Multisensory Learning Tools for Improving Skills in Children with Down Syndrome

24-25J-096

Project Proposal Report

Kamesh Diviyanajan

IT21155352

B.Sc. (Hons) Degree in Information Technology Specializing in Software Engineering

Department of Information Technology Sri Lanka Institute of Information Technology Sri Lanka

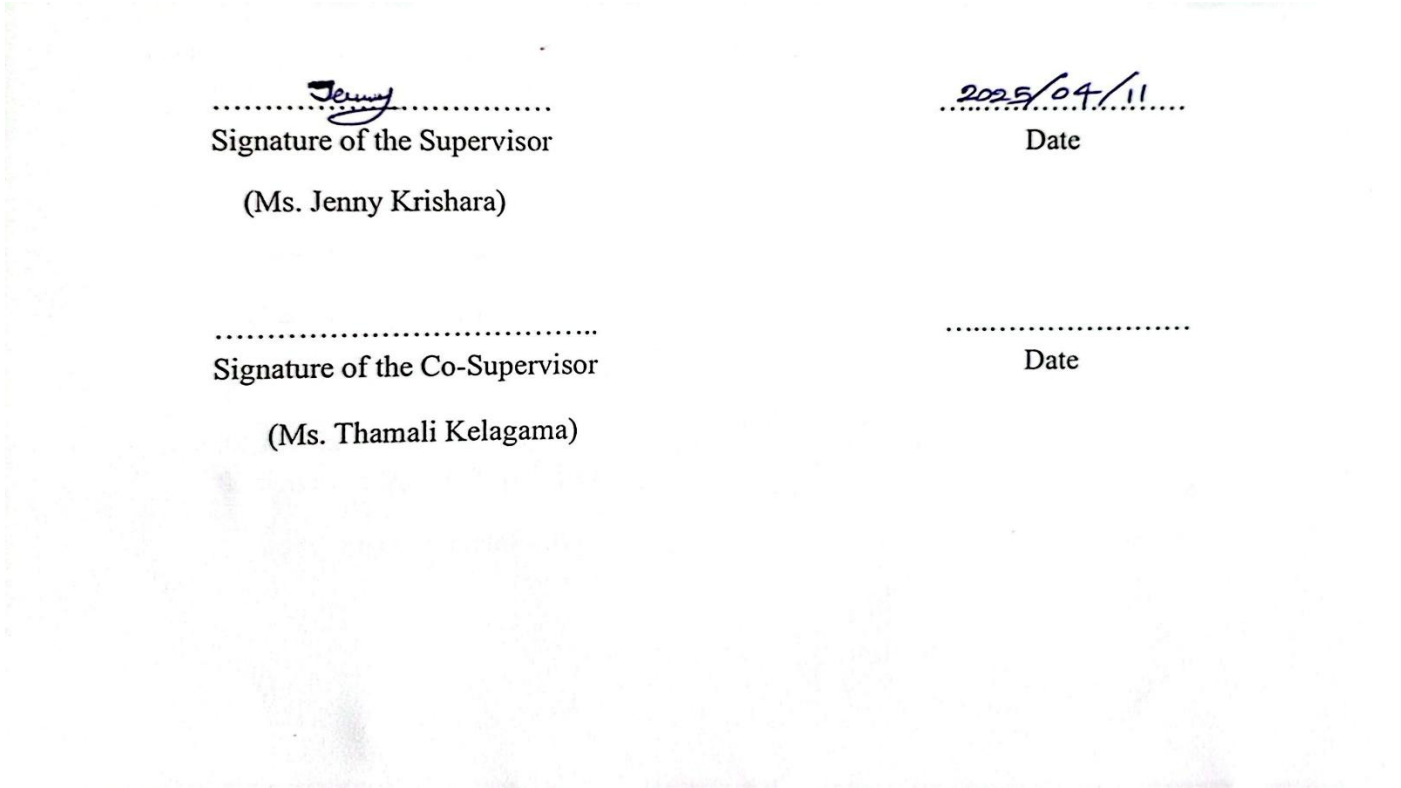
June 2024 /2025

DECLARATION

I declare that this is my 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
H . K . Diviyanjana	IT21155352	

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



.....Jenny.....
Signature of the Supervisor

(Ms. Jenny Krishara)

2025/04/11
Date

.....
Signature of the Co-Supervisor

(Ms. Thamali Kelagama)

.....
Date

ABSTRACTION

The purpose of this study was to explore the impact of targeted speech training on vocabulary development in children aged 5 to 12 with Down syndrome. The research aimed to address the specific challenges these children face in acquiring and using a diverse vocabulary, which is crucial for effective communication and cognitive development. The study was designed as an intervention-based program, incorporating various speech therapy techniques and vocabulary-building exercises over a period of six months.

Initial findings suggest that consistent speech training, with a focus on introducing and practicing different words, significantly enhances the children's ability to understand and use new vocabulary in everyday communication. Notable trends included improved articulation, increased confidence in speaking, and a broader range of vocabulary used in spontaneous speech.

the study underscores the importance of develop by speech training programs that cater to the unique needs of children with Down syndrome. The findings suggest that such programs can lead to measurable improvements in language skills, supporting these children in their academic and social environments.

Contents

BACKGROUND	1
RESEARCH PROBLEM	2
OBJECTIVES	3
METHODOLOGY (OVERALL TEACHING VOCABULARY TRAINING DIAGRAM)	4
USER CASE DIAGRAM	4.1
USER CASES	4.1.2
WIREFRAME (LOW LEVEL)	4.2
DATASET	4.3
WORK BREAKDOWN CHART	4.4
PROJECT REQUIREMENTS	5
GAN-CHART	6
BUDGET AND BUGET JUSTIFICATION	7
TESTING AND IMPLEMENTATION	8
RESULTS & DISCUSSION	9
CONCLUSION	10
REFERENCES	11

BACKGROUND

what is down syndrome ? No it is An additional copy of chromosome 21 is the genetic cause of down syndrome. We can identify 3 type down syndrome type Trisomy 21 , Translocation and Mosaicism of Approximately 1 in every 700 live newborns are case by this syndrome, which is linked to a number of developmental, cognitive, and physical difficulties. The speech and language development impairments that children with Down syndrome frequently face can have a substantial negative influence on their capacity for successful communication. Children with Down syndrome children need early intervention to improve their communication abilities

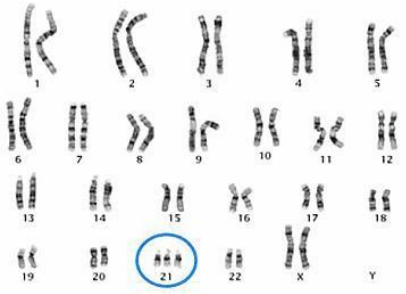


figure 1

challenge of this Down syndrome children Although speech treatment approaches have advanced, a multiple number of case still exist. Speech training can be made more difficult for children with Down syndrome since they frequently co-occur with disorders like hearing loss or attention deficit hyperactivity disorder[1]. Moreover, the cultural and linguistic heterogeneity among children diagnosed with Down syndrome necessitates customized strategies to address specific requirements

Speech and Language Development Down Syndrome Children

Children identified as having Down syndrome frequently show impairments in language and speech development. A number of causes, such as cognitive decline, hearing sound, and anatomical variations in the oral and vocal systems, are frequently blamed for these delays.
according reference research show Down syndrome children have good ability to understanding receptive language[2]

Speech Therapy Techniques

Children with Down syndrome have been studied using a variety of speech treatment approaches. Activities that enhance phonological awareness, oral motor abilities, and articulation are frequently included in traditional speech therapy. Furthermore, methods like the Picture learn Communication System , a visual support system, and auditory-verbal therapy, which stresses listening skills, have been found to be successful.[3]

How to help Down syndrome user technology

Current variety of software App and websites available down syndrome student develop example Leatstake Manus line like different App and websites available these app using different Deep laring module using help Down syndrome improve speech

2.RESEARCH PROBLEM

About 1 in 700 born worldwide are affected by Down syndrome, a genetic disorder caused by an extra copy of chromosome 21. Speech and language deficits are particularly common among the developmental obstacles that are frequently associated with this illness. Communication is extremely difficult for children with Down syndrome since they usually show delayed speech and language development. These level of education, social engament, and general quality of life are all significantly impacted by this problem. These kids' reason to speech and language issues include limited expressive language skills, a delayed initiation of speech, articulation and phonological challenges, and a few of vocabulary knowledge . main reason of these are lack of comprehension and clarity in their speech

Even while these speech and language difficulties are acknowledged, there is still a significant knowledge gap on the underlying causes of these problems and the efficacy of different treatments. A one-size-fits-all approach to intervention may not be beneficial given the variety in speech and language abilities among children with Down syndrome, despite several studies examining the anatomical and physiological components of speech production in this population. Furthermore, little is known about how these children's speech and language development is influenced by the interplay of cognitive, motor, and sensory elements.

Given the of effective communication in all areas of life, explore new solutions and interventions that can be tailored to meet the needs of children with Down syndrome. all referenced research shows that early intervention and consistent speech therapy can have a positive impact on speech outcomes; however, but it not a the long-term sloution of these approaches.

This proposed component address how to down syndrome children vocabulary knowledge improve and how help to engage with other people

2.1 RESEARCH GAP

Features	Proposed System	Research 1	Research 2	Research 3	Research 4
UI/UX implementation	✓	✓	✗	✓	✓
Word correction detection	✓	✗	✓	✓	✓
Scalability	✓	✗	✓	✗	✗
Customer Word adding	✓	✗	✗	✓	✗
Progress Detections	✓	✓	✓	✓	✓

Research 1 Virtual Personal Assistant By Using Hand Gesture And Voice Assistant For Disabilities compare with Proposed System this research user to a system that focused on voice detection using text-to-word recognition that user in CNN module. This research system not provide word adding option This already includes progress detection .interface

Research 2

Reflection of the Emotional State in the Characteristics of Voice and Speech of Children with Down Syndrome Compare with the Proposed System using also this system have progress tracking according prominence and CNN neural network use to voices frequency detection

UI/UX experience interface implement to use in react native this project mobile app compare with proposed system develop technology different not provide word adding options

Research 3

EduPlanner –Best Teaching Method for Students with Down Syndrome this research develop by game to down syndrome children .Audio learning develop one game this research main focus on down syndrome identify comapay with propose system focus on develop speck skill but it use in voice DNN mobile use to training voice detection and this research system available listing word and progress tracking individual

Research 4

E-Learning Education System For Children With Down Syndrome this research system consider 3 level voice of different 3 give identify each word pronation level proposal system consider given one word according pronunciation skill decide level need audio.

3. OBJECTIVES

This suggested web application component's main objective is to develop a dependable and user-friendly platform that aids in speech therapy and word pronunciation improvement for kids with Down syndrome. The software provides interactive activities tailored to specifically target the articulation difficulties young kids have. The platform offers engaging and effective speech development support through constant practice and personalized feedback. Both kids and careers may utilize the user-friendly design, which facilitates a seamless and pleasurable learning experience while improving communication skills.

2.2 Specific Objective

- Word pronation and correctness check :

this task provides Down syndrome children to list of word pronation user can select word and listing and looking that word relate image word and star speech user speech using LSTM Module recognize correctness of word of user speech

- Audio speed recognition:

this sub task main consider user speak recognition correct or not .example we are provide 200 word user can select one word practice word sound fist give user to listing word user can listing word many of time after listing word user system ask user “are you ready to speech “ now user can speech this time start speech resonation process this process identify sound frequency using check correct or wrong . correct word adding database

- Customer word adding

This task give to user add different they want. Fist user add word store database and next start generate sound according to given word .user can listing generate sound start taring add word this also fallowing voice recognition process.

- Correct word pronunciation examine
  A tool to check and evaluate word pronunciation accuracy will be included in the proposed application. This function will assess the number of words the youngster can pronounce correctly and monitor their development over time. The number of words the child knows and can pronounce correctly will be determined by the system through testing their vocabulary and pronunciation abilities using the results of these exams.

4. METHODOLOGY (OVERALL TEACHING VOCABULARY TRAINING DIAGRAM)

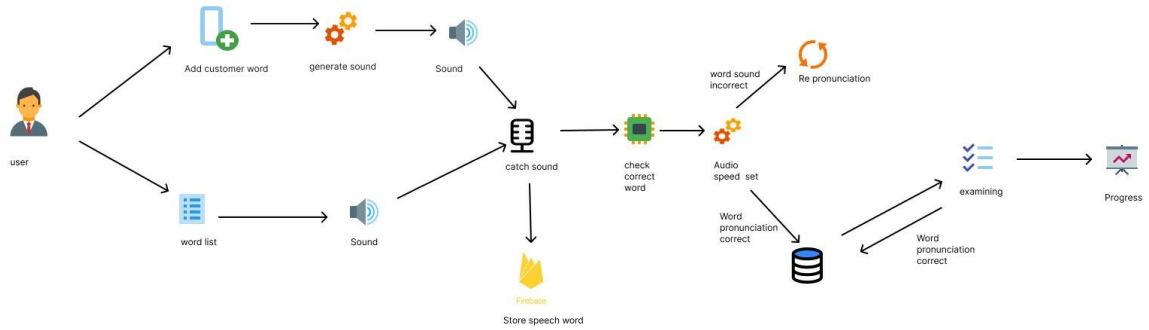


figure 2

This component expects to train different word sounds Down syndrome students these two option provide user one add customer word insert and list of listing to different word sound. This custom word adding the main purpose is help to medical therapy Some of down syndrome students doctors recommend different word training .this purpose archive help this part. The overall process of custom word adding fist user have option add word add word store database and start generating correct sound this user google sound generation after generate sound give children to listing. initial level sound speed is normal (there are three sound levels in this system one is normal, middle, and low) initial level gives list after the listing system display massage "late speech with me" The user can start speech that time start voice correction detection; fist identify sound using mic identify sound pass to correction detection module using this module identify the pronation correctness if it is not correct pronation correctness system give 3-time chances this three time module through suggestion Audio speed level. Pronation correctly in word store word in the database. Give option to examine word in daily and improvement progress

list of listing to different word sound this also happens same prosody in the above user one add customer word insert; this system word correct nation popup different emoji like below. reason to use this like to different colorful thinks

figure 2.1

figure 2.3

More explain other provide task this like below;

speech recognition

this sub task main consider user speech regonation correct or not .example we are provide 200 word user can select one word practice word sound fist give user to listing word user can listing word many of time after listing word user system ask user “are you ready to speech “ now user can speech this time start speech recognition process this process identify sound frequency using check correct or wrong . correct word adding database

Different word adding

This task give to user add different they want. Fist user add word store database and next start generate sound according to given word .user can listing generate sound start taring add word this also fallowing voice recognition process.

Examine word

This task given user to examine different word example this already using user correctly pronation word . this give five word daily to user to fist user say word that relevant image that add looking at this user can pronaos ; this time same process fallowing speech recognition.

4.1 USER CASE DIAGRAM

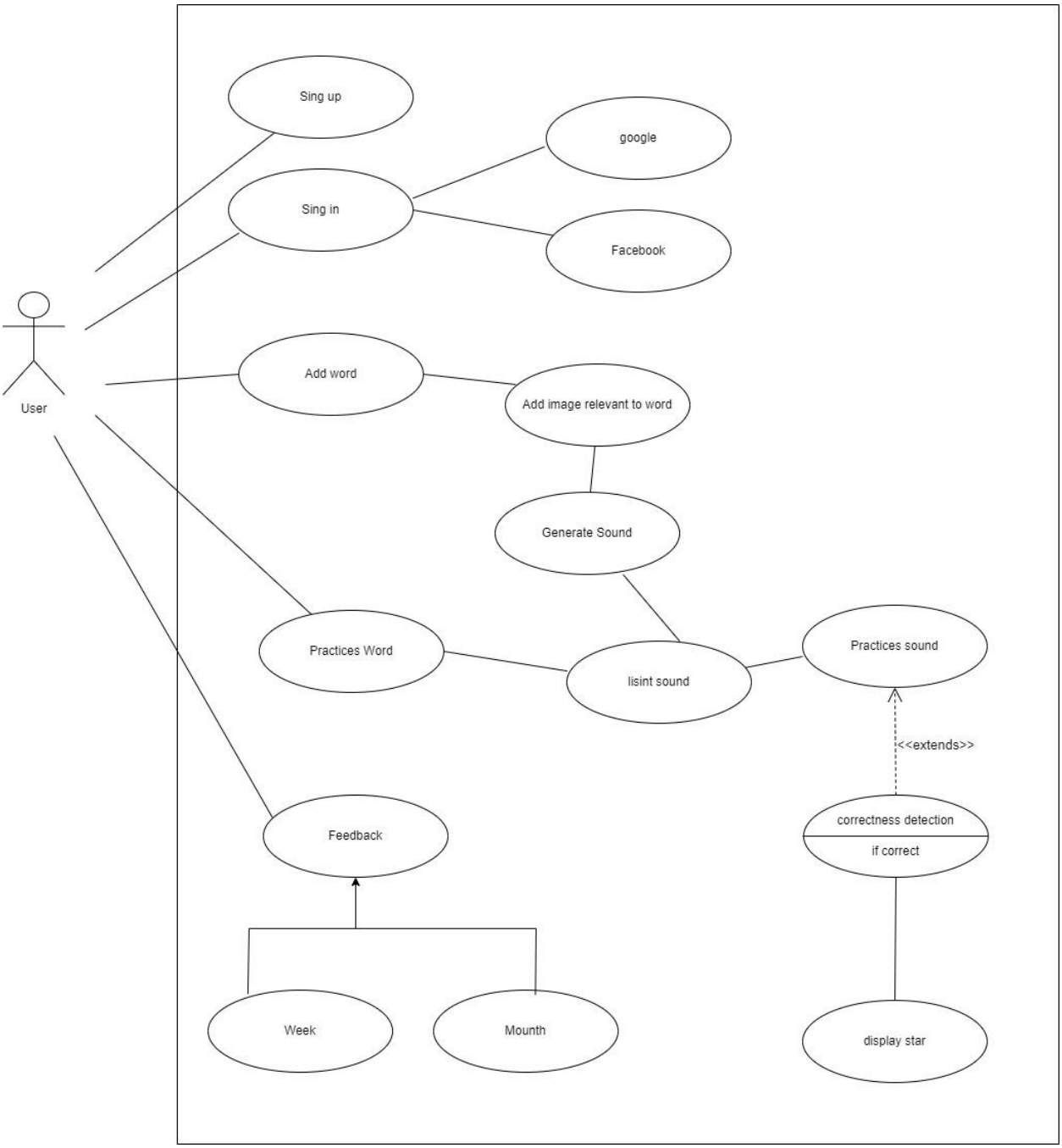


figure 3

4.2 USER CASES

user case	1	
Test title	Practice word sound	
Test Priority	5	
Module Name	Word learn	
Description	User start practice/learn different word	
Per-condition	Valid user	
Primary Actor(s)	user	
Trigger	User need learn new word	
Main Scenario	Step	Action
	1	User need to login system
	2	Check valid user or not
	3	Click button learn
	4	Navigate word set selection page
	5	Select word set
	6	Display first word
	7	Give listing word children
	8	Display massage start learn
	9	Take user speech sound
	10	Start validation
	11	Display correct word and start
Extensions		
	7A	Sound speed change option
	11A	Sound incorrect re pronation massage displaye

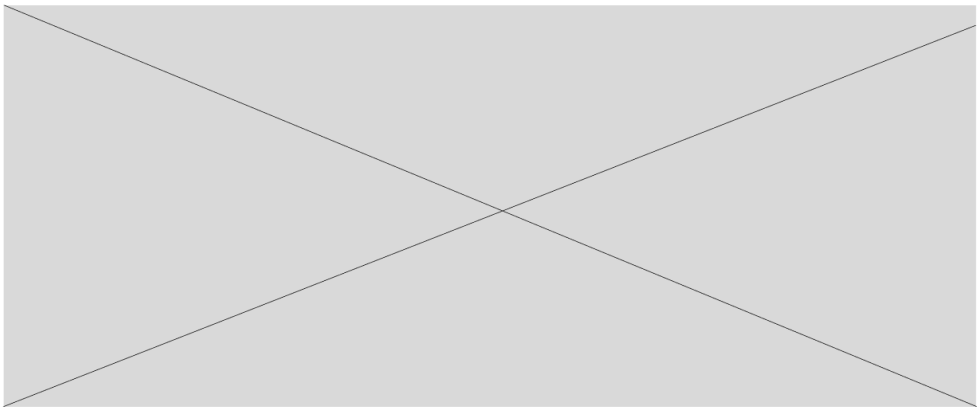
user case	2	
Test title	Add word and great sound	
Test Priority	5	
Module Name	Add user want word	
Description	User have option to add different words	
Per-condition	Valid user	
Primary Actor(s)	user	
Trigger		
Main Scenario	Step	Action
	1	User need to login system
	2	Check valid user or not
	3	Click button add word
	4	Display massage add image
	5	Display word adding page
	6	Add word and save
	7	Generate sound
Extensions		
	2A	Not valid user display error massage
	6A	Add successful massage display

user case	3	
Test title	Feedback check and download	
Test Priority	5	
Module Name	Add user want word	
Description	User can download or check progress	
Per-condition	Valid user	
Primary Actor(s)	user	
Trigger	User need to progress check	
Main Scenario	Step	Action

	1	User need to login system
	2	Check valid user or not
	3	Click progress check
	4	Navigate progress page
	5	Click download
	6	Start download
Extensions		

user case	4	
Test title	Learning word examining	
Test Priority	5	
Module Name	Word examining	
Description	User can download or check progress	
Per-condition	Valid user	
Primary Actor(s)	user	
Trigger	User need to progress check	
Main Scenario	Step	Action
	1	User need to login system
	2	Check valid user or not
	3	Click button start exam
	4	Navigate page
	5	Displayer number question need
	6	Select number question
	7	Start exam
Extensions		
	3A	If not select number question display 5 questions

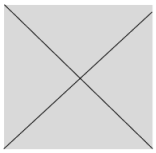
4.3 WIREFRAME (LOW LEVEL)



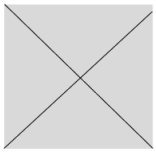
Start Learn

Lorem ipsum dolor sit amet, consectetur
adipiscing elit, sed do eiusmod tempor incididunt
ut labore et dolore magna aliqua.

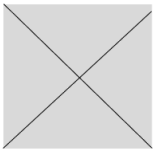
Start free trial



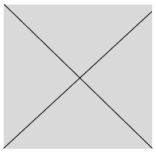
Drawing



Speak



Mathematics



Gross moto skills

TRADE

© 2010 — 2020

Privacy — Terms

Product	Features	Resources	Company
Product	Feature	Resource	Company
Product	Feature	Resource	Company
Product	Feature	Resource	Company

Start Speech

Practise Your word

"If you don't try this app, you won't become the superhero you were meant to be"

List Word
Subtitle
Start Now

Add Word
Subtitle
Start Now

Examine Word
Subtitle
Start Now

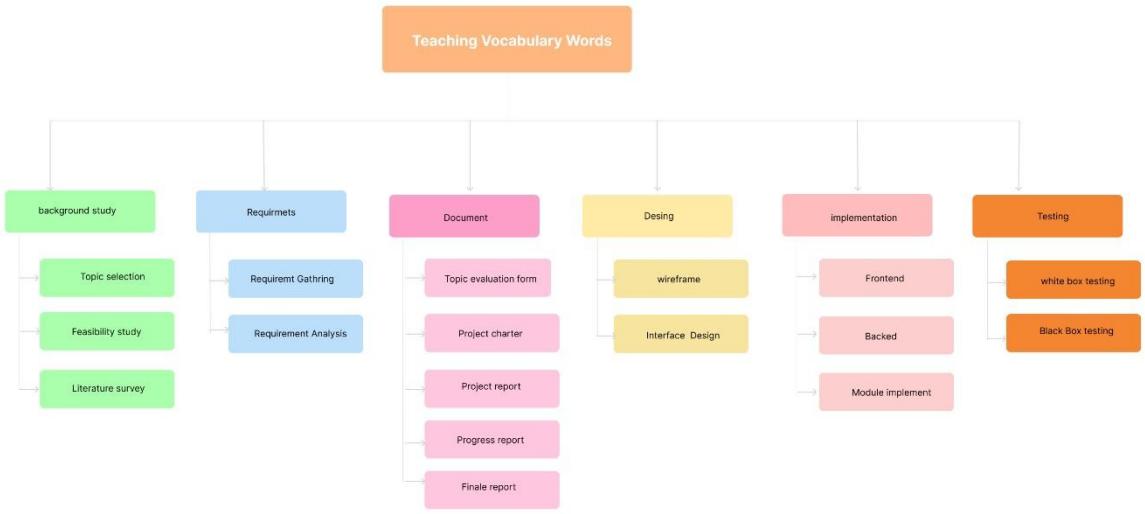
Speech Start

Well Done Great Achievement

4.4 DATASET

This component need dataset are different word pronunciation main goal is teach correct word sound therefor mostly need to word correct sound .expect data collecting visit filed (meeting different down syndrome centers) and find correct sound using internet example Kaggle

WORK BREAKDOWN CHART



PROJECT REQUIREMENTS

Functional Requirement

The system should be able to generate different sound give any word

The system should able to give feedback to children pronation improvement

Different user set different Audio speed set facility relevant image display to Generate sound user expected and any image can be add system with word

Non-functional requirements

1. User-friendliness - The system should provide a wed base application users while maintaining an attractive, responsive and well comprehensive interphase.
2. Reliability - The wed application should not fail or slow at any time throughout the run time User data must be protected and provide better experience wiled use.
3. Performance – provide correct information fastly
4. Availability - The system should be able to be used whenever it is needed any time ant place.

User requirements

- User expected add different word and generate sound it.
- User expected display image every word.
- User expected to Audio balance option different levels give children listing

SYSTEM REQUIREMENTS

- Node -backend
- React – frontend
- Mongo dB
- Firebase – image / video store
- Module taring
    - Python
    - TensorFlow

Expected software Design Architecture is Layard architecture reason to use each lay have specific work and suitable for cloud base implement application

8.Testing and Implementation

8.1 Implementation

The project is developed with a central focus on developing vocabulary skills in children, particularly children with speech and language impairment with down syndrome. With the use of deep learning techniques, the system learns to recognize and evaluate the pronunciation of children and gives feedback and guidelines in real time. the model trained using a most commonly used word corpus for children and ensures age-related content maintained. Activities of the children are engaged and actively involved with the interactive screens and with interactive graphics. Training in deep learning models such as Long Short-Term Memory (LSTM) is done for acceptance of speech inputs with the task of learning pronunciation patterns. So the system does not just make word learning easier, but also allows kids to practice words in a fun and interactive manner for continuous learning and language development

This e-learning tool is designed to improve children's vocabulary knowledge, especially for those with speech and language difficulties, using advanced deep learning techniques. The system utilizes speech recognition and pronunciation analysis to provide real-time feedback, supporting children in learning new words effectively. The web application is developed using a modern technology stack—**React** is used for building a responsive and interactive frontend interface, while the backend logic is handled by both **Express (Node.js)** and **Flask (Python)** to manage APIs, user data, and deep learning model integration. Deep learning models, such as LSTM, are trained to recognize speech patterns and evaluate pronunciation accuracy. This intelligent system delivers an engaging and accessible learning experience, helping children expand their vocabulary in a fun and supportive environment.

For the implementation of the system, the following technologies were used:

- Visual Studio Code
- Google Colab
- MongoDB
- Docker
- Docker Hub
- GitHub
- Firebase

This model uses LSTM. This script aims to implement an LSTM (Long Short-Term Memory) autoencoder because it functions optimally with sequence data by maintaining temporal dependencies. The included code lacks the necessary function build_autoencoder, which should outline the design of the autoencoder. A standard part of this function creates an encoder with LSTM layers that compresses input sequences into latent spaces of specified length (latent_dim) and generates a decoder to recover the original sequences from the compressed data. After compilation, the model is ready for training with suitable optimization and loss functions. Number input 64 dense layer 34 this extract features fist layer using activate function is `relu`

```
# Step 3: Build the LSTM Autoencoder
def build_autoencoder(timesteps, features, latent_dim=32):
    # Encoder
    inputs = Input(shape=(timesteps, features))
    encoded = LSTM(64, return_sequences=False)(inputs)
    encoded = Dense(latent_dim, activation='relu')(encoded)

    # Decoder
    decoded = RepeatVector(timesteps)(encoded)
    decoded = LSTM(64, return_sequences=True)(decoded)
    outputs = TimeDistributed(Dense(features))(decoded)

    # Autoencoder
    autoencoder = Model(inputs, outputs)
    autoencoder.compile(optimizer='adam', loss='mse')
    return autoencoder
```

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 100, 13)	0
lstm (LSTM)	(None, 64)	19,968
dense (Dense)	(None, 32)	2,080
repeat_vector (RepeatVector)	(None, 100, 32)	0
lstm_1 (LSTM)	(None, 100, 64)	24,832
time_distributed (TimeDistributed)	(None, 100, 13)	845

Total params: 47,725 (186.43 KB)
Trainable params: 47,725 (186.43 KB)
Non-trainable params: 0 (0.00 B)

Backend Implementation

At the endpoint route the Flask API receives audio files followed by confident cluster predictions together with their score values. The following is a detailed description of code operation:

Blueprint Setup:

The audio_bp blueprint exists to control all routes related to audio processing functions. By using this structure the program becomes more flexible and simpler to monitor.

Handling POST Request:

The prediction endpoint supports only POST method requests. The route analyzes incoming files through the 'file' not in request.files condition. The endpoint returns a 400 status code together with "No file provided" message when no input file exists.

Preprocessing Audio File:

The function makes an request to process audio preprocess_audio function. The MFCCs extracted by this function provide the model with the correct shaped features for processing. When the preprocessing operation ends in failure or invalid data format (audio processing failure) the function returns a 400 error message.

Latent Feature Extraction:

The processor feeds preprocessed data through encoder_model.predict() to derive the latent features of the audio data. The extracted features allow for predicting the cluster membership of the audio.

Cluster Prediction:

The K-means model receives extracted latent features for cluster prediction generation. The kmeans.predict() function returns cluster indexes predicated by the model. The API uses distance The k-means model receives extracted latent features for generating cluster predictions. The kmeans.predict() function returns cluster indices predicted by the model. The API uses distance measurements to assess prediction accuracy by analyzing the relationship between latent features and cluster center points.

```python
from flask import Blueprint, request, jsonify
from models.lstm_encoder import encoder_model, kmeans
from utils.audio_processing import preprocess_audio

audio_bp = Blueprint("audio_bp", __name__)

@audio_bp.route('/predict', methods=['POST'])
def predict_cluster():
    if 'file' not in request.files:
        return jsonify({"error": "No file provided"}), 400

    file = request.files['file']
    if file.filename == '':
        return jsonify({"error": "No selected file"}), 400

    try:
        # Preprocess the audio file directly from memory
        audio_data = preprocess_audio(file)
        if audio_data is None:
            return jsonify({"error": "Invalid audio file"}), 400

        # Extract latent features using the encoder
        latent_features = encoder_model.predict(audio_data)

        # Predict cluster
        cluster = kmeans.predict(latent_features)[0]

        # Calculate distances to cluster centroids for confidence
        distances = kmeans.transform(latent_features)[0]
        closest_distance = distances[cluster]
        max_distance = distances.max()
        confidence = 100 * (1 - (closest_distance / max_distance))
```

Audio File Preprocess

The preprocess_audio function handles audio processing by extracting MFCCs from files then standardizes their duration before executing normalization on their numerical values. The function includes these core functions which work as follows:

Loading the Audio File:

The function utilizes librosa.load to load the audio file from its input file-like object (file). Librosa.load fails to read audio data when it comes from an io.BytesIO object. Reading audio data from file-like objects becomes easier by employing the soundfile library as a recommended tool for such operations.

Extracting MFCC Features:

The audio signal input generates MFCC feature vectors using librosa.feature.mfcc that create a matrix structure one row per frame.

Padding or Truncating Sequences:

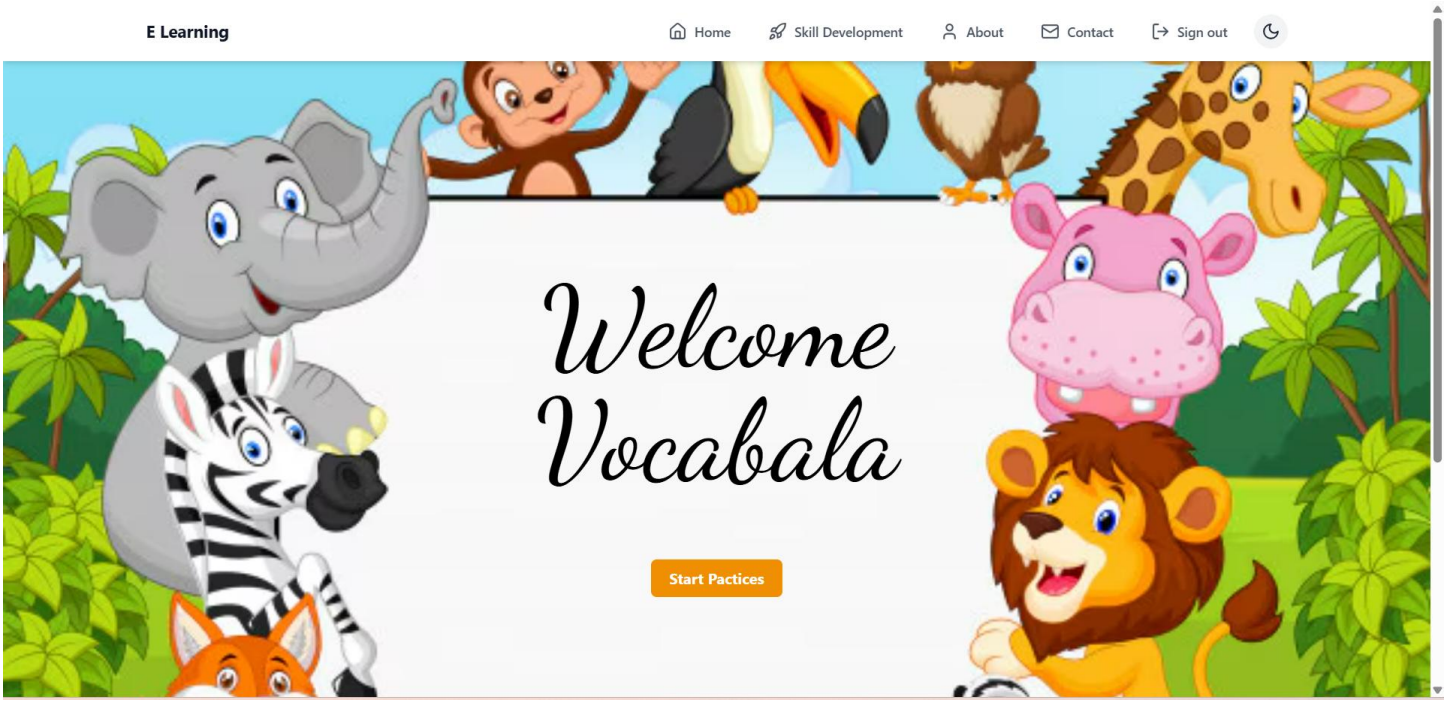
The function adopts a strategy that enforces all MFCC sequences to achieve standard sequence length (max_timesteps). The function pads sequences that measure shorter than max_timesteps with zeros through the use of the numpy.pad function. The time series gets truncated when its duration exceeds max_timesteps.

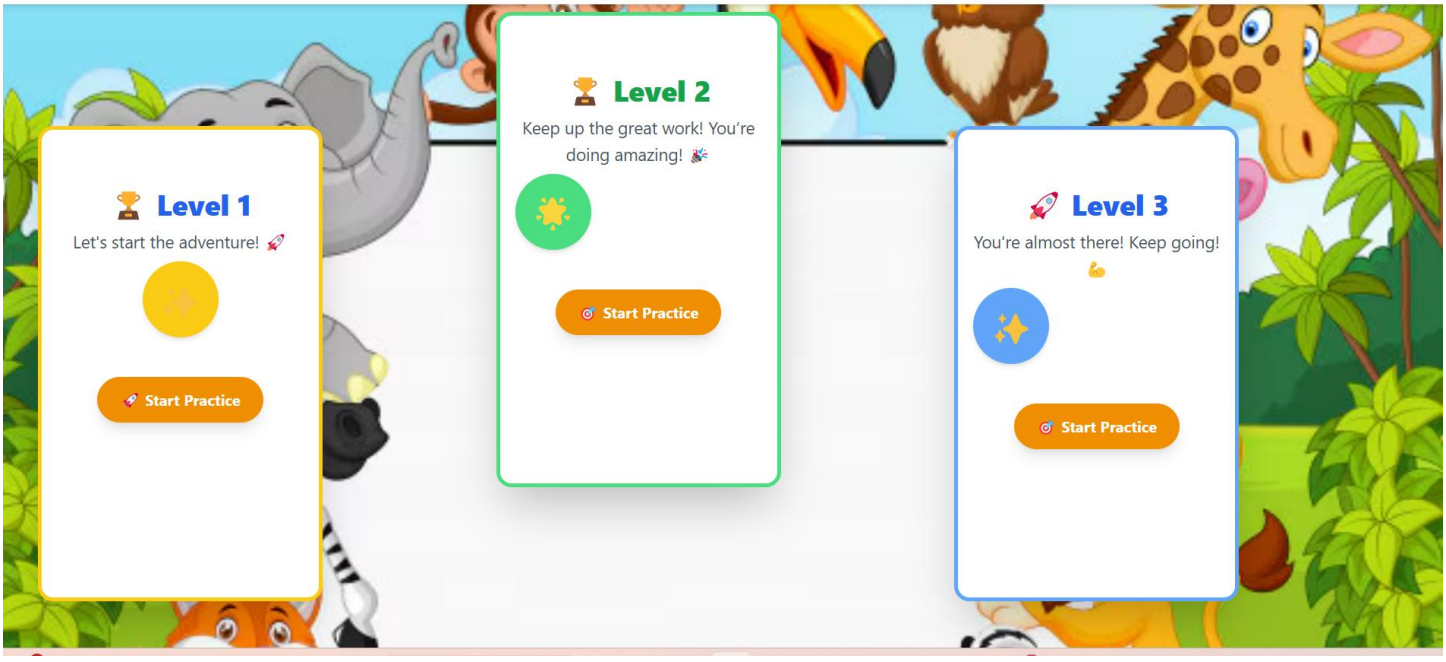
Normalizing the MFCCs:

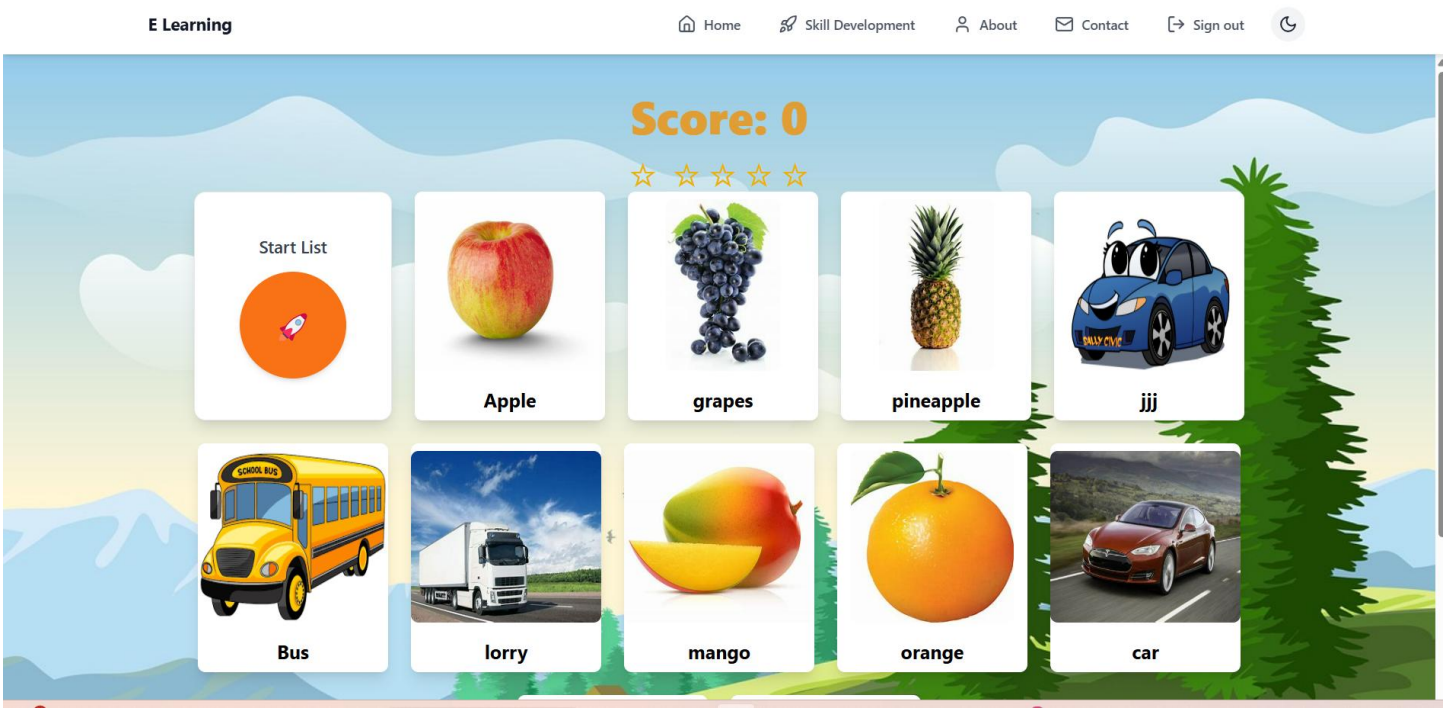
The normalization process utilizes MinMaxScaler to scale MFCC sequence values between 0 and 1 because this helps increase machine learning model performance.

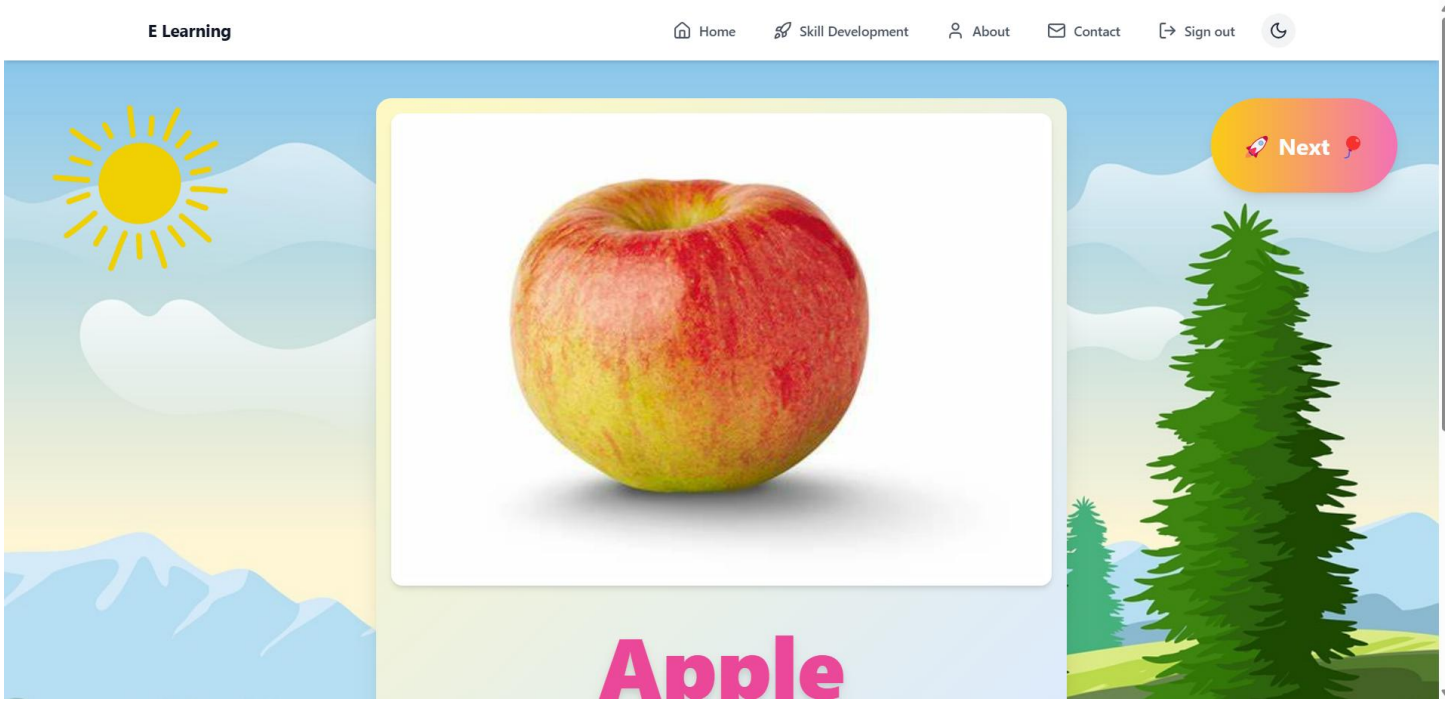
```python
import io
import numpy as np
import librosa
from sklearn.preprocessing import MinMaxScaler

def preprocess_audio(file, sr=16000, n_mfcc=13, max_timesteps=100):
    try:
        audio, _ = librosa.load(io.BytesIO(file.read()), sr=sr)
        mfccs = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=n_mfcc).T

        if len(mfccs) < max_timesteps:
            pad_width = max_timesteps - len(mfccs)
            mfccs = np.pad(mfccs, ((0, pad_width), (0, 0)), mode='constant')
        else:
            mfccs = mfccs[:max_timesteps]

        scaler = MinMaxScaler()
        return np.expand_dims(scaler.fit_transform(mfccs), axis=0)
    except Exception as e:
        print(f"Error processing audio file: {e}")
        return None
```

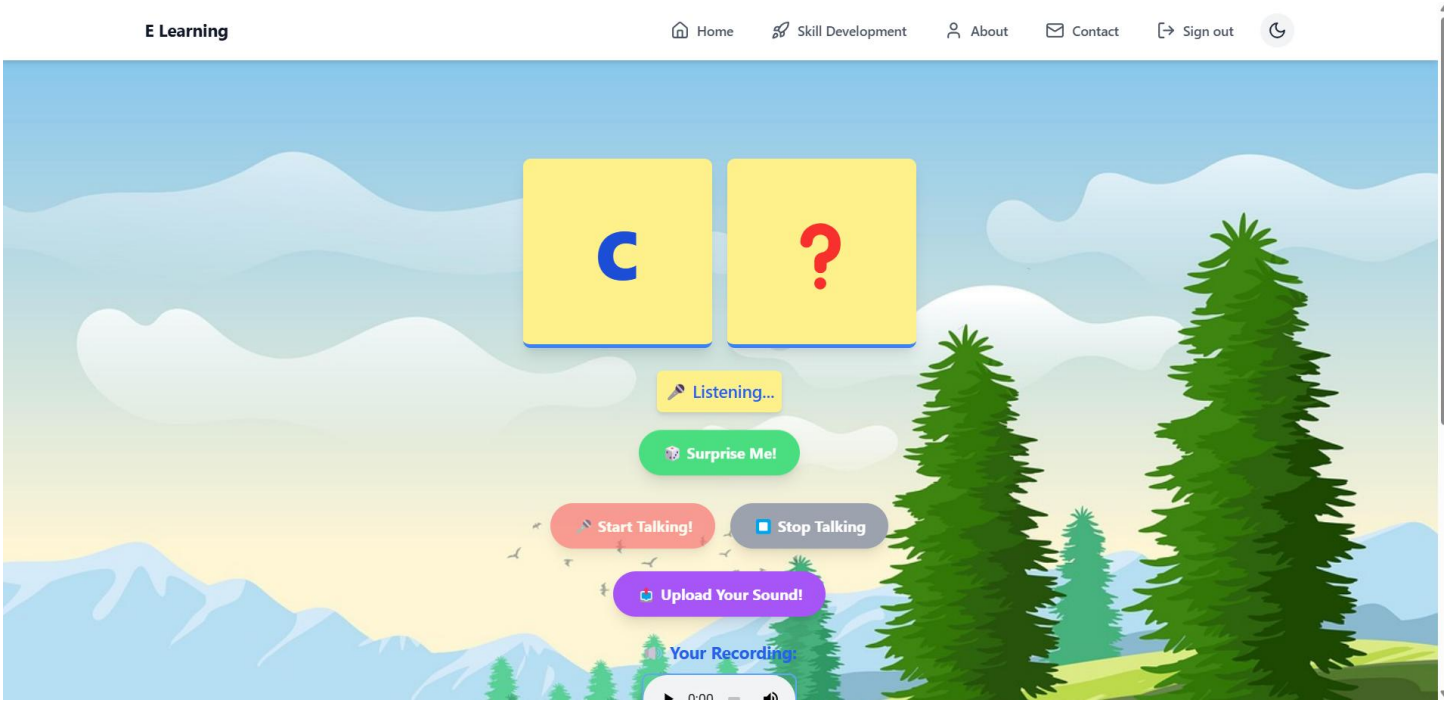
Frontend Implementation

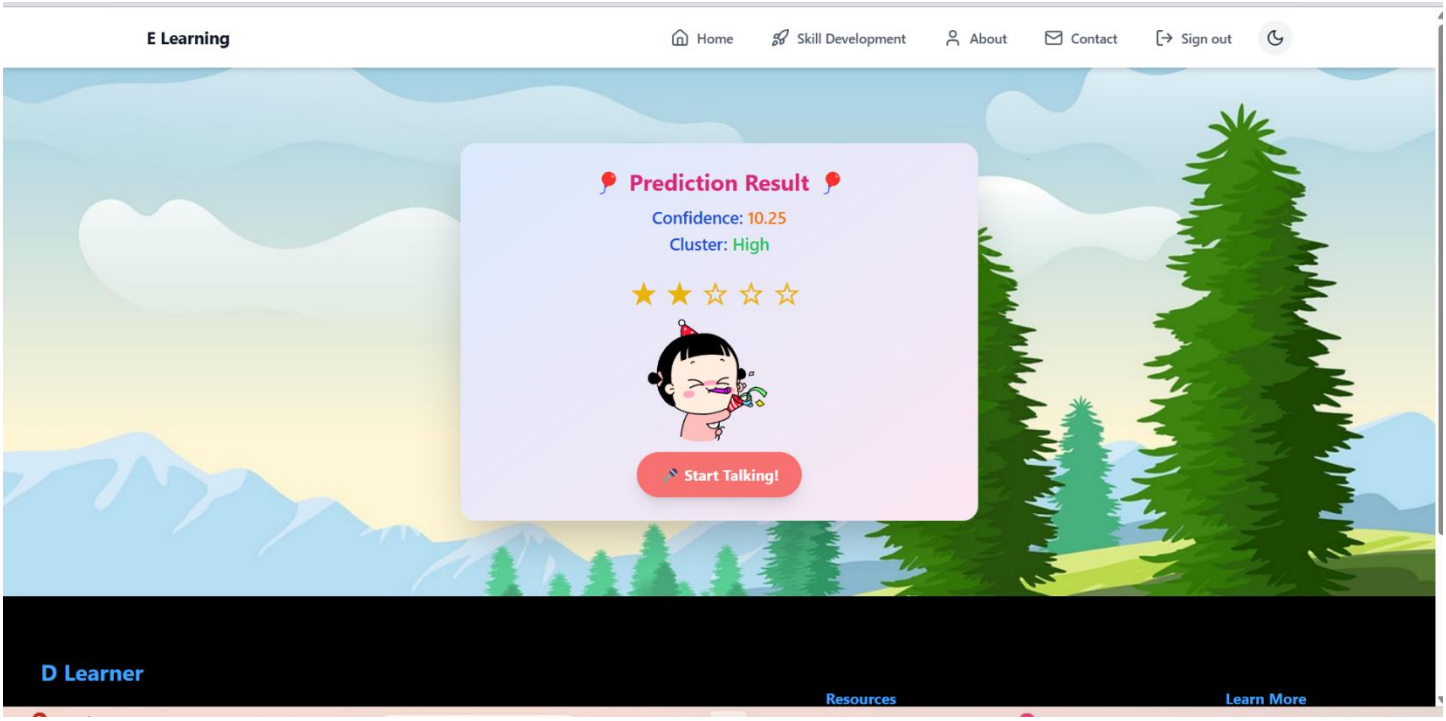





E Learning
Home
Skill Development
About
Contact
Sign out
Score: 0
Start List
Apple
grapes
pineapple
jjj
Bus
lorry
mango
orange
car


E Learning
Home
Skill Development
About
Contact
Sign out
Next
Apple


E Learning
Home
Skill Development
About
Contact
Sign out
c
?
Listening...
Surprise Me!
Start Talking!
Stop Talking
Upload Your Sound!
Your Recording:


E Learning
Home
Skill Development
About
Contact
Sign out
Prediction Result
Confidence: 10.25
Cluster: High
Start Talking!
D Learner
Resources
Learn More

Testing

Test ID	Test Scenario	Input	Expected Outcome
01	After login Navigate Learning page	Enter user name and password click login	Navigate page
02	Click the vocabulary training to navigate the vocabulary training	Click button	Display page successfully
03	Select option practices Level 1 navigate page and display word list	Click the button Level 1 start	Successfully navigate and display word list
04	New Word Adding	Add an image and Word	Successfully add display image
05	Word pronunciation Sound hearing	Click word	Hearing sound
06	Speech word record or not testing	Start learning word Recode voices	Successfully recode voices and display
07	Predict sound quality reactively	Click prediction upload button	Display predication quality level
08	May time faille pronunciation navigate practices level message display	Practices word and predication results low Five time	Successfully display message do you want practice
09	One word learning to navigate the next word	Click next word	Successfully navigate new word
10	Correct answer gives a reward	Correct pronation word sound passing backend to teat	High give five start ,medium give 3 star Low give 1 star display message

Commercialization

Target Audience

- Speech & Language Therapists
- Special Education Teachers
- Parents of Children with Speech Delays
- Inclusive Schools & Therapy Centers

Subscription Plans

Basic :

this start prices range 1500 to 5000 this three pack ages gold , silver and premium

Premium :

This plane several option this cost 8000 only , download audio ,Add new Audio like new facility given

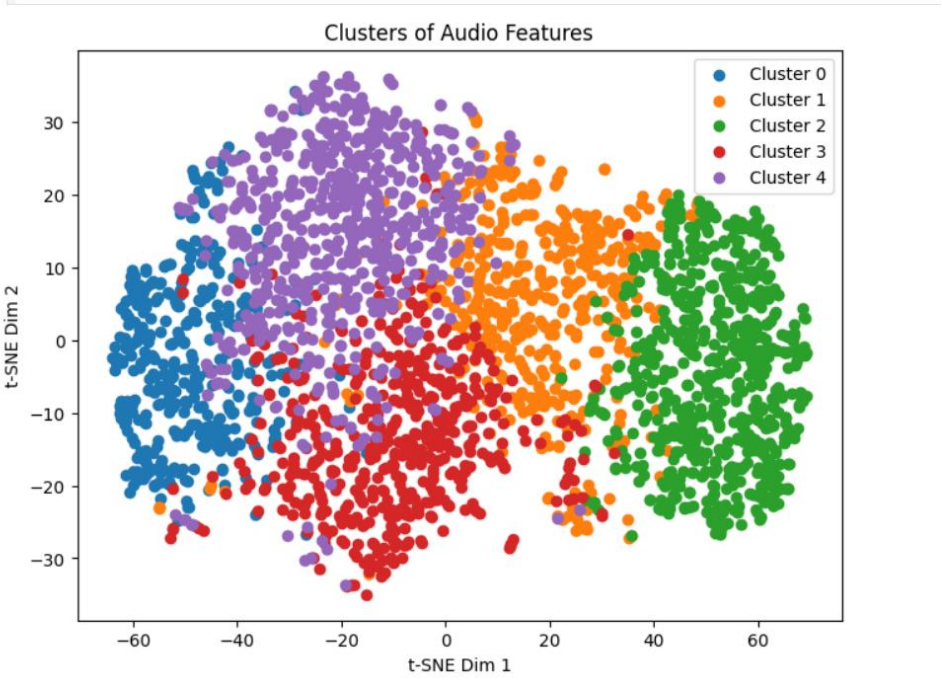
Marketplace:

- No need of advanced knowledge in technology.
- No need of prior knowledge

9 Results & Discussion

9.1 Results

The vocabulary training step targets children with Down syndrome by helping them develop their vocabulary to enhance their pronunciation, focusing on three-stage improvement in word pronunciation. stage one involves learning children how Proper mouth movements, which is essential for developing accurate sound production. It begins with low-intensity sounds and gradually transitions to higher sound ranges, while also teaching correct breathing techniques to optimize word pronunciation
.The second stage emphasizes learning correct word pronunciation. Children improve their efficiency in producing words by continuing to develop their sound production skills until they can speak words with increasing precision.
As illustrated in the diagram shows four clusters. The first three clusters detect sounds with low pronunciation but cluster four analyzes sound production from moderate to high levels. The cluster examines pronunciation levels to award children points according to their achievement.
The third step in developing vocabulary knowledge while improving its quality. This level compares task learned before and day-to-day to day seeing object . learned word set food in learning stages one and two , vocabulary development provided before the never-seeded food list reaches this. this stage every word provide a maximum 10 seconds using this much of time can easy identify think style improvement



9. 2Discussion

The vocabulary training addresses Down syndrome children through programs designed to enhance their vocabulary understanding and speech clarity. The word pronunciation training consists of three incremental stages which assist language development.

Stage One of the learning process teaches children proper movements of their mouth.

The initial teaching phase instructs children about proper facial expressions needed for precise sound generation. At the beginning of this stage the children learn to produce sounds using low volume, then progress to producing sounds of increasing intensity. Proper breathing techniques form part of the curriculum because they help students develop optimal pronunciation skills.

The second stage concentrates on making children form accurate words through proper pronunciation.

During the second stage learning focuses on developing improved word articulation among children. The additional practice of sound production leads students to become both more efficient and precise speakers.Four distinct clusters organize the progression of pronunciation according to the diagram.The focus of Clusters 1 to 3 involves identifying low-pronunciation sounds for improvement.The fourth cluster analyzes elements connected to the production of moderate and strong sound output. The application assesses pronunciation levels and rewards points to the child depending on their advancement.

The development of vocabulary occurs through the third stage of speech development by reinforcing established vocabulary.The third phase combines vocabulary development with word quality enhancement which is carried out through real-world use. At this phase new lesson words are matched with everyday objects which children experience. The newly learned food-related words from stages one and two continue to develop through actual environment associations.

Each word receives no longer than ten seconds of exposure which allows children to quickly recognize and connect the word in their mind. The exposure limit enables better thinking skills development while strengthening recognition and recall for vocabulary items.

10 Conclusion

This project delivers a comprehensive vocabulary training platform which specifically addresses children who face speech and language disabilities together with Down syndrome individuals. LSTM (Long Short-Term Memory) models allow the system to detect speech patterns while evaluating pronunciation accuracy which provides immediate feedback to promote language development. The model training phase employed a word collection containing educational terms which are also frequently used by young students to maintain both educational value and student engagement.

The system merges speech recognition with interactive graphics as well as immediate feedback to develop an interactive e-learning platform for users. The system offers children an interface that simplifies interaction and pronunciation training through a platform that mixes education with fun and accessibility. The tool builds ongoing vocabulary proficiency and speaking competence by providing regular session-based guidance to learners.

The platform was developed with modern technologies where React handles front-end responsiveness and Flask and Express handle back-end processing and model integration. The development and deployment of the system relied heavily on Google Colab, MongoDB, Docker, Firebase along with GitHub and Docker.

An LSTM autoencoder model processed and learned temporal speech features through an implementation with a compression scheme that transforms input sequences of latent variables before reconstruction for validation. The dense layer containing 64 units worked with the ReLU activation function to enhance detail extraction from features.

This AI-driven smart solution makes vocabulary learning easier for speech-challenged children while creating a technological basis that can boost future assistive educational technology development. This development creates conditions to establish customized accessible therapy tools which require scalability for assisting various types of learners.

References

1. S. Silva, A. Jayasinghe, R. Dilshan, C. Dissanayake, S. Vidhanaarachchi, and D. De Silva, "Effectiveness of Deep Learning Technologies to Assist Educational Inclusion for Children with Down Syndrome," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023.

2. H. M. C. H. Herath, N. A. S. B. N. Nissanka, I. M. A. Illankoon, R. P. Madusanka, J. Krishara, and W. Tissera, "EduPlanner –Best Teaching Method for Students with Down Syndrome," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023.

3. A. S. T. Sampath, M. W. Vidanapathirana, T. B. A. Gunawardana, P. W. H. Sandeepani, L. H. S. S. Chandrasiri, and B. Attanayaka, "E-Learning Education System For Children With Down Syndrome," 2022 IEEE 10th Region 10 Humanitarian Technology Conference (R10-HTC), Hyderabad, India, 2022.

4. C. R. Galindo-López, J. Beltrán, C. B. Pérez, A. Macias, and L. A. Castro, "Classifying interactions of parents and children with Down syndrome in educational environments using deep learning," 2022 IEEE Mexican International Conference on Computer Science (ENC), Xalapa, Veracruz, Mexico, 2022.

5. S. Wellala, S. A. Thathsarani, D. Senaratne, P. Samaranayake, and A. Jayakody, "Assistive Learning Platform for Children with Down Syndrome," 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2020.

6. W. M. Shalash, S. AlTamimi, E. Abdu, and A. Barom, "No Limit: A Down Syndrome Children Educational Game," 2018 IEEE Games, Entertainment, Media Conference (GEM), Galway, Ireland, 2018.

7. C. Cunha-Pérez, M. Arevalillo-Herráez, and D. Arnau, "Design and Evaluation of a Set of Methodological Strategies for Learning a Second Language in Students With Down Syndrome Using Computer-Based Instruction," in IEEE Transactions on Learning Technologies, vol. 17, pp. 172-180, 2024.

8. O. Agbolade, A. Nazri, R. Yaakob, and Y. K. Cheah, "Homologous Anatomical-Based Facial-Metrics Application to Down Syndrome Face Recognition," in IEEE Access, vol. 11, pp. 104879-104889, 2023.

9. E. Lyakso, O. Frolova, V. Gorodnyi, A. Grigorev, A. Nikolaev and Y. Matveev, "Reflection of the Emotional State in the Characteristics of Voice and Speech of Children with Down Syndrome," 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Timisoara, Romania, 2019.

10.Bertoli, M., Biasini, G., Calignano, M. T., Celani, G., De Grossi, G., Digilio, M. C., … Zuccalà, G. (2011). Needs and challenges of daily life for people with Down syndrome residing in the city of Rome, Italy. Journal of Intellectual Disability Research, 55(8), 801–820.

11.Caselli, M. C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. Journal of Speech, Language, and Hearing Research, 41, 1125–1135. doi:10.1044/jslhr.4105.1125.

12.I. Escobar, A. Guamán, J. Montaluisa, E. Pruna and Y. Marín, "Mobile application for vowel learning in children with Down Syndrome “LVDS-App”," 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), Pucon, Chile, 2017.