LEXAYUDHA : AI-BASED PERSONALIZED REHABILITATION FOR DYSLEXIA AND DYSCALCULIA ADOLESCENTS

Project ID : 24-25J_233

Thalangama T.P.

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

August 2024

LEXAYUDHA : AI-BASED PERSONALIZED REHABILITATION FOR DYSLEXIA AND DYSCALCULIA ADOLESCENTS

Project Proposal Report

Project ID : 24-25J_233

Thalangama T.P.
IT21223594

Bachelor of Science (Hons) in Information Technology Specializing in Software Engineering

Department of Computer Science & Software Engineering

August 2024

DECLARATION

We declare that this is our own work, and this proposal 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 our 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.

Name	Student ID	Signature
Thalangama T.P	IT21223594	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date: 2024.08.23

Abstract

Dyslexia and dyscalculia are prevalent learning disorders that significantly impair an individual's ability to comprehend written and spoken language. These disorders not only involve difficulties with reading but also pose substantial challenges in speech comprehension. Adolescents with dyslexia and dyscalculia often struggle to process spoken language at typical conversational speeds, leading to difficulties in understanding and following verbal instructions. Addressing these challenges is critical, as speech comprehension is fundamental to effective communication and learning. The overall purpose of this study is to develop a deep learning-based tool that customizes speech pace to improve speech comprehension for dyslexic and dyscalculic adolescents aged 8–12 years. This study explores the implementation of a hybrid AI model that integrates Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to analyze spectrogram images derived from speech audio data. The methodology involves collecting speech audio through a series of calibration tests, converting the audio into spectrogram images, and using a CNN to extract spatial features related to speech pace. These features are then stored in a vector map and further analyzed by the RNN-GRU model to estimate the optimal speech pace. The estimated speech pace is stored in a database for use in a Text-to-Speech (TTS) system. By utilizing this hybrid model approach, the study aims to leverage the distinct advantages of CNNs in extracting relevant visual patterns from spectrogram images and the temporal processing capabilities of RNNs. The integration of these technologies is expected to yield a highly accurate speech pace prediction that aligns closely with the individual needs of dyslexic and dyscalculic adolescents. This tailored approach has the potential to significantly enhance speech comprehension by providing a personalized learning experience. Furthermore, the study's findings could contribute to broader educational tools designed to assist individuals with learning disabilities, ultimately fostering more inclusive learning environments.

Keywords*: Dyslexia, Deep Learning, CNN, RNN-GRU, Spectrogram*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Description
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
GRU	Gated Recurrent Network
SPDACM	Speech Pace Detection and Audio Customization Module
TTS	Text To Speech

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In today's world, education and learning have become crucial for personal growth and societal advancement. As technology evolves, the demand for educated individuals has never been higher. This technological progress has also increased our awareness of concepts like neurodiversity, highlighting challenges faced by segments of society that have long been overlooked. Among these challenges, learning disorders associated with neurodiversity have gained significant attention.

According to the Diagnosing Learning Disorders: From Science to Practice, 3rd ed. by the American Psychological Association, six types of learning disorders are recognized: speech and language disorders, dyslexia, mathematics disorder, attention-deficit/hyperactivity disorder, autism spectrum disorder, and intellectual disability[1]. Dyslexia, the most common of these, affects approximately 20% of the population[2]. Adolescents aged 8–12 years with dyslexia often face misunderstandings due to their difficulty keeping up with standard learning paces, particularly if their condition goes undiagnosed. Despite being mistakenly perceived as less intelligent, studies have shown that dyslexic adolescents often exhibit enhanced creativity and intelligence in specific areas[3].

As awareness of learning disorders grows, substantial research has been conducted on dyslexia and speech-related challenges. Existing studies primarily focus on understanding the disorder's causes, exploring the relationship between dyslexia and speech, and documenting real-life cases to inform educators and practitioners[4][5]. Additionally, research has explored technological applications to support dyslexic learners, ranging from tools for identifying dyslexia to wearable IoT devices that aid multisensory learning[6]. However, the application of emerging technologies such as artificial intelligence (AI) in developing language-aiding tools for dyslexic individuals remains underexplored.

This study focuses on a critical area of struggle for dyslexic adolescents: speech comprehension. While conventional methods for measuring speech pace in healthy individuals are available, few methodologies consider the temporal and spatial features of speech in this context. The advent of AI has shifted the focus towards predicting dyslexic tendencies by analyzing audio data. However, this study aims to utilize deep learning techniques to develop a personalized speech comprehension tool that detects speech pace through audio analysis and subsequently delivers a customized speech pace via an integrated text-to-speech (TTS) model.

To achieve this, researchers must master Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for extracting audio features and converting audio to spectrogram images. Additionally, integrating a TTS model that leverages deep learning outputs is essential. This study will discuss existing knowledge, identify gaps, and highlight key areas for improvement. As the study progresses, the main and sub-objectives will be outlined, along with the current state of knowledge in the research domain. The study will also address the system requirements needed to realize the proposed solution, culminating in a detailed budget and budget justification.

1.1 Background

The field of education is rapidly transforming, driven by technological advances and an increasing understanding of diverse learning needs. A key focus in recent years has been neurodiversity, which recognizes that neurological differences, such as dyslexia, ADHD, and autism spectrum disorders, are natural variations of the human brain. Among these, dyslexia stand out as one of the most common learning disorders, affecting approximately 20% of the global population.

Many technological solutions have emerged to address the challenges associated with these disorders, each offering unique strengths and weaknesses. While these systems incorporate the latest innovations, they often lack personalization. For dyslexic adolescents, learning can be an exhausting task, and this lack of personalization can further discourage them from using existing systems. The proposed system, Lexayudha, addresses this by providing a personalized learning experience designed to reduce exhaustion in the learning environment. This is achieved with chromatic variations, personalized speech pace, the touch-math approach to enhance mathematical skills, and real-time adaptation based on emotion detection. The personalization of speech pace delivery is further explored in the following literature survey.

1.2 Literature Survey

Speech Rate Meter

Speech Rate Meter (SRM+) is an advanced tool designed to measure speech tempo as a complex prosodic component of intonation. The latest version (2021-10-11) introduces a significant new feature for detecting "FILLER" sounds, allowing for an assessment of the "pollution" of analyzed speech.

Key features

- Speech rate detection (words per minute method)
- Articulation Rate detection
- Pauses Score by calculating mean duration of inter-phrase pauses.

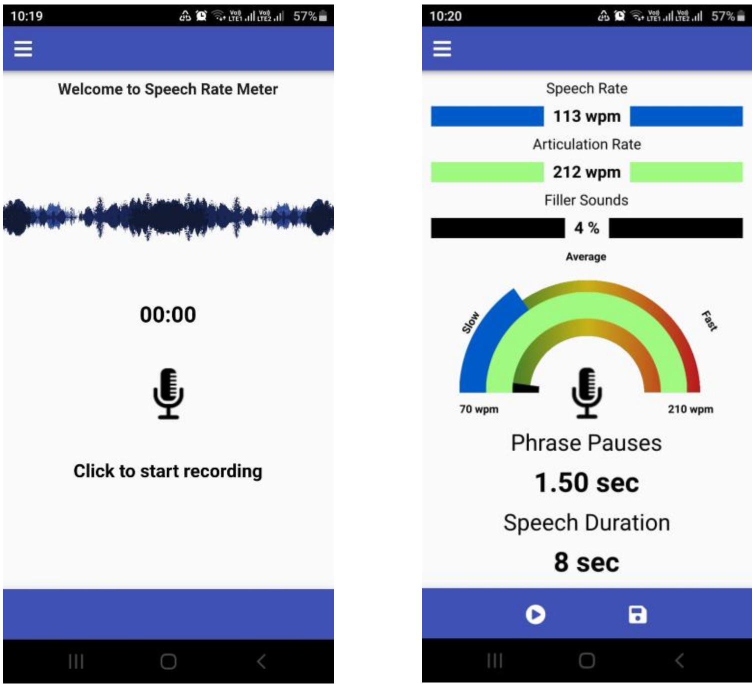

Figure 1.1: Speech Rate Meter app audio detection UI

Speechify

Speechify is a popular text-to-speech (TTS) application designed to help individuals with dyslexia and other reading difficulties. The app transforms written text into spoken words, making reading more accessible for users who struggle with traditional reading methods. This app also allows users to customize speech rate.

Key features

- Text to speech conversion.
- Customizable speech features such as speech rate.
- Multi-platform availability.

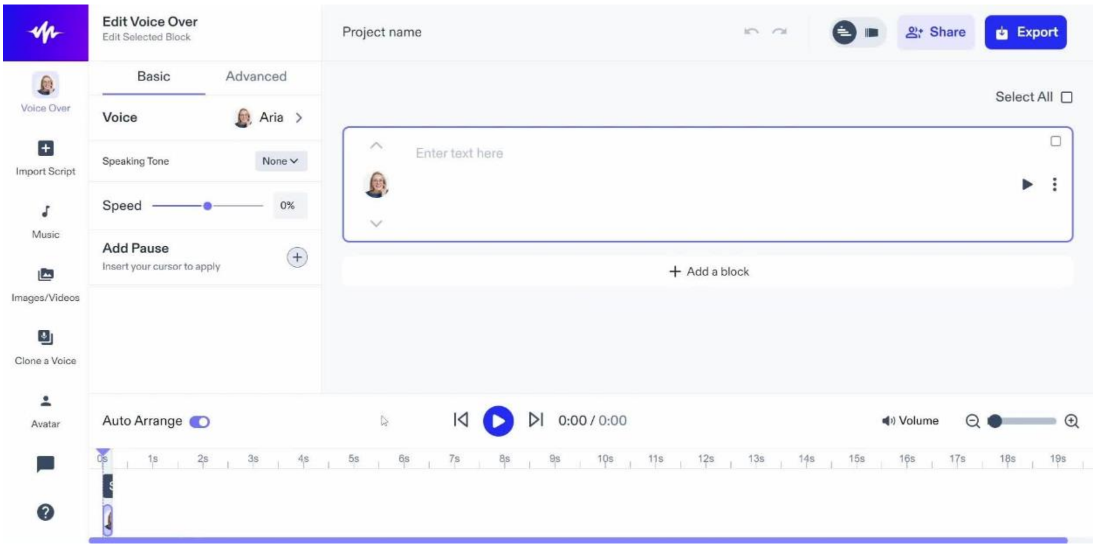

Figure 1.2: Speecchify dashboard

TextAid by ReadSpeaker

TextAid by ReadSpeaker is a comprehensive text-to-speech (TTS) application designed to support individuals with dyslexia and other reading challenges. This tool is particularly beneficial in educational settings, helping users to better engage with written content through auditory support.

Key features

- Adjustable reading speed
- Integration with learning management systems
- Text to speech capabilities
- Word and sentence highlighting

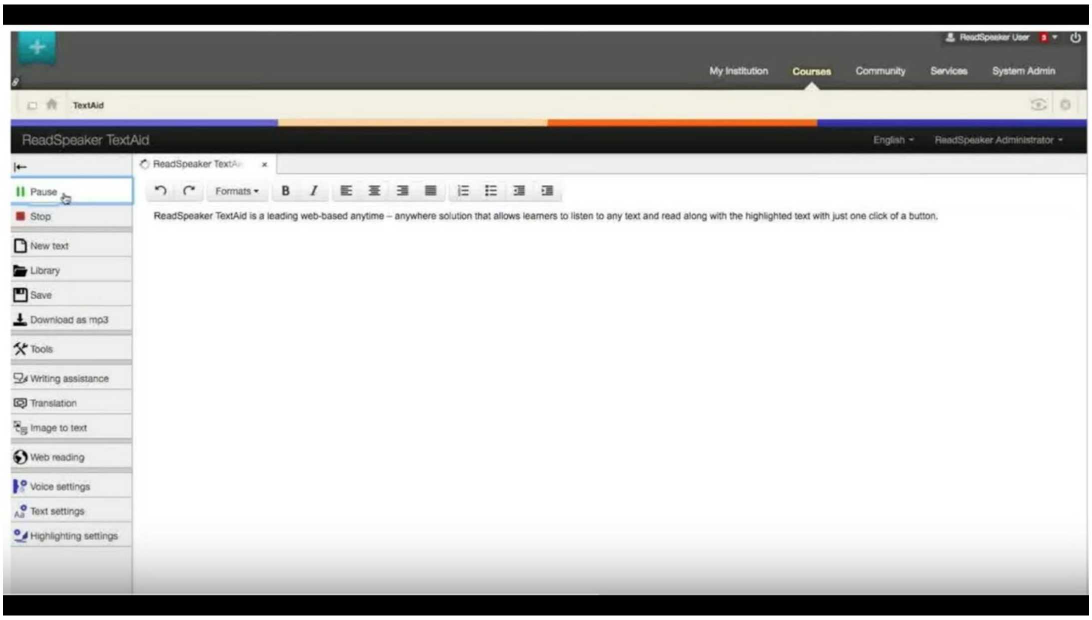

Figure 1.3: TextAid dashboard

Although existing software solutions address some issues effectively, they lack critical features. Lexayudha's personalized speech pace detection fills these gaps, providing tailored support.

Table 1.1 System comparison

Feature	**Speech Rate meter (SRM+)**	**Speechify**	**TextAid**	**Lexayudha**
Detects speech pace	✓	✗	✗	✓
Adjustable speech pace	✗	✓	✓	✓
Automated speech pace adjustment	✗	✗	✗	✓
Use of deep learning methods for high accurate prediction	✗	✗	✗	✓

Provides personalized speech pace	✗	✗	✗	✓

1.3 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. 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.

Although the detection of speech features has gained attention in recent years, resulting in promising developments in areas like hate speech detection [7] and emotion detection in speech [8], the application of modern technology to accurately determine speech pace for dyslexic adolescents remains lacking. This study aims to address this gap by exploring the use of cutting-edge technology, including deep learning methods, to develop personalized speech pace solutions for dyslexic adolescents. Specifically, this research will focus on detecting speech patterns in spectrogram images [9], utilizing hybrid CNN-RNN models [10] for data analysis, and integrating these findings with a Text-to-Speech (TTS) model.

1.4 Research Problem

Dyslexia is one of the most common learning disorders in modern society, affecting nearly 20% of the world's population. It poses a significant challenge to personal development, often hindering academic and social growth. Fortunately, studies have shown that if dyslexia are diagnosed early, its adverse effects can be mitigated, potentially enhancing the learning process later in life [11].

Among the challenges faced by dyslexic adolescents, speech comprehension stands out as a critical obstacle in typical learning environments. Effective speech comprehension is essential for better learning and overall comprehension. However, studies indicate that individuals with dyslexia often struggle with speech comprehension, which in turn impairs their ability to learn efficiently [12].

This study addresses the research problem by proposing a personalized speech pace delivery system. This system aims to serve as both a training and aiding tool, designed to improve speech comprehension skills in dyslexic adolescents. By tailoring the speech pace to individual needs, this solution seeks to enhance the learning experience and provide meaningful support for those affected by dyslexia.

2. OBJECTIVES

2.1 Main Objective

2.1.1 Develop and implement an AI-powered, personalized learning aid tool to enhance the learning experience and address the unique needs of dyslexic adolescents.

The main objective of this study is to develop an AI-powered learning aid tool that offers personalized learning experiences tailored for dyslexic adolescents. By incorporating personalization techniques, the tool aims to reduce fatigue commonly experienced in conventional learning environments and improve learning outcomes by keeping adolescents more engaged with the system.

The objective centers on developing a learning tool that customizes educational experiences to meet the specific needs of dyslexic adolescents through the use of advanced AI technologies. The project's progress will be tracked by achieving key milestones, such as the development of AI models, integration of various system components, and comprehensive user testing. These milestones are planned to be completed within a one-year timeframe. By addressing the significant challenges that dyslexic students face in traditional education settings, this project aims to offer a relevant and impactful solution. The one-year deadline ensures the tool is delivered in a timely manner, aligning with both the project's goals and broader societal needs.

2.2 Sub Objectives

In order to support and achieve the main objective, the following sub-objectives must be realized to fill the discussed research gaps and to provide a solution that addresses the research problem.

3.1.1 Convert audio files to spectrogram images.

This initial step 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. This conversion is crucial for preparing the data for advanced analysis and ensures that the subsequent processing accurately reflects the unique speech patterns of dyslexic individuals.

3.1.2 Analyze voice features for specific patterns to predict the speech pace

Once the audio is converted into spectrogram images, these images are analyzed to identify specific patterns related to speech pace. By leveraging a hybrid CNN-RNN model, the system can detect distinctive features and temporal patterns associated with dyslexic speech. This analysis addresses the research gap by providing insights into how speech pace differs between dyslexic and non-dyslexic individuals, leading to more accurate and personalized predictions.

3.1.3 Integrating a Text to Speech (TTS) model to incorporate predicted speech pace

After predicting the speech pace, the next step 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. This sub-objective directly addresses the need for a tailored learning experience, making it easier for users to engage with the content.

3.1.4 System evaluation and validation.

The final sub-objective focuses on evaluating and validating the system to ensure its effectiveness and reliability. This involves rigorous testing to confirm that the personalized speech pace is accurate and beneficial for the target audience. By validating the system, the project ensures that the proposed solution meets its objectives and provides a measurable improvement in learning outcomes for dyslexic adolescents.

3. METHODOLOGY

The proposed system, Lexayudha is aimed at delivering a personalized learning experience to dyslexic adolescents. To fulfill the desired objectives, Lexayudha utilizes cutting edge technologies such as deep learning techniques in combination with advanced methodologies to deliver the best performance while having the least performance bottleneck. This ensures a seamless user experience for the users of this proposed system.

Figure 4 demonstrates the overall system diagram and how different components of the system interact with each other to deliver the above-mentioned advantages.

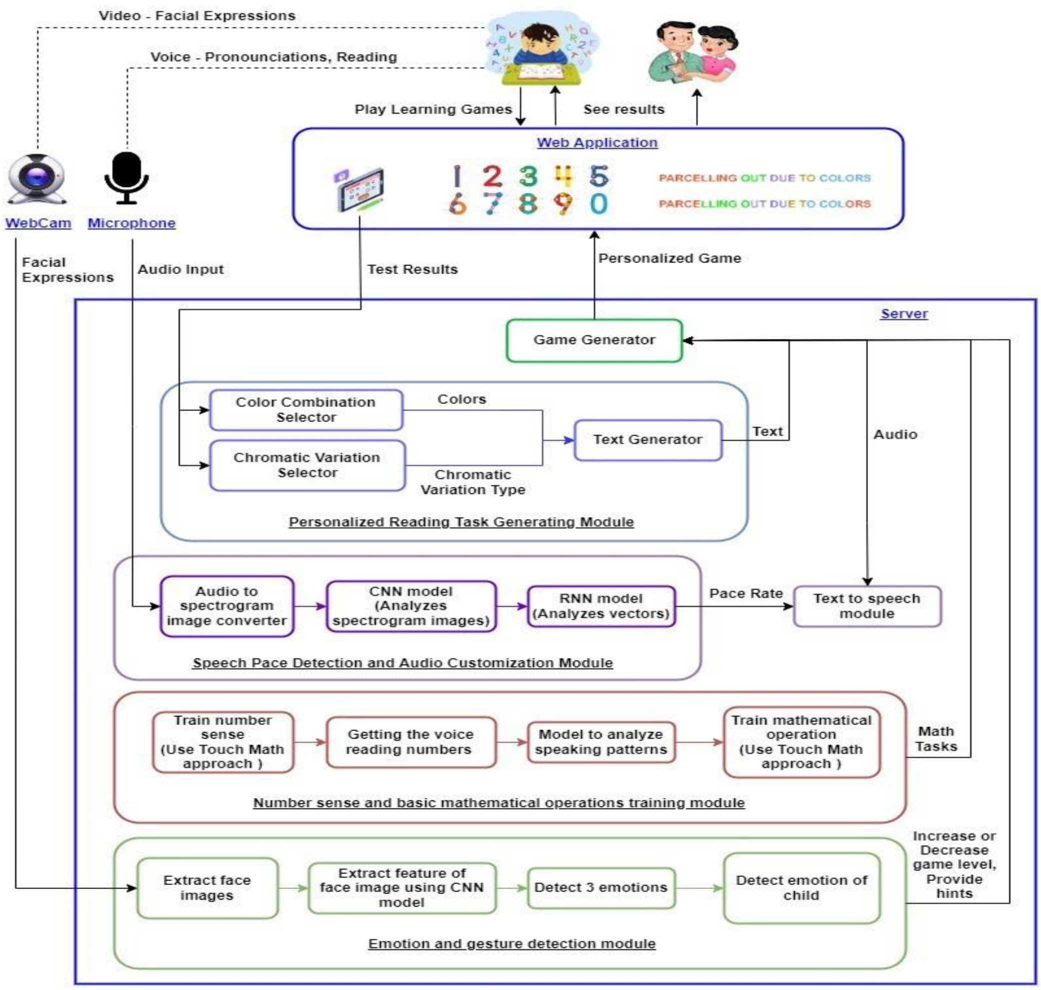

Figure 3.1 System overview diagram of Lexayudha

As shown in figure 4, Lexayudha consists of four major components, each component delivering a different learning experience to the users. While the server is represented as a single unit in figure 4, it is technically proposed to be developed using a microservice architecture. This approach enhances load balancing, security, and fault tolerance, preventing single points of failure and facilitating the integration of various technologies within the system. For data management, the proposed system will utilize MongoDB Atlas, a cost-effective and reliable cloud-based solution, which eliminates the need for maintaining a dedicated database infrastructure. The server of the system is intended to be hosted on Render, leveraging Docker for containerization and Kubernetes as the orchestration platform. The front-end web application of this purposed system is intended to be hosted on Vercel hosting platform. This setup ensures efficient management, scalability, and ease of deployment.

As shown in figure 5, the following overall system use case diagram captures the high-level system requirements.

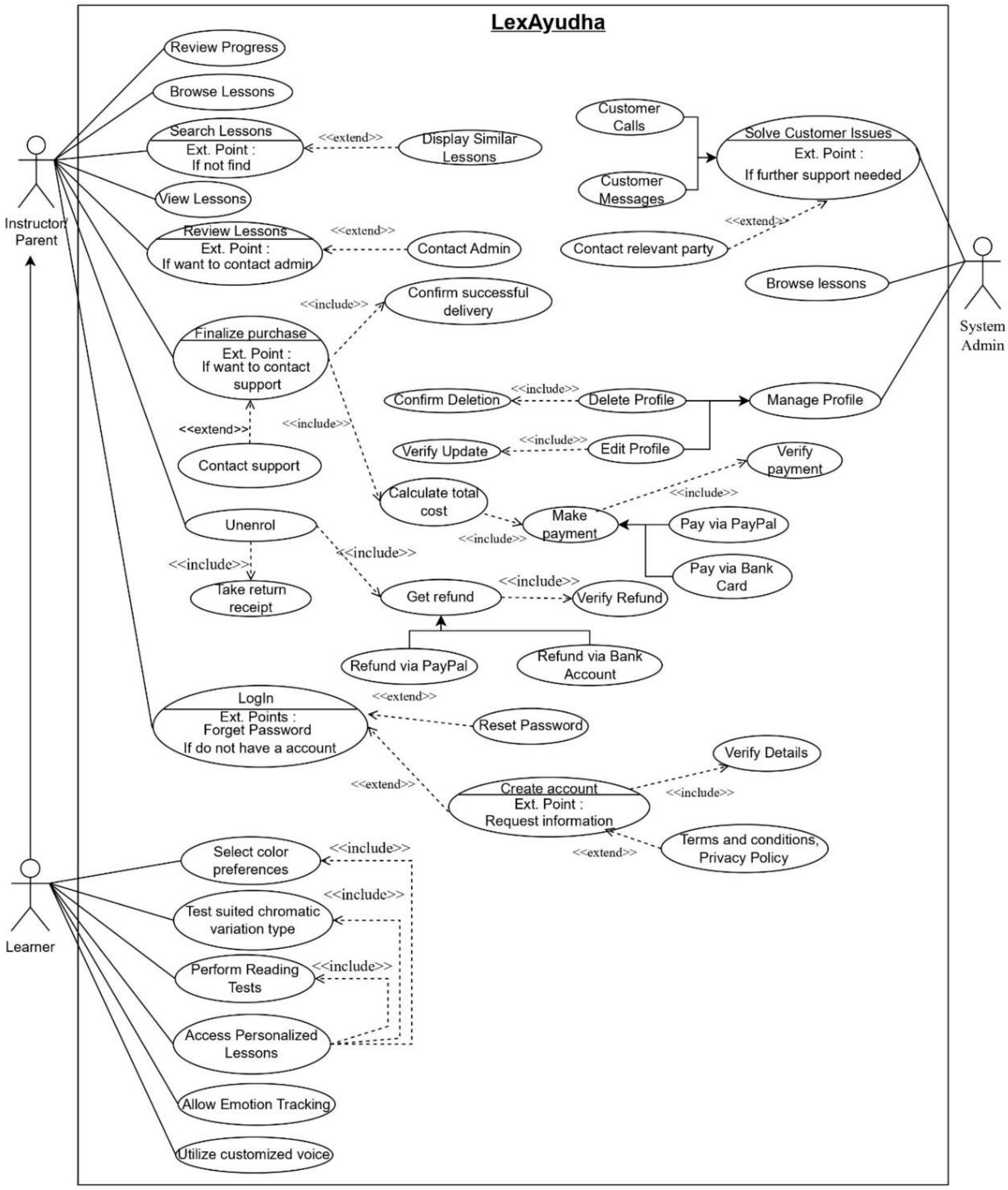

Figure 3.2 Use case diagram of Lexayudha

Apart from the common system overview, each component can also be further broken down into smaller sub-units with distinct technological layers. In this study, the speech pace detection and audio customization module (SPDACM) will be explored in greater detail.

Figure 6 illustrates the sub-component distribution of the SPDACM component.

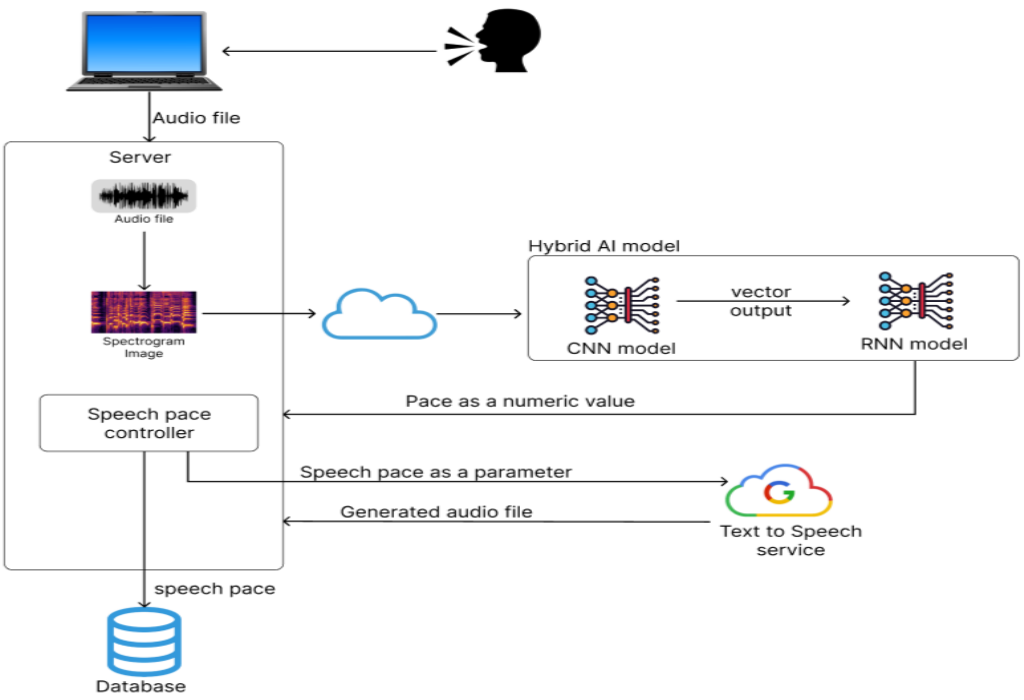

Figure 3.3 SPDACM component overview diagram

The primary objective of the SPDACM component is to deliver a personalized speech pace for dyslexic adolescents. To achieve this, the component combines deep learning with audio processing techniques. Initially, the browser captures audio through recording devices such as microphones. This audio capture process is initiated during a series of calibration tests presented in the system's initial user onboarding phase. The captured audio data is then sent to the server for preprocessing, where it is converted into spectrogram images using Librosa and Matplotlib Python libraries. These spectrogram images, which capture specific patterns, have been effectively used in previous studies to analyze voice features using CNN models [9]. Figure 7 shows a visualization of a spectrogram image of a converted audio sample.

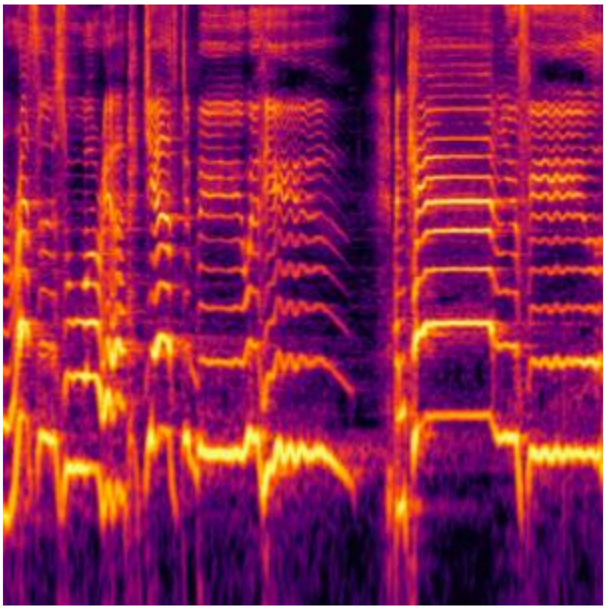

Figure 3.4 Spectrogram image visualization

After the spectrogram images are converted, they are sent to a hybrid CNN-RNN deep learning model to analyze and detect speech pace-related patterns. This hybrid approach, inspired by research on detecting brain abnormalities, leverages the strengths of both CNN and RNN models to enhance pattern detection [10]. In the SPDACM component, the hybrid model uses a CNN to detect and extract specific spatial and temporal features from dyslexic speech. These features are then mapped as a vector, which is processed by an RNN model to identify time-series patterns relevant to speech pace. Specifically, a Gated Recurrent Unit (GRU), a variant of RNN, is employed due to its lower computational demand compared to the Long Short-Term Memory (LSTM) variant. While both variants address issues such as the gradient vanishing problem commonly faced by conventional RNNs, the GRU variant is expected to deliver better performance relative to its computational cost.

Once the spectrogram images have been analyzed by the hybrid models, the predicted speech pace is sent back to the server. The server stores this value in the database under each individual's profile. This stored speech pace is then used as a parameter to adjust the speech rate of the integrated Text-to-Speech model, providing a personalized speech pace for dyslexic adolescents.

While delivering personalized speech pace is the primary objective, this component has four critical sub-objectives that must be achieved to accomplish the main goal.

These sub-objectives include audio-to-spectrogram conversion, hybrid model development and fine-tuning, integration with the TTS model, and finally, system evaluation and validation. Successfully realizing these sub-objectives is essential to meeting the overall objective of the proposed system.

To effectively implement the SPDACM component, the primary Node.js server will be supported by a dedicated Python microservice responsible for converting audio input into spectrogram images. This microservice handles the preprocessing step crucial for accurate analysis. The training of the hybrid deep learning model, which includes both CNN and RNN architectures, will be conducted using Google Colab. This platform provides the necessary computational resources and environment for developing and refining the model. Once the predicted speech pace is generated by the model, this information will be stored in Azure Cosmos DB, ensuring reliable and scalable data management. Finally, for the Text-to-Speech functionality, the Google Text-to-Speech model will be integrated into the system, providing a personalized speech pace experience to the of dyslexic adolescents.

To successfully execute this project, several key requirements must be met. First, a development device is needed to configure the development environment. Adequate computational power is also essential for training and developing the proposed models, which can be effectively achieved using cloud-based platforms such as Google Colab. For initial data collection, a high-quality audio recording device capable of capturing sensitive vocal features is required. Additionally, a data storage solution with sufficient capacity is necessary to manage the collected data. Finally, access to relevant documentation and study materials is crucial for acquiring the necessary knowledge to carry out the project effectively.

To train the hybrid models for detecting dyslexic speech pace, two distinct sets of voice data are required. The first set should consist of speech data from typical adolescents, while the second set should include speech data from dyslexic adolescents. These datasets will enable the models to differentiate between dyslexic and normal speech features. The dataset for normal speech is already available on Kaggle [13]. The dyslexic speech dataset will need to be collected through interviews at centers specializing in dyslexia, with proper authorization from relevant authorities.

In order to accomplish the given sub-objectives, it is crucial to follow a well-structured timeline with a proper work breakdown, especially since the project is time-bound and must be completed within a one-year period. Figure 8 shows the work breakdown chart of the SPDACM component.

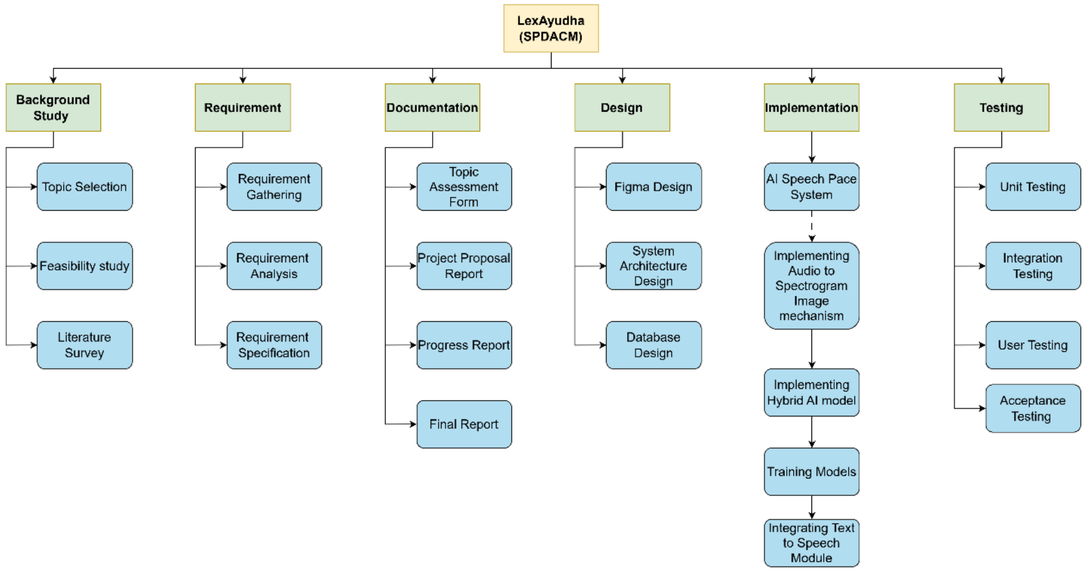

Figure 3.5 Work breakdown structure of SPDACM component

In terms of timeline allocated to realize the sub-objectives, hybrid model training takes up the largest portion of the timeline due to the complexity of development and the time need for the fine-tuning process to deliver accurate predictions. Figure 9 presents the Gantt chart illustrating the project timeline.

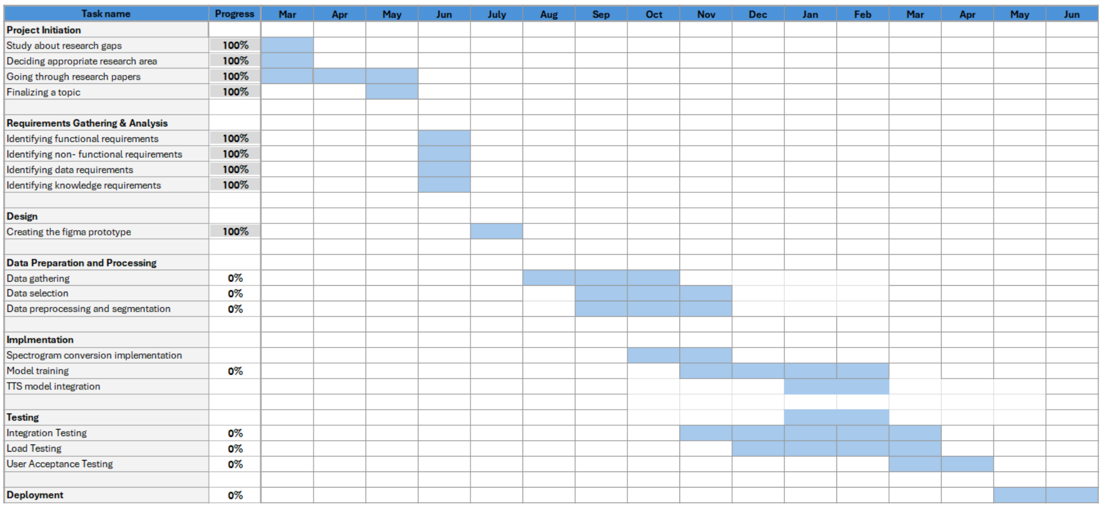

Figure 3.6 Gannt chart of SPDACM component

In conclusion, the SPDACM component aims to provide personalized speech pace for dyslexic adolescents, easing their learning challenges and reducing fatigue. This methodology can also be applied in real-world situations where personalized speech pace improves user experience. For example, a voice-guided tour app could use this approach to help users with different language skills better understand instructions. There are many other scenarios where this method can enhance user experience.

Commercialization Aspects of the Product

It is in fact a business opportunity for the proposed AI-driven adaptive web-based platform, as it uniquely fills the gap in the education sector empowered by personalization in adaptive learning experiences for students with dyslexia. The potential number of people with dyslexia is fairly large, and interest in effective and scalable educational solutions has been quite pervasive and growing at a very fast rate. By providing a solution using artificial intelligence, chromatic variations, and natural language processing, the platform will be positioned as part of the best in class for this EdTech marketplace.

Target Market

The two target markets for this solution will primarily be schools and other educational facilities, as well as parents of dyslexic students. Key Target Audiences for these target markets include:

- Primary and Secondary Schools: Public and private schools that afford special education programs have a great opportunity. These will most frequently be looking out for a package that can help teach such children with learning disabilities, and what the platform will offer in terms of personalized and adaptive learning will effectively meet this demand.
- Parents: Parents of dyslexic children make another crucial segment of the market. In general, parents are more than glad to pay for materials or services which would enable their children to climb up the ladder educationally, especially if the regular learning methodologies have failed.

Educational technology companies: Suitable educational technology companies can also become collaborative to scale the reach of this platform. Many of these companies work to expand their product lines with creative solutions; in this light, the inclusion of this platform in their existing suites will be beneficial for both parties.
Dyslexia Advocacy Groups: It can form alliances with independent, non-profit making dyslexia advocacy groups dedicated to making lives with dyslexia better and echo this platform for more visibility and credibility in general.

Pricing models

This would be a flexible pricing strategy that caters to the needs and budgets of the target market in order to be successful financially. For the purpose of pricing, it may use any one of the following models:

- Subscription based model: The plan will apply a tiered subscription model at a fee charged, plus the features required. With the next tier of features being the advanced tiers, which include advance analytics and an extended library of content with further customization options, the example of one basic tier including the core features, such as individual learning paths and chromatic adjustments. Therefore, depending on the level of service needed, the price will range from $5 and above monthly.
- Freemium Model: That is why, for instance, on this platform, a freemium model could be used where all the basic functionalities would be free of charge. Therefore, users would have experience of the platform and subsequently transfer to the paid version. By subscription, premium feature advanced AI-driven content customization, detailed progress reports, and more support options could be unlock.

Market Strategy

A marketing strategy will be put in place to commercialize the platform successfully .

- Marketing and Outreach: Participating in educational technology related meetings and visiting places offering education for dyslexics and dyscalculics can be of importance, which will enable the creation of brand awareness. This helps to network potential clients and other stakeholders. Impressions on the potentials of the platform for benefiting students with dyslexia and dyscalculia can be achieved through the demos or presentations made at these events.

- Online marketing will be conducted through digital campaigns that place ads on social media, educational forums, and search engines for both school and parent buyers. Thought leadership on this subject will be built using content marketing for those in education for whom dyslexia is a concern. Collaboration with schools and special education programs will facilitate early testing for the platform.
- Dyslexia advocacy groups can boost the believability as well as exposure in the community. Support for the platform as well as ongoing development can be provided through these groups based on input.

Sales Channels

- Direct Sales: Working on a face-to-face basis with schools, colleges, and school districts to make aware the platform and close contracts.
- Online Sales Portal: The channel will also be purchasable and subscribed to through online sales portals, thus being easily accessible to individual customers and smaller institutions.

Only high customer support will help in the retention of the user base and, thus, ensure word-of-mouth publicity. The support services will consist of onboarding assistance, training sessions for educators, and continuous technical support.

Regular updates and improvements, based on user feedback, will maintain the relevance and effectiveness of the platform and thereby secure customer satisfaction in the long term.

Global Expansion

In future, after the platform is developed for the English-speaking markets, the company will get further development of the platform, which will provide an opportunity to cover many languages and educational systems across the globe. The content and the user interface of the platform will be localized based on the specific needs of the international markets.

Revenue Projections and Long-Term Vision

This is so because, if well priced and marketed, large potential can be realized in terms of revenue generation on this platform. Most of the revenues in the first few years will come from subscription sale. Further with the growth, when the user base has increased, more ways can be figured out to better monetize the platform, such as data analytics services provided to institutions or premium content and features.

4. PROJECT REQUIREMENTS

4.1 Functional Requirements

- **Voice input processing**
    - The system must allow users to input spoken words or phrases through a microphone.
    - The input voice data must be captured and processed in real-time.
- **Audio to spectrogram image conversion**
    - The system must convert the audio input into a spectrogram image on the server.
    - The conversion process should occur quickly to minimize delay.
- **Speech pace prediction**
    - The system must analyze the spectrogram image using a hybrid CNN-RNN model to detect the speech pace.
    - The system must accurately predict the speech pace based on temporal and spatial features in the audio.
- **Store predicated speech pace and audio data**
    - The system must store the predicted speech pace and corresponding audio data in a MongoDB database.
    - The system must allow for retrieval of stored data for future use.
- **Integrated Text to Speech model**
    - The system must integrate with Google Text-to-Speech (TTS) to provide personalized speech playback.
    - The TTS model must use the predicted speech pace to adjust the output speech.
- **Progressive personalization of speech pace**
    - The system must allow for the customization of speech pace over time.
    - The system should adapt to the user's progress, assessing speech pace periodically and adjusting as necessary over time.
- **Error handling mechanisms**

- The system must handle errors gracefully, providing clear feedback to the user in case of issues with input processing, data storage, or playback.

4.2 User Requirements

- **Voice input capability**
  - Users must be able to input speech via a microphone or audio recording device, with support for various audio formats.
- **Real-Time feedback**
  - Users should receive real-time feedback on their speech pace during or immediately after input, with clear and actionable information.
- **Secure access**
  - Users should have secure access to their personal data and settings, with privacy and data protection measures in place
- **Help and support**
  - Users should have access to help resources, including tutorials and support documentation, with customer support available for technical issues.
- **Error feedback**
  - Users should receive clear and understandable error messages if issues arise, with guidance on how to correct them and resume normal operation.

4.3 System Requirements

- **Processing power**
  - The system must be equipped with sufficient computational resources to handle real-time audio processing and spectrogram conversion efficiently.
- **Data storage capacity**
  - The system must have sufficient storage capacity to manage large volumes of audio files, spectrogram images, and speech pace data over time.
- **Network bandwidth**

  - The system must ensure sufficient network bandwidth to handle the transmission of audio files, spectrogram images, and predictions between the server and users without delays.
- **System integration**
  - The system must be able to integrate seamlessly with existing software and hardware components, including TTS models, databases, and user interfaces.
- **Scalability and load balancing**
  - The system must support scalability and load balancing to handle increased user traffic and data processing demands as the demand grows.
- **User authentication and authorization**
  - The system must include secure user authentication and authorization features to manage user access and protect sensitive data.
- **Data backup and recovery**
  - The system must include mechanisms for regular data backup and recovery to prevent data loss and ensure system reliability

4.4 Non-functional Requirements

- **Accuracy of predicted speech pace**
  - The system must achieve high accuracy in detecting speech pace, with an error margin of less than 5% compared to expert human assessment. This level of precision is essential to ensure that the personalized speech pace recommendations are reliable and effectively support the learning needs of dyslexic adolescents
- **Minimum time on pace detection process**
  - The system must process and return the predicted speech pace within 2 seconds after receiving the voice input to ensure a smooth user experience.
- **Scalability of the system**
  - The system must be scalable to handle multiple users simultaneously without a significant drop in performance.

- **Reliability of the system**
    - The system must be reliable, with an uptime of 99.9%, ensuring it is available whenever needed.
- **Security measures**
    - The system must ensure data encryption for all user data, both in transit and at rest, to protect sensitive information. Also security measures must be in place to prevent any cyber attack which could affect the users and availability of the system.
- **Maintainability of the system**
    - The system must ensure data encryption for all user data, both in transit and at rest, to protect sensitive information.
- **Compatibility**
    - The system must be compatible with various devices.

5. BUDGET AND BUDGET JUSTIFICATION

The proposed system was planned on a budget in a way that the cost can be reduced and hence could provide a feature rich application to the end users for an affordable price for them. Below is the planned budget for the proposed system and the budget justification. Please note that the budget allocation and the selection of the vendors can be slightly changed with future findings and requirements.

Proposed Budget and allocation

Table 2: Proposed Budget

Description	Cost
Azure Blob Storage	$ 0.018 per GB (first 5GB free)
Azure Cosmos DB	Free tier (50 GB free for year)
AWS SageMaker	$ 0.10/hour (16 GB RAM)
Azure Virtual Machines	Free tier (750h)
Azure Speech to Text	$ 1 /hour (5 hours free per month)
AWS SES	Free (3000 emails per month)
Vercel – Frontend Deployment	Free tier
Render – Backend Deployment	Free tier

Budget justification

Azure Blob storage: Azure blob storage was chosen for it competitive pricing and the free tier. It provides $ 0.018 per GB after the free first 5GB. Blob storage has one of the lowest storage prices compared to AWS S3 bucket $ 0.023 per GB and google clouds $ 0.026 per GB. Blob Storage is well-suited due to its ease of integration, flexibility, and ability to handle unstructured data and also Azure provides a more affordable entry point while still maintaining high performance and scalability.

Azure Cosmos DB: Azure cosmos DB was selected as No SQL Database for its cost effective and extensive free tier offering 50GB of free storage per month for the first year. This makes it's the most suitable option compared to the AWS dynamodb (25 GB per month for a year) and google firestore (1 GB per month for a year). Cosmos DB also provides globally distributed, low-latency access to data, which is ideal for a real-time application like Lexayudha, where performance is critical for storing and retrieving results and feedback. Additionally, Cosmos DB supports multiple NoSQL APIs, making it flexible.

AWS SageMaker: AWS SageMaker can select as the best choice to address the machine learning parts due to its robust platform and flexible pricing options. At $0. 10/hour for an ml. t3. medium instance, SageMaker also presents a a highly cost-effective solution to running machine learning computations. It also easily interconnects with other AWS services, so this is advantageous if you decide to upscale system's AI models.

Azure Speech to Text: Azure Speech to Text was chosen for its low cost ($1/hour) and a free monthly allowance of 5 hours, which is beneficial for projects with moderate usage. Also compared to AWS and Google speech to text prices which is $1.44/hour and free 60 minutes for month, Azure speech to Text is more suitable for the system. The service's high accuracy in transcribing audio to text, along with its integration capabilities within the Azure ecosystem, makes it ideal for project's potential need for speech processing

6. 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] “Learning Disabilities Statistics and Prevalence | HealthyPlace,” Jan. 17, 2022. https://www.healthyplace.com/parenting/learning-disabilities/learning-disabilities-statistics-and-prevalence.

[3] “People with dyslexia have ‘enhanced abilities’, according to a new study,” World Economic Forum, Mar. 07, 2023. https://www.weforum.org/agenda/2022/07/dyslexia-enhanced-abilities-studies/

[4] J. Roitsch, “Dyslexia and the speech pathologist,” in IntechOpen eBooks, 2021. doi: 10.5772/intechopen.93690.

[5] K. L. Cabbage, K. Farquharson, J. Iuzzini-Seigel, J. Zuk, and T. P. Hogan, “Exploring the overlap between dyslexia and speech sound production deficits,” Language Speech and Hearing Services in Schools, vol. 49, no. 4, pp. 774–786, Oct. 2018, doi: 10.1044/2018_lshss-dyslc-18-0008.

[6] M. Mohana, A. C. Da Silveira, V. Suvetha, P. Subashini, G. Ghinea, and C. A. S. Santos, “Technology Enhanced Mulsemedia Learning (TEML) for Learners with Dyslexia,” ACM International Conference on Interactive Media Experiences Workshop, Jun. 2023, doi: 10.1145/3604321.3604357.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] G. K. Georgiou, D. Martinez, A. P. A. Vieira, A. Antoniuk, S. Romero, and K. Guo, “A meta-analytic review of comprehension deficits in students with dyslexia,” Annals of Dyslexia, vol. 72, no. 2, pp. 204–248, Sep. 2021, doi: 10.1007/s11881-021-00244-y.

[13] “Non-Native Children English Speech (NNCES) corpus,” Kaggle, Oct. 29, 2022. https://www.kaggle.com/datasets/kodaliradha20phd7093/nonnative-children-english-speech-nnces-corpus