



## IMPLEMENTASI ARTIFICIAL INTELLIGENCE MELALUI SPEECH-TO-TEXT SEBAGAI ALAT BANTU TUNARUNGU BERKOMUNIKASI

Efrans Firdaus<sup>1</sup>.

<sup>1</sup>Fakultas Teknik, Universitas Singaperbangsa Karawang

Email: [2010631160055@student.unsika.ac.id](mailto:2010631160055@student.unsika.ac.id)

| Informasi Artikel   | ABSTRACT   |
|---|--|
| <p><b>Riwayat artikel :</b><br/>           Disubmit : 11 Juli 2024<br/>           Direvisi : 18 Juli 2024<br/>           Diterima : 23 November 2024<br/>           Dipublikasi : 20 Desember 2024</p> <p><b>Keywords:</b><br/>           ASL, Artificial Intelligence, NLP, Speech-to-Text, STT, Raspberry Pi.</p> | <p><i>As a reason for research based on the author's interview with one of the Special Assistant Teachers (GPK) of SLB B Tunas Harapan Karawang on Friday, November 10 2023, it was said that Deaf students experienced several obstacles in the areas of socialization, social interaction, communication and cooperation because they were still minimal in vocabulary, Difficulty interpreting words or sentences that contain figurative meaning, and Difficulty interpreting abstract words. The aim of the research is to make things easier when individuals experience limitations in communication, in this case they have problems with hearing abilities or are deaf, which creates their own obstacles to carrying out the social interaction process. The research method or approach based on previous researchers carried out on the ASL language translation system is that it is able to display letters resulting from data processing in the system onto an LCD contained in the 1Sheeld application. The research results are to find out the implementation of Speech-to-Text Software, how to initialize and record sound, testing variations in speaking speed, configuring the Speech-to-Text program with Raspberry Pi, testing the Speech-to-Text system, and implementing hardware in the Speech-to-Text subsystem. The conclusion of this research is to answer the problems faced by Deaf students at SLB B Tunas Harapan Karawang in communicating with normal people.</i></p>   |
| <p><b>Kata Kunci:</b><br/>           ASL, Kecerdasan Buatan, NLP, Pengenalan Suara, STT, Raspberry Pi.</p>  | <p><b>ABSTRAK</b></p> <p>Sebagai alasan penelitian yang berdasarkan wawancara penulis dengan salah seorang Guru Pendamping Khusus (GPK) SLB B Tunas Harapan Karawang pada Jumat, 10 November 2023, dikatakan bahwa siswa Tunarungu mengalami beberapa kendala dalam bidang sosialisasi, interaksi sosial, komunikasi dan kerja sama dikarenakan masih minim dalam kosa kata, Sulit mengartikan kata atau kalimat yang mengandung arti kiasan, dan Sulit mengartikan kata-kata yang bersifat abstrak. Adapun tujuan penelitian supaya mempermudah ketika individu mengalami keterbatasan dalam komunikasi, pada hal ini memiliki permasalahan kemampuan mendengar atau Tunarungu sehingga memunculkan hambatan tersendiri untuk melakukan proses interaksi sosial. Metode atau pendekatan penelitian berdasarkan peneliti sebelumnya yang dilakukan terhadap sistem penerjemah bahasa ASL yakni mampu menampilkan huruf hasil pengolahan data pada sistem ke dalam sebuah LCD yang terdapat pada aplikasi 1Sheeld. Adapun hasil Penelitian untuk mengetahui implementasi <i>Software Speech-to-Text</i>, bagaimana inialisasi dan perekaman suara, pengujian variasi kecepatan berbicara, konfigurasi program <i>Speech-to-Text</i> dengan Raspberry Pi, pengujian sistem <i>Speech-to-Text</i>, hingga implementasi <i>Hardware</i> pada subsistem <i>Speech-to-Text</i>. Kesimpulan dari penelitian ini untuk menjawab permasalahan yang dihadapi siswa Tunarungu di SLB B Tunas Harapan Karawang dalam berkomunikasi dengan orang-orang yang normal.</p> |





## PENDAHULUAN

Kemajuan teknologi dalam bidang kecerdasan buatan (*Artificial Intelligence/AI*) telah membawa dampak signifikan pada berbagai aspek kehidupan, termasuk dalam mendukung kebutuhan khusus penyandang disabilitas. Salah satu implementasi teknologi AI yang berkembang pesat adalah sistem penerjemah bahasa isyarat yang mampu mengubah isyarat tangan menjadi teks atau suara. Penelitian sebelumnya menunjukkan bahwa sistem penerjemah bahasa *American Sign Language* (ASL) berhasil menampilkan hasil pengolahan data berupa huruf pada layar LCD melalui aplikasi seperti 1Sheeld. Fitur *Text-to-Speech* yang tersedia dalam aplikasi tersebut memungkinkan hasil terjemahan teks untuk diubah menjadi suara, menciptakan komunikasi yang lebih inklusif.

Salah satu inovasi yang relevan adalah pengembangan subsistem antarmuka berbasis AI yang melibatkan perangkat seperti speaker dan layar ponsel. Sistem ini memungkinkan penerjemahan bahasa isyarat menjadi audio dan teks secara *real-time*, sehingga memperluas akses komunikasi bagi penyandang tunarungu dan tuna wicara. Teknologi ini tidak hanya memudahkan pengguna dalam berkomunikasi, tetapi juga membantu masyarakat umum untuk lebih memahami kebutuhan komunikasi penyandang disabilitas. Hal ini mencerminkan potensi besar AI dalam mengurangi hambatan komunikasi dan meningkatkan inklusivitas sosial.

Selain aplikasi berbasis perangkat lunak, implementasi teknologi ini juga telah dilakukan melalui perangkat keras, seperti sarung tangan pintar berbasis Raspberry Pi. Sarung tangan pintar tersebut dilengkapi dengan sensor dan sistem pemrosesan data untuk mendeteksi gerakan tangan, yang kemudian diubah menjadi format yang dapat dipahami oleh orang lain, baik berupa teks maupun suara. Inovasi ini telah menunjukkan manfaat signifikan dalam meningkatkan efisiensi komunikasi antarindividu dengan kebutuhan komunikasi khusus.

Namun, meskipun inovasi ini telah menunjukkan hasil yang menjanjikan, beberapa tantangan masih perlu diatasi, seperti akurasi penerjemahan, kecepatan pemrosesan, dan adaptasi terhadap berbagai dialek bahasa isyarat. Sistem yang ada sering kali belum sepenuhnya mendukung penerjemahan bahasa isyarat lokal atau regional yang memiliki kekhasan tersendiri. Oleh karena itu, diperlukan penelitian lebih lanjut untuk mengembangkan sistem yang lebih adaptif dan dapat digunakan secara luas di berbagai komunitas.

Penelitian ini bertujuan untuk mengimplementasikan teknologi AI melalui sistem *Speech-to-Text* sebagai alat bantu komunikasi bagi penyandang tunarungu. Berbeda dari pendekatan sebelumnya, penelitian ini fokus pada pengembangan sistem yang lebih praktis dan fleksibel dengan memanfaatkan perangkat portabel. Sistem ini dirancang untuk menerjemahkan percakapan langsung ke dalam teks





secara *real-time*, sehingga penyandang tunarungu dapat lebih mudah memahami percakapan tanpa memerlukan perantara.

Dengan pengembangan ini, diharapkan teknologi yang dihasilkan tidak hanya memberikan solusi komunikasi yang efektif bagi penyandang tunarungu, tetapi juga dapat diintegrasikan dengan berbagai perangkat pintar lainnya. Hal ini sejalan dengan tujuan jangka panjang untuk menciptakan lingkungan yang lebih inklusif dan memberdayakan penyandang disabilitas melalui penerapan kecerdasan buatan. Artikel ini akan membahas proses implementasi sistem, hasil pengujian, serta potensi pengembangan lebih lanjut.

## METODE PENELITIAN

Metodologi penelitian meliputi metode kombinasi dengan menggabungkan berbagai metode pencarian Studi Literatur, Desain dan Pengembangan Sistem, serta Eksperimen dan Pengujian untuk mendapatkan hasil yang komprehensif dengan mengintegrasikan studi literatur dengan eksperimen lapangan dan menggunakan hasil evaluasi kinerja sistem untuk menginformasikan desain dan pengembangan lebih lanjut.

Pencarian Studi Literatur untuk Mengidentifikasi teknologi *Speech-to-Text* yang ada dan bagaimana mereka telah digunakan untuk membantu komunikasi tunarungu dengan langkah-langkah seperti melakukan pencarian literatur di database akademik meliputi menganalisis artikel jurnal, konferensi, dan buku yang relevan.

Kemudian, Desain dan Pengembangan Sistem untuk mengembangkan prototipe aplikasi atau perangkat *Speech-to-Text* berbasis AI dengan platform dan bahasa pemrograman yang mengintegrasikan API *Speech-to-Text*, yakni *Google Speech-to-Text*. Selanjutnya, mengembangkan antarmuka pengguna yang mudah digunakan oleh tunarungu dan melakukan pengujian internal untuk memastikan fungsionalitas dasar.

Pada Eksperimen dan Pengujian untuk menguji efektivitas sistem yang dikembangkan dalam membantu komunikasi tunarungu dengan merancang skenario komunikasi yang akan diuji seperti percakapan sehari-hari, situasi formal, dan lain sebagainya menggunakan metode kuantitatif untuk mengevaluasi performa sistem dan menganalisis data hasil pengujian untuk menilai efektivitas dan area yang perlu diperbaiki.

Terakhir, Evaluasi Kinerja Sistem untuk mengevaluasi kinerja teknis sistem *Speech-to-Text* dengan menilai kecepatan proses konversi suara ke teks dan mengidentifikasi faktor-faktor yang mempengaruhi kinerja sistem, diantaranya kualitas suara dan kebisingan latar belakang.





## HASIL DAN PEMBAHASAN

### A. Studi Literatur Sistem *Speech-to-Text* (STT)

*Speech-to-Text* adalah sebuah teknologi yang tercipta dari perpaduan antara Artificial Intelligence (AI) dan *Natural Language Processing* (NLP), yang dapat mengenali, menganalisis serta mensintesis percakapan manusia asli. Secara singkat cara kerjanya adalah sistem akan menganalisis lalu mengkonversi audio menjadi teks atau tulisan. Kehadiran *Speech-to-Text* terlebih lagi *Speech-to-Text* bahasa Indonesia, tentu menjadi angin segar bagi dosen dan mahasiswa disabilitas runtu.

### B. Implementasi Sistem *Speech-to-Text* (STT) pada Raspberry Pi

Sistem *Speech-to-Text* (STT) pada Raspberry Pi dirancang untuk mengonversi input suara dari pengguna menjadi teks yang dapat diproses lebih lanjut oleh aplikasi atau sistem lain. Proses ini dimulai dengan inisialisasi komponen-komponen yang dibutuhkan, seperti modul STT dan *microphone*. Setelah itu, suara dari pengguna direkam dan diproses untuk menghilangkan noise serta meningkatkan kualitas audio.

Teknologi pengenalan suara kemudian digunakan untuk menganalisis audio yang telah diproses, mengubahnya menjadi teks yang dapat dimengerti oleh komputer. Hasil teks ini dapat ditampilkan atau disimpan untuk digunakan dalam aplikasi yang lebih kompleks, seperti integrasi dengan sistem sign language atau platform IoT seperti Firebase. Dengan alur ini, sistem STT pada Raspberry Pi mampu mengenali dan mengonversi suara dengan akurasi tinggi, mendukung berbagai aplikasi yang memerlukan interaksi suara-ke-teks yang efektif dan efisien.

Tabel 1. Algoritma Sistem *Speech-to-Text*

| No | Algoritma          | Analisis   |
|----|--------------------|--|
| 1  | Inisialisasi suara | Inisialisasi memastikan bahwa semua komponen sistem siap menerima input suara. Hal ini penting untuk mencegah kesalahan selama operasi dan memastikan stabilitas sistem. |
| 2  | Perekaman Suara    | Proses perekaman suara memungkinkan sistem menangkap input suara pengguna secara real-time, namun kualitas rekaman bisa dipengaruhi oleh lingkungan sekitar.             |
| 3  | Preprocessing      | Preprocessing meningkatkan kualitas audio yang akan dianalisis, namun membutuhkan waktu tambahan yang bisa memperlambat respons sistem.                                  |
| 4  | Speech Recognition | Penggunaan teknologi pengenalan suara memungkinkan konversi audio ke teks dengan akurasi tinggi, tetapi memerlukan model yang dilatih dengan baik.                       |
| 5  | Output Text        | Output teks memungkinkan integrasi mudah dengan aplikasi lain, tetapi akurasinya tergantung pada kualitas perekaman dan preprocessing.                                   |

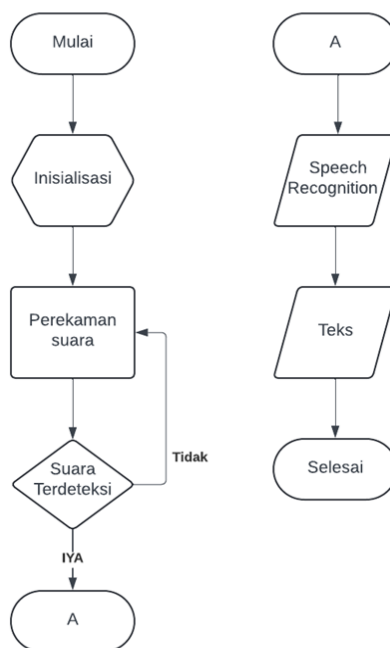




Kelebihan dan kekurangan dari beberapa algoritma berikut :

Tabel 2. Perbandingan Penggunaan Algoritma Sistem *Speech-to-Text*

| No | Algoritma          | Kelebihan   | Kekurangan  |
|----|--------------------|---|---|
| 1  | Inisialisasi Suara | Memastikan semua komponen siap untuk menerima input suara.                                | Memakan waktu yang bisa menyebabkan penundaan dalam respons awal.                                   |
| 2  | Perekaman Suara    | Memungkinkan interaksi langsung antara pengguna dan sistem.                               | Lingkungan sekitar dapat mempengaruhi kualitas rekaman.   |
| 3  | Preprocessing      | Menghilangkan noise dan gangguan lainnya, sehingga meningkatkan akurasi pengenalan suara. | Membutuhkan waktu tambahan yang bisa memperlambat respons sistem.                                   |
| 4  | Speech Recognition | Mengubah audio menjadi teks dengan tingkat akurasi yang tinggi.                           | Memerlukan model yang dilatih dengan baik, membutuhkan data latihan dalam jumlah besar dan beragam. |
| 5  | Output Text        | Hasil teks dapat dengan mudah diintegrasikan ke dalam aplikasi lain.                      | Akurasi teks sangat tergantung pada kualitas perekaman dan preprocessing.                           |



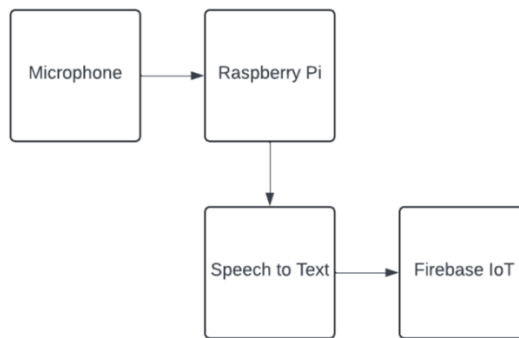
Gambar 1. Flowchart *Speech-to-Text*





Sistem *Speech-to-Text* dalam diagram ini melibatkan proses di mana sistem memulai dengan inialisasi perangkat, seperti mikrofon, untuk merekam suara pengguna. Setelah suara direkam, sistem mendeteksi adanya sinyal suara. Jika suara terdeteksi, proses berlanjut ke pengenalan suara menggunakan teknologi *speech recognition*, yang menganalisis pola suara dan mengubahnya menjadi teks. Teks yang dihasilkan merupakan representasi dari apa yang diucapkan dan dapat digunakan untuk keperluan lebih lanjut, seperti dokumentasi atau sebagai input ke sistem lain. Proses ini berakhir dengan penyelesaian, di mana teks siap digunakan atau disimpan.

Berikut Subsistem *Speech-to-Text* :



Gambar 2. Subsistem *Speech-to-Text*

Tabel 3. Penjelasan Subsistem *Speech-to-Text*

| No | Parameter | Keterangan   |
|----|-----------|--|
| 1  | Input     | <ul style="list-style-type: none"> <li>• Ucapan dari pengguna.</li> <li>• Data suara dari Microphone.</li> <li>• Teks hasil konversi suara dari Speech-to-Text</li> </ul>  |
| 2  | Output    | <ul style="list-style-type: none"> <li>• Data suara.</li> <li>• Teks hasil konversi suara.</li> <li>• Data yang disimpan di Firestore.</li> </ul>  |
| 3  | Fungsi    | <ul style="list-style-type: none"> <li>• Mengubah suara menjadi sinyal digital dan mengirimkan data ini ke Speech-to-Text</li> <li>• Mengubah suara menjadi teks menggunakan pustaka Speech Recognition.</li> <li>• Mengirim data ke Firestore IoT untuk penyimpanan dan monitoring</li> </ul> |

C. Implementasi *Hardware* Subsistem *Speech-to-Text*



Implementasi *hardware* subsistem *Speech-to-Text* pada proyek sistem mencakup pemasangan dan konfigurasi berbagai komponen elektronik yang memungkinkan pengumpulan dan pengolahan data suara dari pengguna.



Gambar 3. Pemasangan *Microphone*

#### D. Konfigurasi Program *Speech-to-Text* dengan Raspberry Pi

Program ini adalah sebuah skrip Python yang menggunakan modul *speech\_recognition* untuk mendeteksi suara dari mikrofon.

Berikut skrip program dibawah ini :

```
import speech_recognition as sr

def main():
    # Inisialisasi recognizer
    recognizer = sr.Recognizer()

    # Set threshold energi dinamis untuk menyesuaikan kepekaan mikrofon
    recognizer.dynamic_energy_threshold = True

    try:
        while True:
            # Gunakan microphone sebagai sumber
            with sr.Microphone() as source:
                print("Listening...")
                recognizer.adjust_for_ambient_noise(source, duration=1) # Sesuaikan ambang energi

            # Dengarkan input suara
            audio = recognizer.listen(source)

            try:
                print("Processing...")
                # Mengenali suara menggunakan Google Speech Recognition dalam bahasa Indonesia
                detected_text = recognizer.recognize_google(audio, language="id")
            except sr.UnknownValueError:
                print("Speech Recognition could not understand the audio")
            except sr.RequestError:
                print("Could not request results from Google Speech Recognition service")

    except KeyboardInterrupt:
        print("Program ended by user")
```





```

print(f"User said: {detected_text}")

# Hentikan program jika pengguna mengatakan "stop"
if detected_text.lower() == "stop":
    break

except sr.UnknownValueError:
    print("Google Speech Recognition could not understand audio.")
except sr.RequestError as e:
    print(f"Could not request results from Google Speech Recognition service; {e}")
except Exception as e:
    print(f"An error occurred: {e}")

except KeyboardInterrupt:
    print("Program terminated by user.")
except Exception as e:
    print(f"An unexpected error occurred: {e}")

if __name__ == "__main__":
    main()
    
```

Berikut adalah penjelasan singkat dari bagian-bagian utama dalam program :

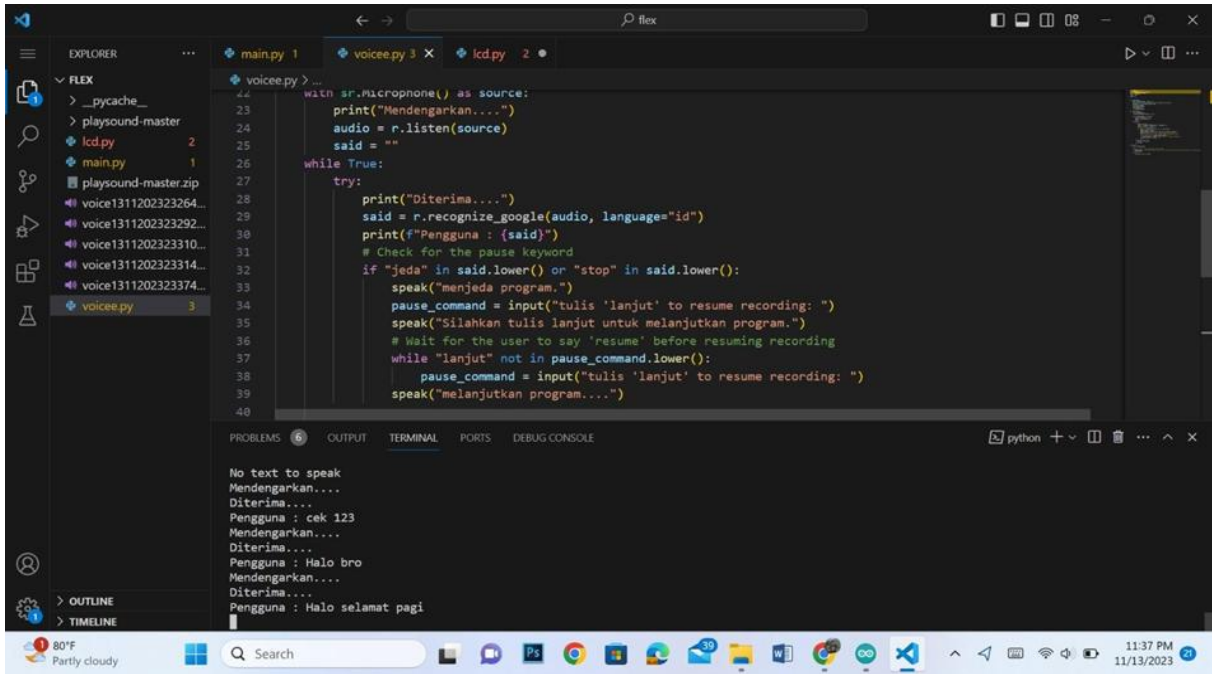
Tabel 4. Penjelasan Skrip *Speech-to-Text*

| No | Skrip Python                                    | Penjelasan  |
|----|---|---|
| 1  | Fungsi tts():                                   | Fungsi utama program yang mengatur proses mendengarkan suara dari mikrofon dan mengirimkan data teks yang terdeteksi ke <i>Firebase</i> . |
| 2  | Recognizer Setup: sr.Recognizer()               | digunakan untuk inialisasi objek recognizer untuk mendeteksi suara.   |
| 3  | while True                                      | Program berjalan dalam loop tak terbatas (while True) untuk terus mendengarkan input suara dari mikrofon.                                 |
| 4  | r.adjust_for_ambient_noise(source, duration=1): | Mendengarkan dan Mengenali Suara  |
| 5  | if said.lower() == "stop": break).              | Penghentian Program   |
| 6  | Main Program (if __name__ == "__main__":):      | Memanggil fungsi tts() untuk menjalankan program utama.   |

E. Pengujian Sistem *Speech-to-Text*

Pengujian sistem *Speech-to-Text* menggunakan Raspberry Pi bertujuan untuk memastikan bahwa teknologi ini dapat secara akurat *mengonversi* suara menjadi teks yang dapat diproses lebih lanjut. Proses pengujian ini dimulai dengan merekam berbagai sampel suara menggunakan microphone yang terhubung ke Raspberry Pi dalam berbagai kondisi lingkungan, termasuk yang memiliki *noise*. Data suara yang telah direkam kemudian menjalani proses preprocessing untuk mengurangi gangguan dan meningkatkan kualitas audio, sehingga memastikan hasil pengenalan suara yang lebih akurat. Berikut dibawah ini merupakan simulasi *Speech-to-Text*.





Gambar 4. Tampilan pemrograman *Voice Recognition* melalui VSCode

F. Implementasi *Software Speech-to-Text*

Implementasi *software* subsistem *Speech-to-Text* dirancang untuk mengubah input suara menjadi teks secara *real-time*, mendukung interaksi lebih lanjut dengan perangkat. Sistem ini menggunakan berbagai komponen dan teknik untuk memastikan keakuratan dan efisiensi dalam proses konversi suara ke teks.

G. Inisialisasi dan Perekaman Suara

Proses dimulai dengan inisialisasi modul-modul perangkat lunak yang diperlukan untuk perekaman suara. Menggunakan pustaka seperti PyAudio untuk menangkap input suara dari *microphone* yang terpasang pada perangkat. Inisialisasi ini mencakup konfigurasi tingkat sensitivitas *microphone* dan pengaturan parameter perekaman, seperti *sample rate* dan *bit depth*, untuk memastikan kualitas suara yang baik.

Tabel 5. Pengujian *Speech-to-Text* berdasarkan variasi volume suara

| No | Teks          | Variasi Volume | Keterangan     |
|----|---------------|----------------|----------------|
| 1  | Selamat Pagi  | 20%            | Tidak Dikenali |
|    |               | 50%            | Dikenali       |
|    |               | 75%            | Dikenali       |
|    |               | 100%           | Dikenali       |
| 2  | Selamat Malam | 20%            | Tidak Dikenali |
|    |               | 50%            | Dikenali       |





|   |               |      |                |
|---|---------------|------|----------------|
| 3 | Selamat Tidur | 75%  | Dikenali       |
|   |               | 100% | Dikenali       |
|   |               | 20%  | Tidak Dikenali |
|   |               | 50%  | Dikenali       |
|   |               | 75%  | Dikenali       |
| 4 | Makan Nasi    | 100% | Dikenali       |
|   |               | 20%  | Tidak Dikenali |
|   |               | 50%  | Dikenali       |
|   |               | 75%  | Dikenali       |
|   |               | 100% | Dikenali       |
| 5 | Darimana      | 20%  | Tidak Dikenali |
|   |               | 50%  | Dikenali       |
|   |               | 75%  | Dikenali       |
|   |               | 100% | Dikenali       |
|   |               | 100% | Dikenali       |

Pengujian ini menunjukkan bahwa sistem *Speech-to-Text* cukup andal dalam mengenali teks pada volume suara yang umum digunakan (50%, 75%, dan 100%). Namun, untuk memastikan kinerja yang optimal, perlu dipertimbangkan untuk meningkatkan kemampuan sistem dalam mengatasi volume suara yang rendah (20%). Ini bisa dilakukan dengan menyesuaikan parameter pengenalan suara atau menggunakan teknik pemrosesan sinyal yang lebih canggih untuk meningkatkan akurasi pada kondisi ekstrem tersebut.

#### H. Pengujian Variasi Kecepatan Berbicara

Pengujian variasi kecepatan berbicara bertujuan untuk mengevaluasi kemampuan sistem dalam mengenali teks saat pengguna berbicara dengan kecepatan yang berbeda-beda. Pengujian ini penting untuk memastikan bahwa sistem dapat bekerja secara efektif dalam berbagai kondisi berbicara yang mungkin ditemui dalam penggunaan sehari-hari.

Tabel 6. Pengujian *Speech-to-Text* berdasarkan variasi kecepatan suara

| No | Teks          | Variasi Kecepatan | Keterangan  |
|----|---------------|-------------------|---|
| 1  | Selamat Pagi  | Normal            | Dikenali dengan baik                              |
|    |               | Cepat             | Dikenali, namun ada beberapa kata yang terpotong  |
|    |               | Lambat            | Dikenali, dengan beberapa jeda yang lebih panjang |
| 2  | Selamat Malam | Normal            | Dikenali dengan baik                              |
|    |               | Cepat             | Dikenali, namun ada beberapa kata yang terpotong  |
|    |               | Lambat            | Dikenali, dengan beberapa jeda yang lebih panjang |
| 3  | Selamat Tidur | Normal            | Dikenali dengan baik                              |
|    |               | Cepat             | Dikenali, namun ada beberapa kata yang terpotong  |





|   |            |        |   |
|---|------------|--------|---|
|   |            | Lambat | Dikenali, dengan beberapa jeda yang lebih panjang |
| 4 | Makan Nasi | Normal | Dikenali dengan baik                              |
|   |            | Cepat  | Dikenali, namun ada beberapa kata yang terpotong  |
|   |            | Lambat | Dikenali, dengan beberapa jeda yang lebih panjang |
| 5 | Darimana   | Normal | Dikenali dengan baik                              |
|   |            | Cepat  | Dikenali, namun ada beberapa kata yang terpotong  |
|   |            | Lambat | Dikenali, dengan beberapa jeda yang lebih panjang |

Pengujian menunjukkan bahwa sistem *Speech-to-Text* pada sistem mampu mengenali teks dengan baik pada kecepatan berbicara normal. Namun, saat kecepatan berbicara ditingkatkan (cepat), terjadi kecenderungan beberapa kata terpotong atau teks menjadi kurang jelas. Sebaliknya, pada kecepatan berbicara yang lebih lambat, meskipun teks dikenali, terjadi penundaan yang lebih besar dalam pengenalan kata-kata. Hal ini menunjukkan bahwa penggunaan optimal sistem ini memperhatikan kecepatan berbicara agar hasil pengenalan teks tetap akurat dan jelas.

#### I. Evaluasi Kinerja Sistem *Speech-to-Text*

Salah satu masalah utama yang ditemukan dalam sistem *Speech-to-Text* adalah ketidakakuratan dalam mengenali teks saat pengguna berbicara dengan kecepatan yang cepat. Hal ini dapat mengakibatkan beberapa kata terpotong atau tidak dikenali dengan baik, yang dapat mengurangi efektivitas komunikasi melalui perangkat. Sistem *Speech-to-Text* sangat sensitif terhadap kebisingan lingkungan. Suara latar belakang seperti percakapan lain, suara mesin, atau suara angin dapat mengganggu kemampuan sistem untuk mengenali ucapan dengan benar. Ini menjadi masalah terutama ketika digunakan di lingkungan yang tidak terkendali atau bising.

Untuk mengatasi masalah ketidakakuratan pada berbicara cepat, parameter pengenalan suara seperti kecepatan pengambilan sampel dan pengaturan ambang batas energi dapat disesuaikan. Ini membantu sistem untuk lebih responsif dan akurat dalam mengenali ucapan yang cepat. Mengumpulkan dan menggunakan dataset pelatihan yang mencakup variasi aksen, dialek, dan kecepatan berbicara yang berbeda dapat membantu dalam meningkatkan ketahanan sistem terhadap variasi dalam ucapan. Pelatihan ulang model dengan data yang lebih beragam akan membuat sistem lebih robust dan akurat dalam berbagai kondisi.





## SIMPULAN

Pada dasarnya, implementasi Raspberry Pi dengan subsistem AI melalui *Speech-to-Text* bertujuan untuk memastikan bahwa teknologi ini dapat menjadi solusi untuk menjawab permasalahan yang dihadapi siswa Tunarungu di SLB B Tunas Harapan Karawang. Adapun tindak lanjutnya mengupayakan subsistem ini senantiasa secara akurat mengonversi suara menjadi teks yang dapat diproses lebih lanjut, sehingga dapat memudahkan tunarungu berkomunikasi dengan orang normal, begitupun sebaliknya.

## DAFTAR RUJUKAN

- Chen, Y., Li, Y., Liu, J., & Tan, C. W. (2020). Enhancing Speech-to-Text Technology for Deaf and Hard-of-Hearing People: A Field Study. \*IEEE Access, 8\*, 183918-183930. <https://doi.org/10.1109/ACCESS.2020.3029836>
- Chmielewski, D., De-la-Torre, M., & Perez-Pinar, M. (2019). Improving Communication for the Deaf and Hard of Hearing with Automatic Speech Recognition. \*Proceedings of the 2019 IEEE Global Humanitarian Technology Conference (GHTC)\*, 1-6. <https://doi.org/10.1109/GHTC46095.2019.9033013>
- Fellinger, J., Holzinger, D., & Pollard, R. (2012). Mental Health of Deaf People. \*The Lancet, 379\*(9820), 1037-1044. [https://doi.org/10.1016/S0140-6736\(11\)61143-4](https://doi.org/10.1016/S0140-6736(11)61143-4)
- Geiger, L., & Bergmann, P. (2018). A Real-Time Speech-to-Text System for Deaf and Hard of Hearing Students. \*Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems (CHI)\*, 1-6. <https://doi.org/10.1145/3173574.3173768>
- Harris, M., & Mohay, H. (1997). Learning to Look in the Right Place: A Comparison of Attentional Behavior in Deaf and Hearing Children. \*Journal of Deaf Studies and Deaf Education, 2\*(2), 95-103. <https://doi.org/10.1093/oxfordjournals.deafed.a014314>
- Juhola, M., & Pyykkö, I. (2005). Speech-to-Text Systems in Schools and the Quality of Their Transcriptions. \*Computer Speech & Language, 19\*(1), 41-60. <https://doi.org/10.1016/j.csl.2004.02.002>
- Karam, M., & Russo, F. (2009). Speech-to-Text Communication System for the Deaf and Hard of Hearing. \*Journal of Deaf Studies and Deaf Education, 14\*(3), 401-411. <https://doi.org/10.1093/deafed/enp006>
- Mehta, K., & Parikh, D. (2021). Artificial Intelligence in Education: Enhancing Communication for the Hearing Impaired. \*International Journal of Educational Technology in Higher Education, 18\*(1), 1-17. <https://doi.org/10.1186/s41239-021-00274-8>
- Morris, M. R., & Paepcke, A. (2006). Designing a Speech-to-Text Interface for the Deaf: A Study of User Preferences. \*Proceedings of the 2006 ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)\*, 239-246. <https://doi.org/10.1145/1168987.1169026>





- Rojas, D., & Capurro, M. (2015). Real-Time Speech-to-Text for Enhancing Communication in Classrooms for the Deaf. \*Proceedings of the 2015 International Conference on Universal Access in Human-Computer Interaction (UAHCI)\*, 129-137. [https://doi.org/10.1007/978-3-319-20684-4\\_13](https://doi.org/10.1007/978-3-319-20684-4_13)
- Scherer, M. J., & Federici, S. (2015). Why People Use and Don't Use Technologies: Introduction to the Special Issue on Assistive Technologies for Cognition/Cognitive Support Technologies. \*NeuroRehabilitation, 37\*(3), 315-319. <https://doi.org/10.3233/NRE-151264>
- Stinson, M. S., & Liu, Y. (1999). Deaf College Students' Perceptions of Communication in Mainstream Classes. \*Journal of Deaf Studies and Deaf Education, 4\*(4), 272-283. <https://doi.org/10.1093/deafed/4.4.272>
- Williams, S. J., & Drigas, A. S. (2020). Speech-to-Text Technology and the Development of Communication Skills in Deaf Students. \*International Journal of Emerging Technologies in Learning, 15\*(11), 197-212. <https://doi.org/10.3991/ijet.v15i11.13257>
- Zeng, F., & Galvin, J. J. (1999). Amplitude Mapping and Phoneme Recognition in Cochlear Implant Users. \*Hearing Research, 139\*(1-2), 56-64. [https://doi.org/10.1016/S0378-5955\(99\)00128-8](https://doi.org/10.1016/S0378-5955(99)00128-8)
- Zhao, Y., & Chai, C. S. (2019). Exploring the Effectiveness of Speech-to-Text Recognition in Assisting Deaf Students' Learning. \*Journal of Computer Assisted Learning, 35\*(1), 47-58. <https://doi.org/10.1111/jcal.12311>.

