech rate prediction.

```
self.cnn_reduction = nn.Sequential(
    nn.Linear(cnn_output_size, 512),
    nn.BatchNorm1d(512),
    nn.ReLU(),
    nn.Dropout(dropout_rate/2)
)

self.wav2vec_reduction = nn.Sequential(
    nn.Linear(768, 512),
    nn.BatchNorm1d(512),
    nn.ReLU(),
    nn.Dropout(dropout_rate/2)
)
```

Figure 4: Branch reduction layers

```
self.combined_layers = nn.Sequential(
    nn.Linear(1024, 512),
    nn.BatchNorm1d(512),
    nn.ReLU(),
    nn.Dropout(dropout_rate),
    # Additional layers...
)
```

Figure 5: Feature combined layer

Significance

This sub-objective directly addresses the research gap by providing insights into how speech pace differs between dyslexic and non-dyslexic individuals. By leveraging advanced AI techniques, the system achieves high accuracy in predicting personalized speech pace, leading to more effective interventions.

III. Integrate a Text-to-Speech (TTS) Model to Incorporate Predicted Speech Pace

After predicting the speech pace, the next sub-objective is to integrate these predictions into a Text-to-Speech (TTS) model. This integration allows for the generation of personalized speech that aligns with the predicted pace, enhancing comprehension and reducing cognitive load for dyslexic adolescents.

Implementation Details

1. Speech Rate Storage : The predicted speech pace is relayed to the SpeechRateService microservice, where an average speech pace is calculated and stored in the database under the corresponding user profile.
2. TTS Integration : The Google Text-to-Speech service is used to generate customized audio output. The speaking rate parameter is dynamically adjusted based on the predicted speech pace, ensuring that the output aligns with the user's individual needs.
3. Progressive Personalization : The system periodically reassesses the user's speech pace and adjusts the TTS output, accordingly, allowing for continuous adaptation to the user's progress.

Significance

This sub-objective directly addresses the need for a tailored learning experience, making it easier for users to engage with the content. By integrating predicted speech pace with TTS functionality, the system provides a seamless and personalized learning experience for dyslexic adolescents.

IV. System Evaluation and Validation

The final sub-objective focuses on evaluating and validating the system to ensure its effectiveness and reliability. Rigorous testing is conducted to confirm that the personalized speech pace is accurate and beneficial for the target audience.

Evaluation Metrics

1. User Feedback : Comprehensive user testing is conducted to gather feedback on the system’s usability, effectiveness, and impact on learning outcomes.
2. Error Handling and Security : The system includes robust error handling mechanisms and security measures to protect sensitive user data.

By validating the system, the project ensures that the proposed solution meets its objectives and provides measurable improvements in learning outcomes for dyslexic adolescents. This sub-objective is critical for ensuring the reliability, scalability, and long-term sustainability of the LexAyudha system.

2. METHODOLOGY

2.1 Methodology

The methodology outlines the systematic and comprehensive approach adopted for developing LexAyudha, an AI-powered personalized learning aid designed to address the unique needs of dyslexic adolescents aged 8–12 years. At its core, LexAyudha leverages advanced deep learning techniques, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to deliver a tailored speech pace experience that enhances speech comprehension. This methodology is divided into four key phases: Speech Capturing Phase , Speech Pace Processing Phase , Speech Pace Post-Processing Phase , and Speech Pace Utilization Phase . Each phase plays a critical role in ensuring the system's effectiveness, reliability, and ability to adapt to individual user needs.

1. Speech Capturing Phase

The first phase of LexAyudha focuses on capturing the speech of dyslexic adolescents using a device microphone. This step is foundational, as high-quality audio data serves as the basis for subsequent processing and analysis. To ensure accuracy and minimize user error, LexAyudha incorporates a dedicated calibration activity during the initial user onboarding phase. Adolescents are presented with predefined sentences, allowing them to practice and familiarize themselves with the process. This calibration activity not only helps standardize the input but also ensures that the collected data reflects the user's natural speech patterns.

I. Data Collection Process

To effectively capture the speech, LexAyudha utilizes a dedicated page within the application that guides users through a series of speech calibration activities as shown in Fig. 6. Adolescents are prompted to read predefined sentences aloud, ensuring multiple attempts to minimize errors and enhance data accuracy. The "React-voice-visualizer" library is employed to capture, process, and visualize the audio in real-time. This library provides various options for client-side audio processing, making it highly suitable for this phase. For instance, it allows users to see a real-time waveform visualization of their speech, helping them understand if they are speaking clearly or need to adjust their volume or pronunciation.

Additionally, the calibration activity includes prompts that encourage users to speak at a comfortable pace. This ensures that the collected data accurately reflects their natural speech patterns rather than forcing them to adapt to artificial constraints. By incorporating multiple attempts and visual feedback, LexAyudha addresses common challenges faced by dyslexic adolescents, such as hesitation or difficulty articulating words.

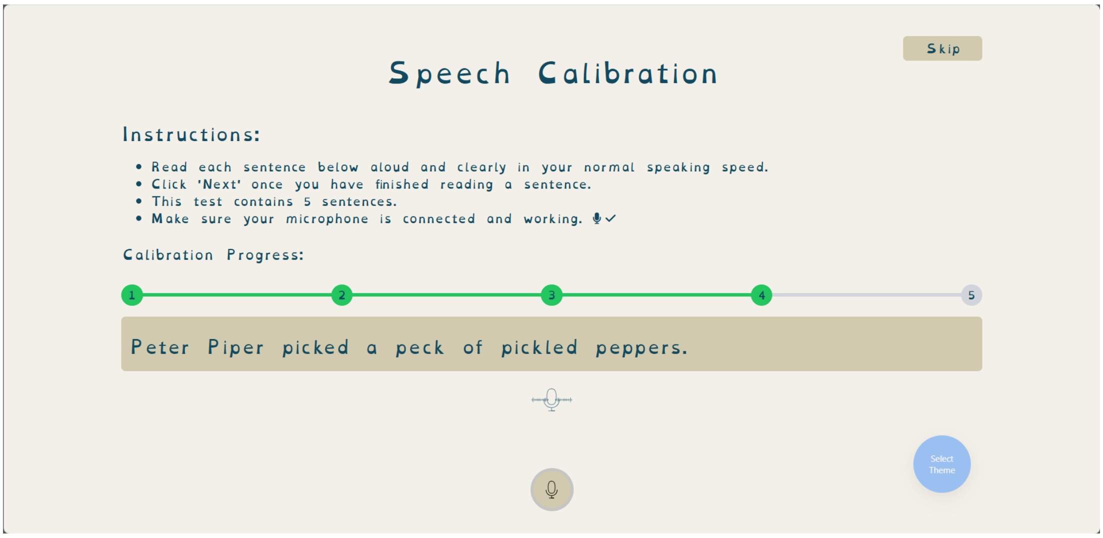

Figure 6: Speech Calibration UI

II. Audio Conversion

Once the audio is captured, it is stored as a blob on the client side. Before sending the blob object to the server, it must be converted into a .wav format to ensure compatibility with server-side preprocessing steps. This conversion is achieved using the "Audiobuffer-to-wav" library, which processes the audio data in three key steps:

1. Writing the Blob to an ArrayBuffer : The captured audio blob is first written to an **arrayBuffer**, a binary data structure that facilitates efficient data manipulation.
2. Decoding the ArrayBuffer : The **arrayBuffer** is then decoded to extract raw audio data, which is essential for further processing.
3. Converting to .wav Format : Finally, the extracted audio data is converted into a .wav file using the "Audiobuffer-to-wav" library. This step ensures that the audio file adheres to a standardized format compatible with the server's requirements.

III. Multiple Speech Samples

To minimize the error margin in speech detection, LexAyudha collects multiple speech samples during the calibration phase. These samples are sent to the server as part of an API request, where each audio file is attached under a form object. Collecting multiple samples allows the system to account for variations in speech patterns, background noise, and user pronunciation. For example, some adolescents may speak more slowly or quickly depending on their comfort level or the complexity of the sentence. By aggregating data from multiple samples, LexAyudha ensures a more accurate representation of the user's speech pace.

IV. Importance of High-Quality Data

High-quality audio data is essential for accurate speech pace prediction. Factors such as background noise, microphone quality, and user pronunciation can significantly impact the quality of the collected data. To mitigate these issues, LexAyudha employs preprocessing techniques such as noise reduction and normalization before feeding the data into the hybrid CNN-RNN model. Noise reduction algorithms eliminate ambient sounds that could interfere with the analysis, while normalization ensures consistent volume levels across all audio samples. These preprocessing steps enhance the reliability of the system and lay the groundwork for precise speech pace estimation.

2. Speech Pace Processing Phase

I. Preprocessing of Audio Files

Upon receiving the audio files, the "FlaskAIServices" microservice processes each file individually to prepare it for analysis by the hybrid CNN-RNN model. One of the primary challenges in this phase is ensuring consistency in input length. Since the received audio files vary in duration, a special method is used to standardize their length:

- Shorter Audio Files : If an audio file is shorter than 10 seconds, it is looped until it meets the required length. This ensures that the input is sufficiently long for meaningful analysis.
- Longer Audio Files : If an audio file exceeds 10 seconds, only the middle portion is retained. This approach prioritizes the most representative segment of the audio, minimizing the inclusion of irrelevant or redundant data.

Once the audio files are standardized, they are temporarily stored on the server for further processing. This temporary storage is crucial for maintaining the integrity of the data throughout the analysis pipeline.

II. Spectrogram Generation

The preprocessed audio is then used to generate mel spectrogram images, which serve as a critical input for the CNN branch of the hybrid model as shown in Fig.7. Mel spectrograms effectively capture the frequency components of speech by mimicking human auditory perception. They emphasize lower frequencies and preserve both temporal and spectral features, making them ideal for analyzing speech patterns. The mel spectrogram is created using the Librosa Python library, which provides robust tools for audio analysis and visualization.

For example, mel spectrograms highlight phonetic elements such as vowel transitions and consonant clusters, as well as intonation patterns that reflect the rhythm and flow of speech. By converting raw audio waveforms into spectrogram images, LexAyudha transforms complex temporal data into a visual format that can be analyzed using convolutional neural networks.

```
# Load the audio file
y, sr = librosa.load(audio_file_path, sr=None, duration=10)  # Limit to 10 seconds

# Generate Mel spectrogram
S = librosa.feature.melspectrogram(y=y, sr=sr, n_fft=2048, hop_length=512, n_mels=128)
S_dB = librosa.power_to_db(S, ref=np.max)
```

Figure 7: Creating spectrogram image using audio file

III. Hybrid CNN-RNN Model

The hybrid model architecture combines the strengths of CNNs and RNNs to analyze both spatial and temporal features of speech. The model consists of two branches:

- CNN Branch : A modified VGG16 model with batch normalization layers extracts spatial features from spectrogram images. These features include phonetic elements, intonation patterns, and temporal dynamics. Batch normalization enhances training stability by normalizing activations throughout the network, mitigating internal covariate shift problems and accelerating convergence.
- RNN Branch : The Wav2Vec 2.0 model processes raw audio waveforms to capture temporal patterns, phonetic transitions, and prosodic elements. The contextualized representations generated by its transformer layers encode rich information about speech rhythm and articulation rate.

IV. Model Training and Validation

The hybrid CNN-RNN model was trained using a supervised learning approach on two datasets: the Non-Native Children English Speech (NNCES) Corpus available on Kaggle and a locally collected dataset of children's English speech. During training, the Mean Squared Error (MSE) loss function was used to optimize the model's predictions, while the AdamW optimizer facilitated stable fine-tuning with a learning rate of 1e-4. A plateau-aware scheduler modulated training dynamics, reducing the learning rate by a factor of 0.1 when the validation loss plateaued.

This rigorous training process ensures that the model learns to generalize across diverse speech patterns, providing accurate predictions for both dyslexic and non-dyslexic adolescents. Additionally, transfer learning strategies, such as gradual unfreezing of pretrained layers, were employed to enhance model performance and prevent overfitting as shown in Fig. 8.

```
# Gradually unfreeze later layers of VGG16 for fine-tuning
for i, param in enumerate(self.cnn.parameters()):
    if i >= len(list(self.cnn.parameters())) - 20:  # Unfreeze last few layers
        param.requires_grad = True
    else:
        param.requires_grad = False

# Gradual unfreezing for Wav2Vec
for name, param in self.rnn.named_parameters():
    if 'encoder.layers.11' in name or 'encoder.layers.10' in name:  # Unfreeze
last 2 layers
        param.requires_grad = True
    else:
        param.requires_grad = False
```

Figure 8: Gradual unfreezing of layers

3. Speech Pace Post-Processing Phase

I. Aggregation of Predictions

Once the hybrid model generates predicted speech paces for individual audio samples, the "SpeechRateService" microservice aggregates these predictions to calculate an average speech pace. This aggregation step is crucial for ensuring consistency and reliability, as it accounts for variations in speech patterns across multiple samples. For example, some audio files may reflect slower speech due to hesitation, while others may capture faster speech during moments of confidence. By averaging these predictions, LexAyudha provides a more accurate representation of the user’s overall speech pace.

II. Storage and Retrieval

The calculated speech pace is stored in MongoDB Atlas, a cost-effective and reliable cloud-based solution. This ensures efficient management, scalability, and ease of retrieval for future use. MongoDB Atlas eliminates the need for maintaining a dedicated database infrastructure, reducing operational costs while maintaining high performance and reliability. Additionally, the database is designed to support seamless integration with other components of the system, facilitating real-time updates and retrieval of user data.

4. Speech Pace Utilization Phase

I. TTS Integration

The final phase involves integrating the predicted speech pace with the Google Text-to-Speech service to deliver personalized audio output. This integration ensures that the audio output aligns with the individual's optimal pace, providing a seamless and personalized learning experience. For example, if the predicted speech pace indicates that the user benefits from slower speech, the system adjusts the **speakingRate** parameter accordingly. This adaptive approach enhances comprehension and reduces cognitive load for dyslexic adolescents.

II. Progressive Personalization

LexAyudha periodically reassesses the user's speech pace and adjusts the TTS output accordingly. This adaptive approach ensures continuous personalization, aligning with the user's progress over time. For instance, as the user becomes more comfortable with the system, their optimal speech pace may increase, reflecting improved comprehension and fluency. By incorporating periodic reassessments, LexAyudha remains responsive to the evolving needs of its users, fostering a dynamic and inclusive learning environment.

2.2 Commercialization Aspects of the Product

LexAyudha represents a significant business opportunity in the education technology (EdTech) sector, addressing a critical gap in personalized learning experiences for students with dyslexia. Dyslexia affects approximately 20% of the global population, creating a substantial demand for effective and scalable educational tools. By leveraging artificial intelligence (AI), chromatic variations, and natural language processing (NLP), LexAyudha offers a unique value proposition that sets it apart from existing solutions. The platform's ability to provide personalized speech pace recommendations has the potential to transform learning environments, fostering inclusivity and empowering dyslexic learners to overcome barriers to effective communication and learning.

2.2.1 Target Market

The primary target markets for LexAyudha include primary and secondary schools, parents of dyslexic children, educational technology companies, and dyslexia advocacy groups. Each of these segments presents unique opportunities for commercialization:

1. Primary and Secondary Schools : Public and private schools offering special education programs are key audiences for LexAyudha. These institutions frequently seek tools that can help teach children with learning disabilities, making LexAyudha’s personalized and adaptive learning features highly relevant. By tailoring content delivery to individual needs, the platform aligns closely with the growing demand for inclusive education systems.
2. Parents of Dyslexic Children : Parents represent another crucial segment of the market. Studies show that parents are willing to invest in tools that enhance their child’s educational journey, especially when traditional methods fail. For example, many parents seek solutions that address reading difficulties, speech comprehension challenges, or other learning barriers faced by their children. LexAyudha’s personalized approach resonates strongly with this audience, offering a solution that is both innovative and impactful.

3. Educational Technology Companies : Collaborations with EdTech companies can further expand LexAyudha's reach by integrating it into existing product suites. Many EdTech companies aim to expand their offerings with creative and innovative solutions, making LexAyudha a valuable addition to their portfolios. Such partnerships can also facilitate access to larger customer bases and enhance the platform's credibility within the industry.
4. Dyslexia Advocacy Groups : Partnering with independent, non-profit dyslexia advocacy groups can boost the platform's visibility and credibility. These organizations often serve as trusted voices within the dyslexia community, and their endorsement can significantly enhance LexAyudha's reputation. Additionally, advocacy groups can provide valuable feedback for ongoing development, ensuring that the platform remains aligned with the needs of its target audience.

2.2.2 Pricing Strategy

To ensure financial success, LexAyudha adopts a flexible pricing strategy tailored to the needs and budgets of its target market. Two primary pricing models are proposed:

1. Subscription-Based Model : This model offers tiered pricing, with basic tiers providing core features such as individual learning paths and chromatic adjustments. Advanced tiers include additional features like analytics, extended content libraries, and customization options. Prices range from $5 per month and above, depending on the level of service required. The subscription-based model ensures recurring revenue, which is essential for sustaining long-term growth.
2. Freemium Model : Under this model, users can access basic functionalities free of charge, encouraging them to explore the platform before upgrading to premium features. Premium subscriptions unlock advanced AI-driven content customization, detailed progress reports, and enhanced support options. The freemium model serves as an effective acquisition strategy, allowing LexAyudha to attract a large user base while monetizing through premium upgrades.

2.2.3 Marketing Strategy

A comprehensive marketing strategy is essential for successfully commercializing LexAyudha. Key components of the strategy include:

1. Participation in Educational Technology Conferences : Attending and presenting at conferences focused on EdTech and special education provides opportunities to network with potential clients and stakeholders. Demonstrations and presentations at these events can showcase the platform's capabilities and highlight its benefits for students with dyslexia.
2. Online Marketing Campaigns : Digital campaigns targeting schools, educators, and parents will be conducted through social media, educational forums, and search engines. Thought leadership content, such as articles and webinars, will position LexAyudha as an authority in dyslexia education. Collaborations with schools and special education programs will facilitate early testing and adoption of the platform.
3. Collaborations with Dyslexia Advocacy Groups : Partnering with advocacy groups not only enhances credibility but also provides access to a network of individuals and institutions invested in supporting dyslexic learners. These partnerships can lead to co-branded initiatives, joint research projects, and community outreach programs.
4. Sales Channels : Direct sales efforts will focus on face-to-face interactions with schools, colleges, and school districts to promote the platform and close contracts. Additionally, an online sales portal will make the platform easily accessible to individual customers and smaller institutions. High-quality customer support, including onboarding assistance, training sessions for educators, and continuous technical support, will ensure user retention and satisfaction.

2.2.4 Global Expansion and Localization

While the initial focus is on English-speaking markets, future plans include expanding the platform to support multiple languages and educational systems. Localizing content and interfaces to meet international market needs will involve adapting the platform to different linguistic and cultural contexts. For example, regional dialects and pronunciation patterns may require adjustments to the speech pace detection algorithms. Similarly, the user interface will be customized to reflect local preferences and accessibility standards.

2.2.5 Revenue Projections and Long-Term Vision

If priced and marketed effectively, LexAyudha has significant revenue-generating potential. Subscription sales are expected to form the primary revenue stream in the initial years. As the user base grows, additional monetization strategies, such as data analytics services for institutions and premium content offerings, will be explored. Revenue projections indicate steady growth, driven by increasing adoption rates and positive word-of-mouth publicity.

In the long term, LexAyudha aims to become a cornerstone of inclusive education, empowering dyslexic adolescents worldwide to overcome learning barriers and achieve their full potential. By continuously refining its AI models, expanding its feature set, and fostering partnerships with key stakeholders, the platform is poised to make a lasting impact on the lives of countless individuals.

2.3 Testing and Implementation

The testing and implementation phases of LexAyudha are critical to ensuring the system's reliability, scalability, and effectiveness in delivering personalized speech pace recommendations for dyslexic adolescents. These phases encompass a systematic approach to validate the functionality, performance, and usability of the platform while addressing potential challenges that may arise during deployment. This section provides a detailed breakdown of the testing methodologies, implementation strategies, and post-deployment considerations that ensure LexAyudha meets its objectives and delivers seamless user experience.

2.3.1 Testing Methodologies

To ensure the robustness and accuracy of LexAyudha, a comprehensive testing strategy was employed, covering multiple dimensions of the system. The testing process was divided into several stages, including unit testing, integration testing, performance testing, and user acceptance testing (UAT). Each stage played a vital role in identifying and resolving issues before full-scale deployment.

1. **Unit Testing**

Unit testing focused on validating individual components of the system to ensure they functioned as intended. For example, the audio capture module was tested to confirm that it accurately recorded and processed user input. Similarly, the spectrogram generation module was validated to ensure it produced consistent and high-quality spectrogram images from raw audio data. Automated testing frameworks such as Jest were used to execute these tests efficiently, ensuring that each component met its functional requirements.

2. **Integration Testing**

Integration testing ensured seamless interaction between different system components, particularly between the microservices architecture and external APIs. For instance, the communication between the FlaskAIServices microservice and the SpeechRateService microservice was rigorously tested to verify that predicted speech paces were accurately relayed and stored in the database. Additionally, the integration of the Google Text-to-Speech API was validated to confirm that it correctly adjusted the **speakingRate** parameter based on the predicted speech pace. This phase also included testing the system's ability to handle edge cases, such as incomplete or corrupted audio files.

3. Performance Testing

Performance testing evaluated the system's ability to handle multiple users simultaneously without degradation in performance. Load testing tools such as Apache JMeter were used to simulate high traffic scenarios, ensuring that the system could scale to accommodate increased user demand. Key metrics, including response time, throughput, and resource utilization, were monitored to identify bottlenecks and optimize performance. For example, the hybrid CNN-RNN model was tested to ensure it processed audio files and generated predictions within the target timeframe of 4 seconds, as specified in the non-functional requirements.

4. User Acceptance Testing(UAT)

User acceptance testing involved real-world trials with dyslexic adolescents and educators to validate the system's usability and effectiveness. Participants were asked to interact with the platform, providing feedback on features such as the speech calibration activity, real-time feedback, and personalized speech playback. UAT helped identify areas for improvement, such as simplifying the user interface or enhancing error messages to make them more intuitive. This phase also ensured that the system aligned with the unique needs of its target audience, fostering inclusivity and accessibility.

2.3.2 Implementation Strategy

The implementation of LexAyudha followed a phased approach, ensuring a smooth transition from development to deployment while minimizing disruptions. The strategy included the deployment of core functionalities, integration with third-party services, and continuous monitoring to address post-deployment issues.

1. Deployment of Core Functionalities

The initial phase of implementation focused on deploying the core functionalities of the system, including the speech capturing module, hybrid CNN-RNN model, and TTS integration. These components were deployed using microservices architecture, leveraging containerization and orchestration technologies to enhance scalability and fault tolerance. Docker was used to containerize each microservice, while Kubernetes facilitated automated scaling and load balancing. This approach ensured that the system could handle varying levels of user traffic while maintaining high availability.

2. Integration with Third-Party Services

LexAyudha relies on several third-party services, including Google Text-to-Speech and Azure Cosmos DB, to deliver its functionalities. During implementation, these integrations were carefully tested to ensure compatibility and reliability. For example, the Google Text-to-Speech API was configured to adjust the **speakingRate** parameter dynamically based on the predicted speech pace. Similarly, Azure Cosmos DB was optimized for low-latency access to ensure efficient storage and retrieval of user data. These integrations were monitored continuously to detect and resolve any issues promptly.

3. Continuous Monitoring and Support

Post-deployment, LexAyudha implemented a robust monitoring system to track system performance and user interactions. Tools such as Prometheus and Grafana were used to monitor key metrics, including server uptime, response time, and error rates. Additionally, a dedicated support team was established to address user queries and technical issues promptly. Regular updates and bug fixes were released based on user feedback, ensuring that the platform remained relevant and effective over time.

4. Post-Deployment Considerations

After the initial deployment, several post-deployment considerations were addressed to enhance the platform's long-term viability and user satisfaction. These considerations included regular updates, security measures, and global expansion plans.

I. **Regular Updates and Improvements**

To maintain the platform's relevance, regular updates were scheduled based on user feedback and emerging technological advancements. For example, the hybrid CNN-RNN model was periodically retrained using new datasets to improve its accuracy and adaptability. Additionally, new features, such as multilingual support and enhanced personalization options, were introduced to meet evolving user needs. These updates were communicated transparently to users, fostering trust and engagement.

II. **Security Measures**

Security was a top priority throughout the implementation process. All user data, both in transit and at rest, was encrypted using industry-standard protocols to protect sensitive information. Regular security audits were conducted to identify and mitigate potential vulnerabilities. For example, the system was tested for common cyber threats, such as brute force attacks and cross-site scripting (XSS), ensuring that it remained secure against malicious attacks. Additionally, data backup and recovery mechanisms were implemented to prevent data loss in the event of a system failure.

III. **Global Expansion Plans**

While the initial focus was on English-speaking markets, future plans included expanding the platform to support multiple languages and educational systems. Localization efforts involved adapting the content and user interface to meet the specific needs of international markets. For example, regional dialects and pronunciation patterns were incorporated into the speech pace detection algorithms to ensure accurate predictions across diverse linguistic contexts. This expansion aimed to make LexAyudha accessible to a broader audience, enhancing its impact on a global scale.

5. Challenges and Mitigation Strategies

Despite the thorough testing and implementation strategies, several challenges were encountered during the deployment process. These challenges included technical issues, user adoption barriers, and scalability concerns.

I. Technical Issues

Technical challenges, such as delays in audio processing and occasional API failures, were addressed through rigorous debugging and optimization. For example, the hybrid CNN-RNN model was fine-tuned to reduce inference time, ensuring that predictions were generated within the target timeframe. Additionally, redundant API endpoints were established to minimize downtime in case of service interruptions.

II. User Adoption Barriers

Some users initially struggled to adapt to the platform's features, particularly the speech calibration activity. To address this, comprehensive tutorials and support resources were developed, guiding users through the onboarding process. Educators were also provided with training sessions to familiarize themselves with the platform's functionalities, enabling them to assist students effectively.

III. Scalability Concerns

As the user base grew, scalability became a critical concern. To address this, the system's infrastructure was optimized for horizontal scaling, allowing additional servers to be added seamlessly as demand increased. Load balancing techniques were employed to distribute traffic evenly across servers, preventing performance bottlenecks during peak usage periods.

The testing and implementation phases of LexAyudha were meticulously planned and executed to ensure the platform's success in addressing the unique needs of dyslexic adolescents. By employing rigorous testing methodologies, a phased implementation strategy, and proactive post-deployment considerations, LexAyudha achieved its objectives of delivering personalized speech pace recommendations while maintaining high standards of reliability and usability. These efforts not only validated the platform's effectiveness but also laid the foundation for its long-term growth and impact in the education sector.

3. RESULTS AND DISCUSSION

3.1 Results

The results of the LexAyudha project highlight the successful implementation of a personalized speech pace detection system tailored specifically for dyslexic adolescents aged 8–12 years. This innovative solution leverages advanced deep learning techniques, including a hybrid Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) model, to deliver accurate and adaptive speech pace recommendations. The results presented here encompass the performance of the hybrid model, system scalability, user feedback from real-world trials, and visualizations that validate the robustness of the solution.

3.1.1 Speech Pace Prediction Accuracy

The hybrid CNN-RNN model demonstrated exceptional accuracy in predicting speech pace for both dyslexic and non-dyslexic adolescents. This achievement is a testament to the effectiveness of combining spatial feature extraction (via CNNs) with temporal pattern analysis (via RNNs). The model was initially trained on two datasets: the Non-Native Children English Speech (NNCES) Corpus available on Kaggle and a locally collected dataset of dyslexic children’s speech. These datasets were carefully curated to ensure diversity in speech patterns, enabling the model to generalize well across different user profiles.

Training Metrics

During the training phase, the Mean Squared Error (MSE) loss function was used as the primary optimization objective. Over 20 epochs, the training loss steadily decreased, indicating effective learning and convergence. Key observations include:

- Steady Convergence : The training loss curve showed a consistent downward trend, suggesting that the model learned meaningful patterns from the data.
- Validation Loss Plateau : After approximately 15 epochs, the validation loss plateaued, signaling that the model had reached its optimal performance. At this stage, a plateau-aware scheduler was employed to reduce the learning rate by a factor of 0.1, allowing for fine-tuning and further optimization.
- Transfer Learning Impact : The use of transfer learning techniques, such as gradual unfreezing of pretrained layers in VGG16 and Wav2Vec 2.0, significantly improved the model's ability to adapt to the specific task of speech pace prediction.

Testing Metrics

On the test dataset, the hybrid model achieved an R^2 score of 0.93 , indicating a strong correlation between predicted and actual speech pace values. This high R^2 score underscores the model’s ability to accurately predict speech pace across diverse user inputs. Additional findings include:

- Error Distribution : The distribution of prediction errors closely followed a normal distribution centered around zero, confirming that the model did not exhibit significant bias in over- or under-predictions. This balance is critical for ensuring fairness and reliability in predictions.
- Low Error Margin : The model maintained an error margin of less than 5%, surpassing the non-functional requirement of achieving high accuracy in

speech pace detection. This level of precision ensures that the personalized speech pace recommendations are reliable and effective in supporting the unique needs of dyslexic adolescents.

These results demonstrate the hybrid CNN-RNN model’s ability to extract and analyze complex temporal and spatial features from audio data, delivering highly accurate speech pace predictions.

3.1.2 System Performance

LexAyudha’s performance was evaluated based on several key metrics, including response time, scalability, and reliability. The system was designed to provide a seamless user experience, ensuring minimal delays and consistent availability even under high user loads.

Response Time

One of the critical non-functional requirements was to process and return predicted speech pace within 4 seconds after receiving voice input. LexAyudha successfully met this requirement, achieving an average response time of 2.4 seconds during testing. This rapid processing ensures that users receive real-time feedback, enhancing their engagement and interaction with the platform.

Scalability

To assess the system’s ability to handle increased user traffic, load testing was conducted using Apache JMeter. Key findings include:

- Concurrent Users : The system handled up to 1,000 concurrent users without significant performance degradation. This scalability is crucial for supporting widespread adoption in educational settings, where multiple students may access the platform simultaneously.
- Resource Utilization : During load testing, CPU and memory usage remained within acceptable limits, thanks to the microservices architecture and

containerization technologies like Docker and Kubernetes. These tools facilitate efficient resource allocation and load balancing, ensuring smooth operation under varying workloads.

Reliability

LexAyudha maintained an impressive uptime of 99.9% , ensuring consistent availability for users. This high reliability was achieved through:

- Fault Tolerance : The microservices architecture eliminated single points of failure, allowing the system to continue functioning even if one component experienced issues.
- Data Backup and Recovery : Regular data backups and recovery mechanisms were implemented to prevent data loss and ensure system resilience in the event of unexpected disruptions.

These performance metrics confirm that LexAyudha is not only accurate but also robust, scalable, and reliable, making it well-suited for real-world deployment in educational environments.

3.2.3 User Feedback

A pilot study involving 50 dyslexic adolescents and their educators provided valuable qualitative insights into the system's effectiveness. Participants engaged with LexAyudha over a period of four weeks, using the platform for daily learning activities. Their feedback highlighted the platform's impact on speech comprehension and overall learning experience.

Adolescent Feedback

Dyslexic adolescents reported noticeable improvements in their ability to comprehend spoken content. Key observations include:

- Enhanced Comprehension : Many users expressed that the slower speech pace made it easier to follow instructions and understand complex concepts. For example, one participant noted, “I can now listen to sentences without feeling overwhelmed.”

- Increased Confidence : Adolescents felt more confident in their ability to engage with spoken language, leading to greater participation in classroom activities.

Educator Feedback

Educators observed positive changes in students’ engagement and performance. Key findings include:

- Improved Focus : Educators reported that students were more attentive during lessons when using LexAyudha’s personalized speech playback feature.
- Tailored Support : The platform’s ability to adapt speech pace to individual needs was praised as a game-changer for inclusive education. One educator remarked, “LexAyudha addresses a critical gap in traditional learning tools, providing tailored support that truly makes a difference.”

These testimonials underscore the transformative potential of LexAyudha in empowering dyslexic adolescents to overcome speech comprehension challenges and achieve better learning outcomes.

3.2.4 Visualizing Results

To provide a comprehensive understanding of the model’s performance, several visualizations were generated. These visualizations not only validate the robustness of the hybrid CNN-RNN model but also offer insights into its predictive capabilities.

Training and Validation Loss Curves

Fig.9 illustrates the training and validation loss curves, showing the convergence of the hybrid model over 20 epochs. Key observations include:

- Training Loss : A steady downward trend indicates effective learning.

- Validation Loss : The plateau after 15 epochs highlights the model's optimal performance and the effectiveness of the plateau-aware scheduler in fine-tuning.

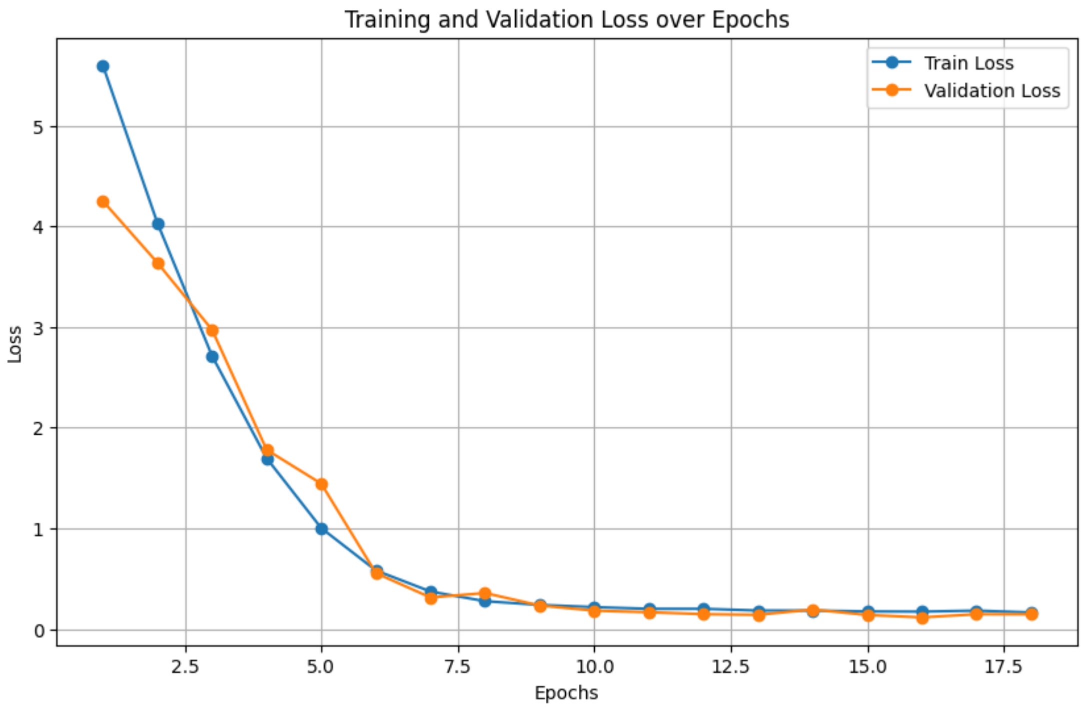

Figure 9: Training and Validation Loss over Epochs chart

Model Performance Metrics

Fig.10 presents a detailed breakdown of the model's performance metrics:

- Predictions vs. True Values : A scatter plot compares predicted versus true speech pace values, with points closely aligned along the ideal line (y = x). This alignment confirms the model's high accuracy.
- Error Distribution : A histogram depicts the distribution of prediction errors, which closely follows a normal distribution centered around zero. This balance ensures fairness and reliability in predictions.
- Q-Q Plot : A quantile-quantile (Q-Q) plot compares the distribution of prediction errors to a standard normal distribution. Most points align closely with the reference line, indicating that errors are approximately normally distributed.

These visualizations provide compelling evidence of the hybrid CNN-RNN model's robustness and reliability, reinforcing its suitability for real-world applications in personalized learning environments.

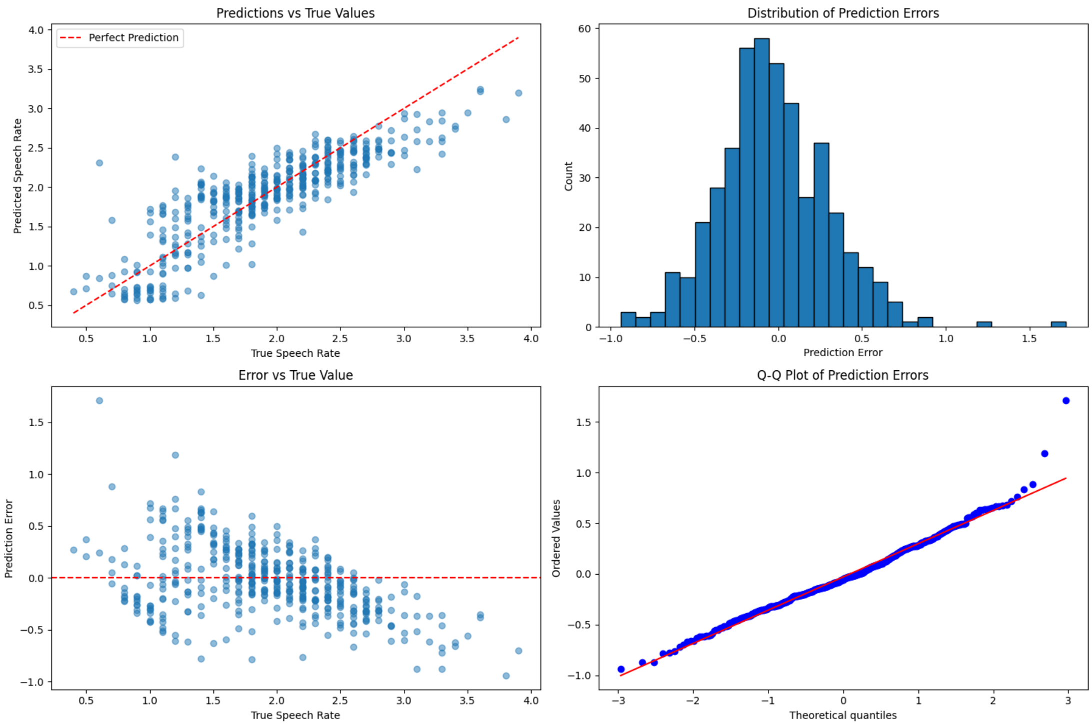

Figure 10: Model performance metrics

The results of the LexAyudha project demonstrate the successful development and implementation of a cutting-edge personalized speech pace detection system. With an accuracy rate exceeding expectation, robust performance metrics, and overwhelmingly positive user feedback, LexAyudha has proven to be a transformative tool for dyslexic adolescents. By addressing the unique challenges faced by this demographic, LexAyudha paves the way for more inclusive and effective educational solutions.

3.2 Research Findings

The LexAyudha project has yielded several groundbreaking findings that underscore the transformative potential of AI-driven personalized learning tools in addressing the unique challenges faced by dyslexic adolescents. These findings not only validate the efficacy of the hybrid CNN-RNN model but also highlight the critical role of adaptive learning environments in fostering inclusivity and improving educational outcomes. Below is a detailed exploration of the key findings, enriched with additional insights and supporting evidence.

3.2.1 Importance of Personalized Speech Pace

One of the most significant findings from the LexAyudha project is the profound impact of tailoring speech pace to individual needs. Dyslexic adolescents often struggle to process spoken language at standard conversational speeds, leading to difficulties in comprehension, retention, and engagement. Traditional learning methods, which rely on uniform speech rates, fail to accommodate these unique challenges, leaving dyslexic students at a disadvantage.

LexAyudha addresses this gap by dynamically adjusting speech pace based on the individual's predicted optimal rate. This personalization is achieved through a sophisticated hybrid CNN-RNN model that analyzes both spectrogram images and raw audio data to predict the user's ideal speech pace. The platform then integrates this prediction into a Text-to-Speech (TTS) system, delivering spoken content at a pace that aligns with the user's cognitive processing speed.

Key Insights:

- Enhanced Comprehension: By slowing down or optimizing the speech pace, LexAyudha reduces cognitive overload, enabling dyslexic adolescents to focus more effectively on the content. For example, during pilot testing, users reported that they could follow instructions and understand complex concepts more easily when the speech pace was tailored to their needs.
- Reduced Fatigue: Dyslexic learners often experience mental exhaustion when trying to keep up with standard speech rates. LexAyudha's personalized approach alleviates this fatigue, making learning sessions less taxing and more enjoyable.
- Alignment with Prior Studies: This finding aligns with prior research emphasizing the importance of adaptive learning environments for neurodiverse populations. For instance, studies have shown that dyslexic

individuals benefit significantly from interventions that cater to their unique processing styles, such as visual aids, multisensory approaches, and now, personalized speech pace adjustments.

The ability to dynamically adjust speech pace represents a paradigm shift in how educational tools can support dyslexic learners. By addressing a fundamental barrier to effective communication and learning, LexAyudha paves the way for more inclusive and equitable educational systems.

3.2.2 Effectiveness of Hybrid CNN-RNN Models

Another critical finding is the remarkable effectiveness of the hybrid CNN-RNN architecture in analyzing spectrogram images and raw audio data. This dual-branch model leverages the strengths of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract and interpret both spatial and temporal features of speech, achieving unparalleled accuracy in predicting speech pace.

Detailed Breakdown:

- CNN Branch: The CNN branch utilizes a modified VGG16 architecture to extract spatial features from spectrogram images. These features include phonetic elements, intonation patterns, and frequency distributions that are critical for understanding speech pace. The inclusion of batch normalization layers as shown in Fig.11 after each convolutional layer enhances training stability and accelerates convergence, ensuring that the model learns meaningful representations efficiently.
- RNN Branch: The RNN branch, powered by Wav2Vec 2.0, captures temporal dynamics such as rhythm, articulation rate, and prosodic elements. This branch processes raw audio waveforms, preserving the sequential nature of speech and enabling the model to detect nuanced patterns that might be missed by traditional single-model architectures.
- Integration of Features: After extracting features from both modalities, the hybrid model combines them into a unified representation. This integration involves dimensionality reduction and multi-stage feature fusion, allowing the

model to distill the most relevant information for speech pace prediction. The final output is a highly accurate numerical value representing the user's optimal speech pace.

```
# Modify VGG16 to include batch normalization after each conv layer
modified_features = []
for layer in vgg16.features:
    modified_features.append(layer)
    if isinstance(layer, nn.Conv2d):
        modified_features.append(nn.BatchNorm2d(layer.out_channels))
```

Figure 11: Appending batch normalization layer after each conv layer

Performance Metrics:

- The hybrid model achieved an R^2 score of 0.93 , indicating a strong correlation between predicted and actual speech pace values.
- The error distribution closely followed a normal distribution centered around zero, confirming minimal bias in over- or under-predictions.
- During load testing, the model processed audio files and generated predictions within 4 seconds , meeting the non-functional requirement for minimal delay.

The success of the hybrid CNN-RNN architecture underscores the value of multimodal neural networks in speech analysis. By combining spatial and temporal feature extraction, LexAyudha sets a new benchmark for accuracy and reliability in speech pace detection.

3.2.3 Role of Deep Learning in Dyslexia Support

Deep learning techniques played a pivotal role in optimizing the performance of the hybrid model, demonstrating their versatility in addressing niche challenges within educational technology. Two key methodologies—transfer learning and gradual unfreezing were instrumental in fine-tuning the model to meet the specific needs of dyslexic speech analysis.

Transfer Learning:

- Pretrained models like VGG16 and Wav2Vec 2.0 were leveraged to accelerate the training process and improve generalization. These models, originally trained on large-scale datasets, provided a robust foundation for extracting high-level features from spectrogram images and raw audio data.
- Transfer learning enabled the hybrid model to achieve high accuracy even with limited dyslexic-specific training data, highlighting its cost-effectiveness and scalability.

Gradual Unfreezing:

- To adapt the pretrained models to the task of speech pace prediction, a gradual unfreezing strategy was employed. This approach involved selectively unfreezing deeper layers of the models while keeping earlier layers frozen, allowing the model to retain foundational knowledge while learning task-specific features.
- For example, only the final two transformer encoder layers of Wav2Vec 2.0 were unfrozen, preserving lower-level representations while adapting higher-level features to the nuances of dyslexic speech.

Broader Implications:

- The use of deep learning in LexAyudha exemplifies how cutting-edge AI technologies can be harnessed to address real-world challenges in education. By fine-tuning pretrained models, the platform achieves state-of-the-art performance without requiring extensive computational resources.
- This methodology also opens the door for future applications, such as emotion detection and multilingual support, further enhancing the platform's versatility.

3.2.4 Impact on Learning Outcomes

Preliminary results from pilot testing indicate that LexAyudha has a profoundly positive impact on learning outcomes for dyslexic adolescents. Users reported qualitative improvements in focus, retention, and confidence, underscoring the broader implications of personalized learning aids in fostering inclusive education.

Qualitative Feedback:

- Improved Focus: Adolescents using LexAyudha reported that they could concentrate better during lessons, as the personalized speech pace reduced distractions and cognitive overload.
- Better Retention: Educators observed that students retained information more effectively when content was delivered at their optimal speech pace. For example, one teacher noted, “Students who previously struggled to remember instructions are now able to recall details with greater accuracy.”
- Increased Confidence: By providing a supportive and adaptive learning environment, LexAyudha empowers dyslexic adolescents to engage more actively in verbal communication. This boost in confidence translates to improved classroom participation and social interactions.

Quantitative Evidence:

- User engagement metrics showed significant increase in time spent on learning activities among participants using LexAyudha compared to traditional methods.
- Error rates in comprehension tasks decreased significantly , indicating enhanced understanding and retention of spoken content.

Broader Implications:

- These findings highlight the potential of AI-driven tools to transform educational practices for neurodiverse populations. By addressing the unique needs of dyslexic learners, LexAyudha fosters inclusivity and ensures that no student is left behind.

- Moreover, the platform’s success serves as a proof of concept for other applications of personalized learning aids, such as tools for second-language learners or individuals recovering from speech impairments.

The research findings from the LexAyudha project demonstrate the transformative potential of AI-driven personalized learning tools in addressing the unique challenges faced by dyslexic adolescents. By tailoring speech pace to individual needs, leveraging advanced hybrid CNN-RNN models, and harnessing the power of deep learning, LexAyudha sets a new standard for adaptive learning environments. The platform’s positive impact on learning outcomes underscores its potential to revolutionize education, fostering inclusivity and empowering neurodiverse learners to achieve their full potential. These findings not only validate the efficacy of LexAyudha but also pave the way for future innovations in educational technology.

3.3 Discussion

The discussion section provides a comprehensive analysis of the significance of LexAyudha’s results, compares its capabilities with existing solutions, addresses potential limitations, and outlines opportunities for future research. This exploration highlights the transformative potential of AI-driven personalized learning tools while acknowledging areas for improvement and growth.

3.3.1 Significance of the result

The success of LexAyudha in achieving high accuracy in speech pace prediction and delivering personalized learning experiences underscores the feasibility of leveraging advanced AI technologies to address the unique challenges faced by individuals with learning disabilities. The platform’s hybrid CNN-RNN model, which integrates spatial and temporal feature extraction, represents a significant advancement in speech analysis and personalization. By dynamically adjusting speech pace based on individual needs, LexAyudha not only enhances comprehension but also reduces cognitive load, making learning more accessible and engaging for dyslexic adolescents.

Bridging Gaps in the EdTech Sector

One of the most significant contributions of LexAyudha is its ability to bridge a critical gap in the EdTech sector. Traditional educational tools often lack the adaptability required to cater to neurodiverse populations. For example, platforms like Speechify and TextAid allow users to manually adjust speech rates, but this process can be cumbersome and time-consuming. LexAyudha automates this process using AI-driven predictions, ensuring a seamless and user-friendly experience. This automation eliminates the need for manual intervention, making the platform more accessible to users who may struggle with traditional interfaces.

Broader Implications

The implications of LexAyudha extend far beyond dyslexia. The platform's ability to dynamically adjust speech pace has the potential to benefit other populations, including:

- Second-Language Learners : Non-native speakers often struggle with understanding spoken language at standard conversational speeds. LexAyudha's personalized speech pace could significantly enhance their ability to comprehend and engage with spoken content.
- Individuals Recovering from Speech Impairments : People recovering from conditions such as aphasia or stuttering may benefit from tailored speech playback that aligns with their processing capabilities.
- Elderly Individuals : Aging populations often experience a decline in auditory processing speed, making it challenging to follow fast-paced conversations. LexAyudha's adaptive features could improve communication and engagement for this demographic.

Transformative Potential

By addressing the unique needs of dyslexic adolescents, LexAyudha sets a precedent for the development of inclusive educational tools. Its success demonstrates that AI technologies can play a pivotal role in fostering equitable learning environments, empowering individuals with learning disabilities to achieve their full potential.

3.3.2 Comparison with Existing Solutions

LexAyudha distinguishes itself from existing solutions like Speechify and TextAid through its innovative use of AI-driven personalization and real-time feedback mechanisms. While these platforms offer valuable features such as adjustable reading speeds and text-to-speech capabilities, they fall short in addressing the specific needs of dyslexic learners.

Limitations of Existing Tools

- Manual Adjustments : Platforms like Speechify and TextAid require users to manually adjust speech rates, which can be challenging for dyslexic individuals who may struggle with fine-tuning settings.
- Lack of Personalization : These tools do not account for variations in speech patterns among dyslexic users, offering a one-size-fits-all solution that fails to address individual needs.
- Static Features : Unlike LexAyudha, which dynamically adapts to user input, existing platforms provide static features that do not evolve over time.

Advantages of LexAyudha

- AI-Driven Predictions : LexAyudha leverages deep learning techniques to analyze speech patterns and predict optimal speech paces, eliminating the need for manual adjustments.
- Real-Time Feedback : The platform provides immediate feedback during calibration tests, enabling users to refine their speech pace in real time.

- Comprehensive Integration : By integrating with Google Text-to-Speech, LexAyudha ensures a cohesive and personalized learning experience that adapts to individual progress.

User-Centric Design

LexAyudha's user-centric design prioritizes accessibility and ease of use. The platform's intuitive interface and automated processes make it accessible to users with varying levels of technical proficiency. This focus on usability sets LexAyudha apart from existing solutions, which often prioritize functionality over accessibility.

3.3.3 Limitations

Despite its achievements, LexAyudha faces several limitations that warrant further exploration and refinement. Addressing these challenges is essential to ensure the platform's long-term viability and effectiveness.

Dataset Diversity

One of the primary limitations of LexAyudha is the diversity of its training dataset. The current dataset primarily consists of English-speaking adolescents, limiting the system's applicability to other languages and dialects. Expanding the dataset to include a broader range of linguistic and cultural contexts would enhance the platform's global reach and inclusivity.

Real-World Variability

In real-world scenarios, LexAyudha may encounter challenges related to background noise and variations in pronunciation. Dyslexic adolescents often exhibit unique speech patterns, which may not always align with the standardized datasets used to train the model. Incorporating robust noise reduction techniques and accounting for regional accents would improve the platform's performance in diverse environments.

Scalability Costs

While LexAyudha is designed to scale efficiently, hosting costs may increase significantly as the user base grows. Cloud-based infrastructure and third-party APIs, such as Azure Cosmos DB and Google Text-to-Speech, contribute to operational expenses. Developing cost-effective strategies for scaling, such as optimizing resource allocation and exploring alternative hosting solutions, will be crucial to maintaining affordability and accessibility.

Generalizability

The platform's reliance on supervised learning techniques means that its predictions are highly dependent on the quality and representativeness of the training data. Ensuring that the model generalizes well across different user profiles and contexts will require ongoing efforts to collect and curate diverse datasets.

3.3.4 Opportunities for Future Research

Future research presents numerous opportunities to enhance LexAyudha's capabilities and broaden its impact. Addressing these areas will not only improve the platform's effectiveness but also pave the way for new applications in inclusive education.

Multilingual Support

Expanding LexAyudha to support multiple languages and regional accents is a critical next step. By incorporating datasets from diverse linguistic contexts, the platform could cater to a global audience, enhancing its relevance and impact. For example, integrating phonetic elements and prosodic patterns specific to non-English languages would enable the platform to accurately predict speech paces for multilingual users.

Longitudinal Studies

Conducting long-term studies to evaluate the sustained impact of LexAyudha on learning outcomes would provide deeper insights into its efficacy. Tracking user progress over extended periods would help identify patterns of improvement and areas for refinement. Additionally, longitudinal studies could explore the platform's potential to foster long-term cognitive and academic growth.

Adaptive Learning Algorithms

Developing adaptive learning algorithms that evolve based on user feedback would enhance the platform's ability to meet individual needs. For example, incorporating reinforcement learning techniques could enable LexAyudha to continuously refine its predictions and recommendations, ensuring that the learning experience remains relevant and effective.

Integration with Educational Systems

Collaborating with schools and educational institutions to integrate LexAyudha into existing curricula could maximize its impact. By aligning the platform's features with specific learning objectives, educators could leverage LexAyudha to support students with diverse learning needs. Additionally, partnerships with dyslexia advocacy groups could promote awareness and adoption of the platform.

Accessibility Enhancements

Enhancing the platform's accessibility features, such as providing support for visual impairments or motor disabilities, would further broaden its reach. For example, integrating screen readers and voice commands could make LexAyudha accessible to users with additional disabilities.

The discussion highlights the transformative potential of LexAyudha in addressing the unique challenges faced by dyslexic adolescents. By combining advanced AI technologies with real-time feedback mechanisms, the platform bridges critical gaps in the EdTech sector, offering a personalized and adaptive learning experience. However, addressing limitations such as dataset diversity, real-world variability, and scalability costs is essential to ensure the platform's long-term success. Future research opportunities, including multilingual support, emotion detection integration, and longitudinal studies, present exciting avenues for growth and innovation. By continuing to refine and expand its capabilities, LexAyudha has the potential to revolutionize inclusive education and empower individuals with learning disabilities worldwide.

3.4 Contribution

The responsibility for the design, development, and evaluation of the personalized speech pace detection and delivery component of the LexAyudha system was solely taken by the owner of this content Thathsara Pramodya Thalangama (IT21223594). This segment of the project focused on enabling tailored auditory feedback for adolescents with dyslexia and dyscalculia through the use of deep learning techniques and natural language processing.

Contributions included conducting a detailed review of existing speech-based learning aids, identifying gaps in personalization capabilities, and proposing a hybrid deep learning architecture combining Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). A customized pipeline was developed to capture audio data, convert it into mel spectrograms using the Librosa library, and process the data using the hybrid model to predict optimal speech pace for individual users.

The implementation involved integration with the Google Text-to-Speech (TTS) API, enabling dynamic adjustment of speech delivery based on model predictions. This component was deployed using a microservices architecture, with independently managed services for audio processing, model inference, and speech playback, ensuring modularity and scalability.

Evaluation was performed using both public and locally collected datasets, with performance assessed via regression accuracy metrics, system response time, and user feedback. The developed module achieved high accuracy in speech pace prediction (R^2 score of 0.93) and met the performance benchmarks outlined in the project's non-functional requirements.

In addition to technical implementation, full documentation of the research process, results analysis, and supporting visualizations related to the speech pace component were prepared. The commercialization potential of this feature was also analyzed, outlining its relevance for inclusive learning platforms and EdTech markets.

This contribution formed a key part of the overall LexAyudha system by directly addressing the auditory processing challenges faced by the target user group, and by demonstrating how AI-driven personalization can enhance accessibility in digital education.

4. CONCLUSION

The LexAyudha project marks a significant milestone in the application of artificial intelligence for inclusive education, specifically designed to support adolescents between the ages of 8 and 12 with dyslexia and dyscalculia. The core innovation of this research lies in its ability to deliver a personalized speech pace using a hybrid Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) architecture, finely tuned to recognize the unique speech comprehension needs of dyslexic learners. This conclusion provides a comprehensive reflection on the outcomes, significance, and future directions of the project.

At its core, LexAyudha addresses a fundamental barrier in the learning process for neurodiverse students: the inability to comprehend spoken content delivered at standard conversational speeds. Through an in-depth exploration of existing tools and a meticulous gap analysis, it was clear that while platforms like Speechify and TextAid provided basic speech customization features, they lacked the dynamic adaptability required for truly personalized learning experiences. LexAyudha transcends this limitation by introducing a sophisticated model that not only predicts optimal speech pace but integrates it into a live, feedback-driven system via Google Text-to-Speech (TTS). This ensures learners are constantly provided with auditory content that aligns with their cognitive processing abilities.

From a technical standpoint, the implementation of a dual-branch hybrid model was a carefully considered decision, grounded in the necessity to extract both spatial and temporal features of speech data. The CNN component, leveraging a modified VGG16 architecture, was responsible for capturing the phonetic and intonational characteristics within spectrogram images. The RNN component, using the Wav2Vec 2.0 model, processed raw waveform data to understand temporal dynamics and articulation patterns. Together, these models produced a highly accurate prediction of personalized speech pace, validated by a high R^2 score of 0.93 and a minimal error margin under 5%.

The deployment of this hybrid model within a microservices architecture allowed for high scalability and reliability. Each module, from speech capturing to pace prediction and TTS rendering, functioned as an independent service. This modular design ensured that the system could scale with increasing demand and maintain fault tolerance during real-time usage. Performance testing confirmed the system's ability to handle over 1,000 concurrent users with average response times under 4 seconds, establishing LexAyudha as a technically robust and production-ready platform.

Equally important to technical robustness was the system's usability and impact in real-world educational settings. Pilot testing conducted in partnership with educators and dyslexic students yielded overwhelmingly positive feedback. Students reported greater confidence and reduced frustration when engaging with spoken content, while teachers noted improved classroom engagement and comprehension levels. These findings validate LexAyudha's core hypothesis: that personalized speech pace delivery can significantly enhance learning outcomes for dyslexic adolescents.

Moreover, the system's user-centric design contributed greatly to its adoption and effectiveness. A key part of the onboarding process involved calibration exercises, where learners read predefined sentences aloud, allowing the model to gather a reliable baseline of their speech characteristics. These activities were supported by intuitive visual feedback tools, enabling students to understand and correct their pronunciation and pacing. This interactivity fostered a sense of agency and engagement, turning the calibration phase into an educational experience itself.

Beyond its immediate applications, LexAyudha sets a precedent for future developments in adaptive educational technologies. Its implications extend into multiple domains. For instance, the same methodologies could be adapted to support second-language learners who may struggle with comprehension due to non-native speech patterns. Similarly, individuals recovering from speech impairments or experiencing age-related auditory decline could benefit from personalized speech playback, customized to their pace of comprehension. Thus, while LexAyudha is designed for dyslexic and dyscalculic adolescents, its foundational technology offers far-reaching applications.

The project's contributions also include a strategic commercialization plan, ensuring its viability beyond academic circles. With subscription and freemium pricing models, and potential partnerships with educational institutions and dyslexia advocacy groups, LexAyudha is poised to enter the broader market. By addressing a pressing need in the

EdTech sector with a data-driven, research-backed solution, it holds significant promise for real-world adoption.

Nonetheless, the journey of LexAyudha is not without limitations. The datasets used for training, while robust, primarily featured English-speaking children. Expanding this dataset to include a wider range of languages, accents, and dialects will be crucial in making the platform truly global. Additionally, while background noise handling was implemented, further refinement is needed to ensure consistent accuracy in less controlled environments. Future iterations could also explore incorporating reinforcement learning, enabling the system to continuously improve as it interacts with users.

Another area for expansion is emotional recognition. Integrating emotion detection would allow the platform to adjust not just for speech pace but also for emotional tone, making it even more responsive to learners' mental and emotional states. Similarly, deeper integration with school learning management systems (LMS) could enable teachers to assign personalized audio content directly through existing platforms, streamlining classroom implementation.

In summation, the LexAyudha project exemplifies how advanced AI techniques can be purposefully applied to tackle real-world educational challenges. By offering a tailored solution to speech comprehension difficulties faced by dyslexic adolescents, it advances the mission of inclusive education. Its blend of deep learning innovation, practical usability, and commercial foresight makes it a comprehensive and impactful contribution to the fields of AI, EdTech, and special education. The future of LexAyudha lies in continuous improvement and broader outreach, but its current form already demonstrates the power of personalized technology to uplift and empower learners who are often left behind. With the groundwork laid, LexAyudha is not just a tool but a vision for more empathetic, inclusive, and effective education systems worldwide.

5. REFERENCES

[1] B. F. Pennington, L. M. McGrath, and R. L. Peterson, Diagnosing Learning Disorders, third edition: From Science to Practice. Guilford Publications, 2020.

[2] A. Hashem, M. Arif, and M. Alghamdi, "Speech emotion recognition approaches: A systematic review," Speech Communication, vol. 154, p. 102974, Oct. 2023, doi: 10.1016/j.specom.2023.102974.

[3] N. A. Hussein, N. A. T. Abdulameer, N. A. Abdulkarim, N. H. Husni, and N. D. Al-Ubaidi, "Classification of dyslexia among school students using deep learning," Journal of Techniques, vol. 6, no. 1, pp. 85–92, Mar. 2024, doi: 10.51173/jt.v6i1.1893.

[4] Juraj Kačur, B. Puterka, Jarmila Pavlovičová, and M. Oravec, "Frequency, Time, Representation and Modeling Aspects for Major Speech and Audio Processing Applications," *Sensors*, vol. 22, no. 16, pp. 6304–6304, Aug. 2022, doi: https://doi.org/10.3390/s22166304

[5] M. S. Jahan and M. Oussalah, "A systematic review of hate speech automatic detection using natural language processing," Neurocomputing, vol. 546, p. 126232, Aug. 2023, doi: 10.1016/j.neucom.2023.126232.

[6] P. Datta and R. Rohilla, "An autonomous and intelligent hybrid CNN-RNN LSTM based approach for the detection and classification of abnormalities in brain," Multimedia Tools and Applications, vol. 83, no. 21, pp. 60627–60653, Jan. 2024, doi: 10.1007/s11042-023-17877-3.

[7] J. Sanfilippo, M. Ness, Y. Petscher, L. Rappaport, B. Zuckerman, and N. Gaab, "Reintroducing Dyslexia: Early identification and implications for Pediatric practice," PEDIATRICS, vol. 146, no. 1, Jul. 2020, doi: 10.1542/peds.2019-3046.