0x.Day

Pondy Edition 1

Speaksi

Presented by: **x2X Coders**

Idea behind Speaksi

Speaksi is an app designed to assist individuals with speech impairments by converting unclear speech into text and providing real-time feedback on pronunciation. It uses deep learning for speech-to-text conversion, Google API for text-to-speech, and offers gamified speech therapy features to track progress. The app enhances communication, boosts confidence, and supports skill development through a user-friendly interface and advanced AI models.

Problem

Millions worldwide face challenges in communication due to speech impairments caused by neurological and developmental disorders like Cerebral Palsy, Parkinson's Disease, Aphasia, Dysarthria, and Multiple Sclerosis. These conditions affect speech clarity and production, significantly impacting everyday interactions.

Cerebral Palsy:

Impairs control of speech muscles, causing slurred, slow, or unclear speech. Severe cases may result in difficulty producing speech altogether.

Parkinson's Disease:

Causes soft, slurred, or monotone speech due to weakened vocal muscles and reduced control.

Aphasia

Impacts the ability to understand or produce speech, often causing word-finding difficulties or comprehension issues.

Dysarthria:

Weak or uncoordinated speech muscles lead to slurred, strained, or slow speech, affecting clarity.

Multiple Sclerosis:

Nerve damage disrupts muscle control, leading to slurred, nasal, or uneven speech patterns.

Cerebral Palsy

- ·Approximately **18 million people** worldwide are affected by some form of cerebral palsy.
- ·Speech Symptoms:
  - Delays in speech development
  - Trouble speaking

Parkinson's Disease

- ·Over **8.5 million individuals** globally are living with Parkinson's disease.
- ·Speech Symptoms:
  - Soft voice: Quieter and harder to hear
  - Flat tone: Dull, little change in pitch
  - Unclear speech: Words slurred or hard to understand
  - Awkward pauses or uneven timing

Dysarthria

- ·It is estimated that **over 1 million people** worldwide are affected by dysarthria,
- ·Speech Symptoms:
  - Slurred or mumbled speech
  - Speaking too quickly or slowly
  - Unintended volume changes
  - Robotic or monotone speech

Aphasia

- It is often associated with other conditions such as stroke, which affects **approximately 15 million people** worldwide
- ·Speech Symptoms:
  - Short or nonsensical sentences
  - Substitution of words or sounds
  - Trouble understanding conversations or reading

Multiple Sclerosis

- Multiple sclerosis affects an estimated 2.8 million people globally.
- Speech Symptoms:
  - Slurred speech
  - Memory and understanding issues
  - Mood changes affecting communication

Approximately 45 to 50 million people worldwide are affected by conditions such as Cerebral Palsy, Parkinson's Disease, Aphasia, Dysarthria, and Multiple Sclerosis, highlighting the significant global impact of disorders associated with speech and communication.

Solution:

Speaksi is an innovative app for individuals with dysarthria, converting speech to accurate text and providing real-time pronunciation feedback. It uses deep learning and Google API for speech-to-text and text-to-speech. The app also offers speech therapy with practice tools, feedback, and progress tracking through points and levels.

Speech
Recognition Model

Sentiment
Analysis

Magic Spell

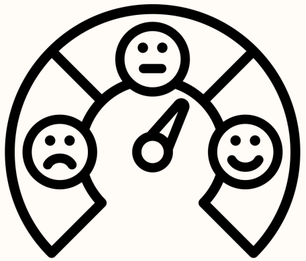
Features:

Speech Recognition Model

Speech recognition models convert spoken language into text using deep learning techniques like RNNs and Transformers. They involve audio preprocessing, feature extraction, and text generation, enabling applications in virtual assistants, transcription services, and accessibility tools.

Magic Spell

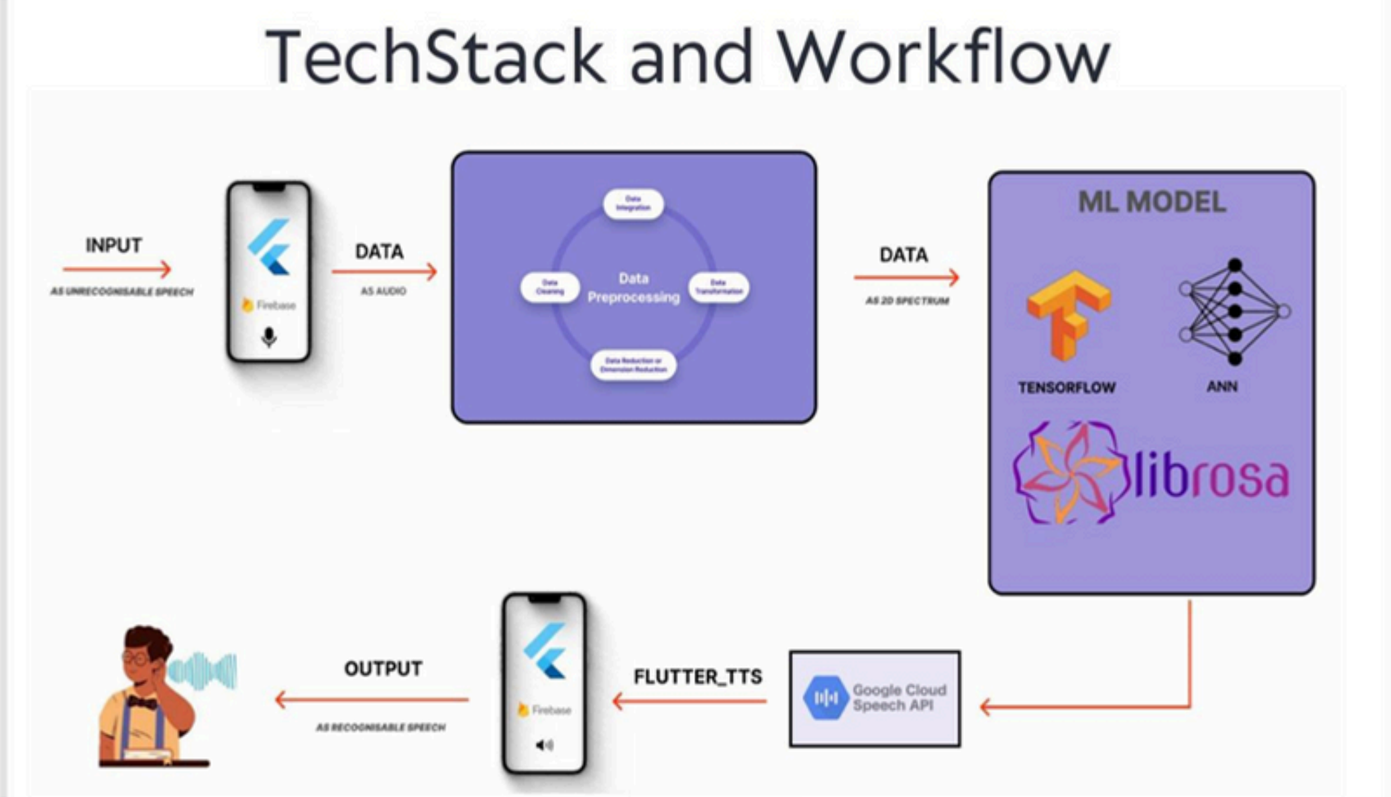
The "Magic Spell" feature enhances interactivity by simulating spellcasting through voice commands or gestures. It uses voice activation, gesture recognition, and visual/audio effects, making it ideal for gaming, AR/VR, and educational environments.



Sentiment Analysis

Sentiment analysis detects emotions in text or speech, classifying them as positive, negative, or neutral. Leveraging NLP techniques, it is used in feedback systems, mental health monitoring, and gamification to enhance user engagement and experienc

Workflow


TechStack and Workflow
INPUT
AS UNRECOGNISABLE SPEECH
Firebase
DATA
AS AUDIO
Data Integration
Data Cleaning
Data Preprocessing
Data Transformation
Data Reduction or Dimension Reduction
DATA
AS 2D SPECTRUM
ML MODEL
TENSORFLOW
ANN
librosa
Google Cloud Speech API
FLUTTER_TTS
Firebase
OUTPUT
AS RECOGNISABLE SPEECH

Tech Stacks

Frontend

Flutter for user interface and interaction.

Backend:

Firebase for audio data storage and management.

Preprocessing

Librosa for audio processing.

Machine Learning

TensorFlow and ANN for speech enhancement.

Output Processing

Google Cloud Speech API and Flutter TTS for text-to-speech conversion.

Future Development

1.Therapy Clinics: Collaborate with speech therapy clinics to integrate the app into their services.

2.Accessibility Partnerships: Work with organizations focused on accessibility to enhance the app’s usability and outreach.

3.Revenue Diversification: Leverage partnerships and premium features to ensure a sustainable and profitable business model.

4.Exploring Subscription Tiers: Offer tiered plans catering to diverse user needs, including basic and premium options.

5.Premium Features: Provide advanced speech therapy tools and personalized feedback for higher-tier subscribers.

6.Healthcare Collaborations: Partner with hospitals and rehabilitation centers to promote the app among patients