

Aplikasi Penerjemah Kegiatan Seminar Menjadi Video Bahasa Isyarat BISINDO Dengan Speech To Text

Marcel Slamet Sugianto¹, Liliana², Anita Nathania P.³

Program Studi Informatika, Fakultas Teknologi Industri, Universitas Kristen Petra

Jl. Siwalankerto 121 – 131 Surabaya 60236

Telp. (031) – 2983455, Fax. (031) – 8417658

Emails: C14180033@john.petra.ac.id¹, lilian@petra.ac.id², anita.nathania@petra.ac.id³

ABSTRAK

Informasi disaat ini merupakan hal yang sangat dibutuhkan untuk menambah wawasan kita. Tetapi, penyampaian tersebut dapat terhalang beberapa kondisi seperti ketidakmampuan mendengar masyarakat tunarungu. Berdasarkan data dari Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia pada tahun 2019, 7.09% dari penduduk Indonesia adalah penderita tunarungu. Selain itu penyampaian informasi di seminar dapat terhalang oleh adanya noise suara dan peserta yang duduknya jauh dari pembicara akan kesulitan mendengar suara pembicara.

Pada penelitian ini akan menggunakan Speech-To-Text pada aplikasi Android yang bertujuan membantu menterjemahkan informasi dalam bentuk suara yang disampaikan seperti pada saat seminar kedalam bentuk teks dan akan diubah kedalam bentuk video bahasa isyarat BISINDO.

Hasil dari pengujian terhadap penggunaan fitur Speech-To-Text pada aplikasi yang telah dibuat menunjukkan mampu menampung kurang lebih 100 kata dalam 1 menit pada saat sekali pembicara bicara tanpa adanya jeda. Fitur Speech-To-Text yang digunakan membutuhkan waktu kurang lebih 2 detik untuk menterjemahkan suara yang diterima dan jeda waktu yang dibutuhkan perangkat pembicara ke perangkat peserta membutuhkan kurang lebih 3-5 detik setelah menggunakan 5 kecepatan internet yang berbeda-beda. Untuk akurasi dari fitur Speech-To-Text yang diuji menggunakan 3 narasi yang dibacakan oleh 4 orang yang berbeda menunjukkan akurasi dari fitur Speech-To-Text memiliki akurasi diatas 80% secara umum meskipun terdapat akurasi yang berada dibawah 80% yang disebabkan ketidakjelasan dari pelafalan kata.

Kata Kunci: Kotlin, Bahasa Isyarat, BISINDO, Mobile Application, Seminar, Speech-To-Text

ABSTRACT

Information at this time is very much needed to increase our knowledge. However, this delivery can be hindered by several conditions such as the inability to hear the deaf community. Based on data from the Data and Information Center of the Ministry of Health of the Republic of Indonesia in 2019, 7.09% of the Indonesian population is deaf. In addition, the delivery of information at the seminar can be hindered by noise and participants sitting far from the speaker will have difficulty hearing the speaker's voice.

In this study, we will use Speech-To-Text on the Android application which aims to help translate information in the form of

voice delivered as at a seminar into text and will be converted into BISINDO sign language video.

The results of testing the use of the Speech-To-Text feature in the application that has been made show that it is able to accommodate approximately 100 words in 1 minute at a time when the speaker speaks without any pause. The Speech-To-Text feature used takes approximately 2 seconds to translate the received voice and the time lag required by the speaker device to the participant's device takes approximately 3-5 seconds after using 5 different internet speeds. For the accuracy of the Speech-To-Text feature that was tested using 3 narrations read by 4 different people, the accuracy of the Speech-To-Text feature has an accuracy of above 80% in general, although there is an accuracy that is below 80% due to the ambiguity of the pronunciation.

Keywords: Kotlin, Hand Sign, BISINDO, Mobile Application, Seminar, Speech-To-Text.

1. PENDAHULUAN

Bahasa isyarat adalah sarana komunikasi yang digunakan oleh masyarakat tunawicara dan tunarungu dikarenakan bahasa isyarat ini dapat digunakan untuk mengidentifikasi diri atau memberikan informasi kepada orang lain. Salah satu bahasa isyarat yang ada di Indonesia adalah BISINDO. BISINDO sendiri merupakan jenis bahasa isyarat yang banyak dipakai dikarenakan kemudahannya untuk dipelajari.

Penyampaian informasi atau ilmu pengetahuan seperti melalui kegiatan seminar atau media online seperti Youtube menjadi hal yang umum saat ini. Tetapi, tidak semua kalangan dapat menerima informasi tersebut karena adanya halangan seperti ketidakmampuan mendengar masyarakat tunarungu. Menurut dari Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia pada tahun 2019, 7.09% dari penduduk Indonesia adalah penderita tunarungu. Oleh karena itu akan dibuatnya sebuah aplikasi yang akan menggunakan Speech-To-Text sebagai perantara dalam penyampaian suara oleh pembicara di suatu seminar kepada peserta tunarungu yang hadir saat itu dengan mengubah suara yang diterima kedalam bentuk teks dan diubah kedalam bentuk video bahasa isyarat BISINDO.

Beberapa aplikasi yang berhubungan dengan bahasa isyarat telah ada diantaranya adalah aplikasi Hear Me, I Sign, dan Hand Talk Translator.

Pada aplikasi Hear Me terdapat fitur speech to text yang mampu menerjemahkan sebuah teks menjadi dalam bentuk video bahasa isyarat BISINDO dalam kehidupan sehari-hari. Tetapi, pada aplikasi ini belum terdapat fitur yang membantu sebagai perantara

penyampaian materi yang diberikan saat seminar berlangsung kepada pengguna tunarungu.

Untuk aplikasi ASL Sign terdapat fitur speech-to-text dan text-to-speech yang bertujuan untuk membantu menjembatani komunikasi antara masyarakat biasa dengan masyarakat tunarungu yang memakai bahasa isyarat American Sign Language. Tetapi, sebagai penjematan komunikasi tersebut perlu secara bergantian dalam menggunakan perangkat untuk dapat menggunakan fitur tersebut. Untuk aplikasi yang terakhir yaitu I Sign hanya bekerja sebagai aplikasi kamus bahasa isyarat.

2. LANDASAN TEORI

2.1 Gesture

Gestur merupakan sebuah cara natural dalam komunikasi antar manusia, dikarenakan gerakan fisik seperti pergerakan tangan, lengan, atau tubuh mempunyai informasi yang berarti [3]. Gestur terbagi menjadi 3 bagian yaitu gestur lengan dan tangan, gestur kepala dan wajah, dan gestur tubuh. Dalam istilah gestur sendiri terbagi menjadi 2 hal yaitu, postur yang memiliki gerakan statik atau tidak bergerak sedangkan gestur yang melibatkan gerakan dinamis seperti melambatkan tangan saat mengucapkan selamat tinggal salah satunya [5], [6].

2.2 Bahasa Isyarat BISINDO

BISINDO merupakan bahasa isyarat yang dikembangkan oleh para tuna rungu di Indonesia dan digunakan sebagai bahasa isyarat dalam kehidupan sehari-hari mereka. Selain itu, BISINDO juga tidak memiliki aturan yang pasti sehingga mempermudah orang untuk mempelajari BISINDO lebih cepat. Dalam alfabet BISINDO terdapat beberapa alfabet yang dibentuk menggunakan 1 tangan saja seperti huruf C, E, I, J, L, O, R, U, V, dan Z sedangkan terdapat beberapa alfabet yang dibentuk menggunakan 2 tangan seperti huruf A, B, D, F, G, H, K, M, N, P, Q, S, T, W, X, dan Y [2]. Penggunaan BISINDO ini banyak digunakan dikarenakan mudah untuk dipelajari dan disesuaikan dengan bahasa keseharian mereka sehingga tiap daerah memiliki perbedaan [1], [10].

2.3 Speech-To-Text

Speech-To-Text merupakan salah satu bidang computer science yang bekerja dengan cara merubah inputan sinyal suara menjadi dalam bentuk teks [4], [7]. Speech-To-Text sendiri dikembangkan untuk membantu memberikan informasi kepada masyarakat disabilitas, memberikan perintah kepada perangkat lainnya, dan beberapa hal lainnya. Dalam proses Speech-To-Text terdapat beberapa hal yang perlu diperhatikan dikarenakan adanya utterance approach, utterance style, types of speaker model, vocabulary, dan channel variability [8], [9]. Hal-hal tersebut bersangkut pautan dengan hal seperti cara pelafalan kata, kosakata yang digunakan, dan berbagai hal eksternal seperti noise dan microphone salah satu contohnya.

Speech-To-Text yang digunakan pada penelitian ini adalah Speech-To-Text yang dibuat oleh Google dan terbagi menjadi 2 jenis yaitu, yang pertama adalah fitur Speech-To-Text yang terdapat pada Android yaitu dengan menggunakan objek speechRecognizer. Untuk yang kedua adalah Google Cloud Speech-To-Text API yang merupakan layanan berbayar yang disediakan oleh Google.

3. DESAIN SISTEM

Untuk mengatasi permasalahan dimana informasi dari pembicara yang tidak bisa tersampaikan dengan jelas kepada peserta acara maka, konsep dari aplikasi yang akan dibuat akan menggunakan

satu perangkat yang berfungsi sebagai penerjemah suara dari pembicara dan mengubahnya kedalam bentuk teks dengan menggunakan fitur Speech-To-Text. Setelah hasil terjemahan suara sudah dihasilkan maka nantinya akan dikirimkan ke seluruh peserta acara yang telah bergabung pada kegiatan melalui aplikasi yang telah dibuat. Hasil yang akan diterima oleh peserta acara nantinya akan ada 2 yaitu teks dan video bahasa isyarat BISINDO untuk membantu teman tuli. Untuk gambaran besar dari aplikasi yang akan dibuat dapat dilihat pada Gambar 1.



Gambar 1. Ilustrasi Sistem

Seperti yang terlihat pada Gambar 1, ilustrasi tersebut menjelaskan garis besar alur dari sistem yang akan dibuat. Alur dimulai dari perangkat pembicara menerima input suara dari pembicara acara dan akan memproses suara tersebut menjadi kedalam bentuk teks dan akan di-upload kedalam database. Setelah data berhasil di-upload kedalam database, perangkat dari para peserta tunarungu yang telah bergabung pada acara menggunakan sistem ini nantinya dapat menerima hasil terjemahan suara pembicara dan ditampilkan pada layar perangkat mereka dalam bentuk video bahasa isyarat dan teks.

Sistem yang akan dibuat akan berbasis Android dan dalam sistem tersebut akan terbagi menjadi 3 jenis user yaitu, pengguna, penyelenggara acara, dan pembicara acara.

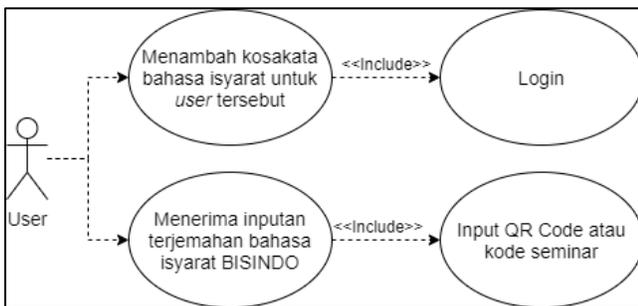
- **Pengguna**, merupakan peserta tunarungu yang hadir pada acara tersebut yang dimana peserta ini dapat menerima materi yang disampaikan oleh pembicara dalam bentuk video bahasa isyarat maupun teks.

- **Penyelenggara acara**, merupakan untuk panitia atau orang yang menyelenggarakan acara tersebut. Role ini ditujukan agar panitia dari acara tersebut dapat mendaftarkan acaranya pada aplikasi ini sehingga nantinya para peserta tunarungu dapat menggunakan fitur terjemahan materi seminar yang disampaikan pembicara.

- **Pembicara acara**, merupakan pembicara acara yang telah diundang oleh penyelenggara acara.

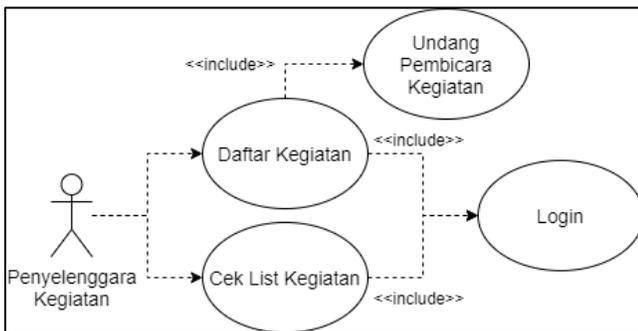
3.1 Use Case Diagram

Use case diagram ini bertujuan untuk menjelaskan gambaran dari berbagai fitur yang dapat digunakan oleh masing-masing pihak pada saat menggunakan sistem yang telah dibuat. Pada use case diagram ini akan terbagi menjadi 3 role yaitu peserta tunarungu acara yang hadir pada acara untuk dapat menerima terjemahan materi dari pembicara dalam bentuk video bahasa isyarat BISINDO, penyelenggara acara yang bertujuan mendaftarkan acaranya dan mengundang pembicara, dan pembicara acara yang akan menggunakan fitur Speech-To-Text.



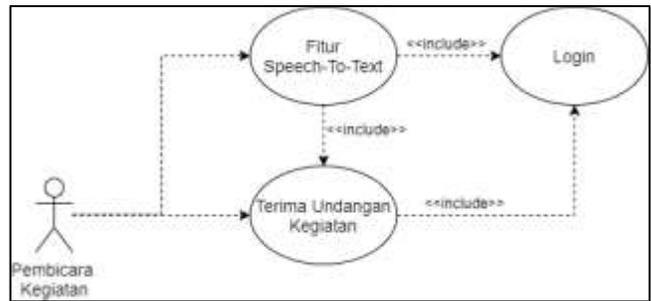
Gambar 2. Use Case Diagram Role Peserta

Pada Gambar 2, *user* mempunyai akses untuk melakukan penambahan kosakata bahasa isyarat dan menerima hasil terjemahan suara pembicara dalam bentuk teks dan video bahasa isyarat.



Gambar 3. Use Case Diagram Role Penyelenggara Kegiatan

Pada Gambar 3, penyelenggara kegiatan setelah melakukan *login* mempunyai akses untuk mengecek *list* kegiatan yang telah didaftar dan mendaftarkan kegiatan sekaligus mengundang pembicara kegiatan dengan cara memasukkan *email* pembicara.

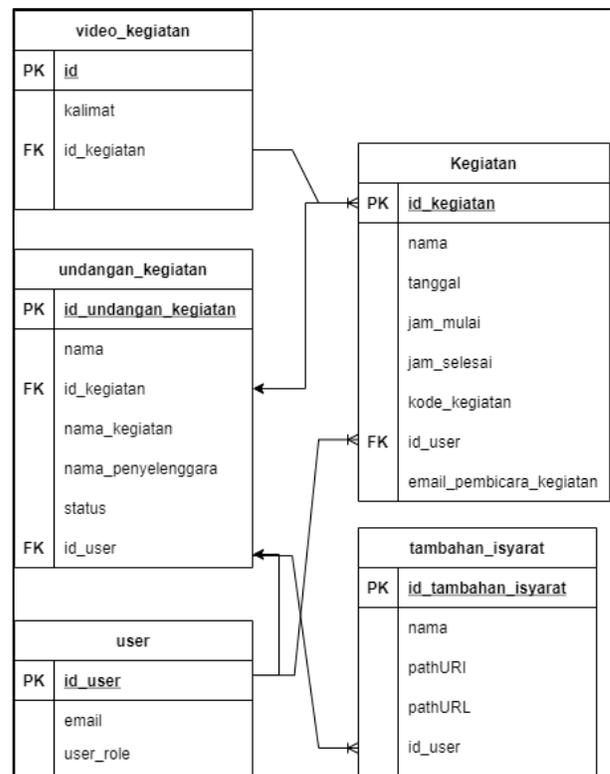


Gambar 4. Use Case Diagram Role Pembicara Kegiatan

Pada Gambar 4, pembicara kegiatan setelah melakukan *login* mempunyai akses untuk menggunakan fitur Speech-To-Text dan menerima undangan kegiatan dari penyelenggara kegiatan.

3.2 Entity Relationship Diagram

Entity Relation Diagram (ERD) ini bertujuan untuk menjelaskan struktur dari database yang akan digunakan oleh sistem seperti yang telah dilampirkan pada Gambar 5. Tabel yang akan digunakan terdiri dari 5 tabel yaitu, tabel user yang bertujuan menyimpan data user tunarungu, penyelenggara acara, dan pembicara acara. Tabel kegiatan bertujuan untuk menyimpan data kegiatan yang telah terdaftar dan email dari pembicara acara tersebut. Tabel video_kegiatan bertujuan menyimpan terjemahan suara dari pembicara menjadi kedalam bentuk teks pada acara tersebut. Tabel tambahan_isyarat yang bertujuan menyimpan kosakata yang telah ditambah oleh setiap user tunarungu beserta lokasi penyimpanannya. Tabel undangan_kegiatan yang bertujuan untuk menyimpan daftar undangan yang ditujukan untuk pembicara acara yang diberikan oleh penyelenggara acara

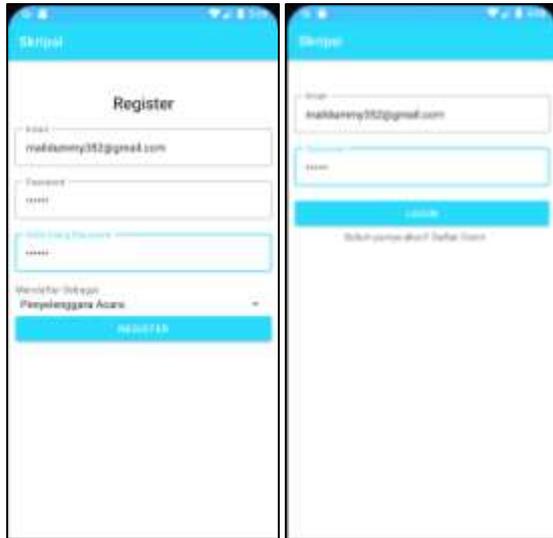


Gambar 5. Entity Relationship Diagram Sistem

4. PENGUJIAN

4.1 Pengujian Aplikasi

Disaat pengguna pertama kali membuka aplikasi setiap pihak perlu melakukan registrasi terlebih dahulu baru dapat melakukan login pada aplikasi seperti pada Gambar 6.



Gambar 6. Halaman Registrasi & Login

4.1.1 Penyelenggara Acara

Pada pihak penyelenggara acara setelah berhasil login pada aplikasi, maka pihak penyelenggara dapat mendaftarkan acaranya dengan menekan tombol *add* pada halaman utama dan akan mengisi data kegiatan dan email dari pembicara acara yang akan diundang. Setelah itu penyelenggara acara akan menerima kode kegiatan dan QR Code yang dapat dibagikan kepada peserta tunarungu yang hadir di acara tersebut seperti pada Gambar 7.



Gambar 7. Halaman registrasi kegiatan & detail kegiatan

4.1.2 Pembicara Acara

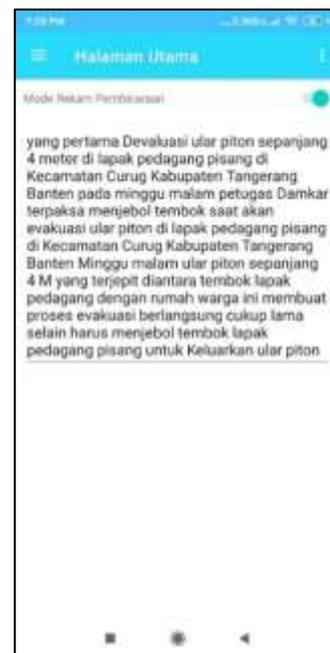
Pada pihak pembicara acara yang telah diundang setelah berhasil login maka dapat menerima undangan yang diberikan dari

penyelenggara pada aplikasi. Hal tersebut dilakukan dengan memilih menu navigasi halaman undangan kegiatan dan pembicara acara dapat menekan tombol untuk menerima undangan kegiatan tersebut seperti pada Gambar 8. Nantinya kegiatan tersebut akan terdaftar pada halaman utama dan pembicara dapat menggunakan fitur Speech-To-Text pada hari kegiatan berlangsung.



Gambar 8. Halaman list undangan kegiatan & halaman utama pembicara acara

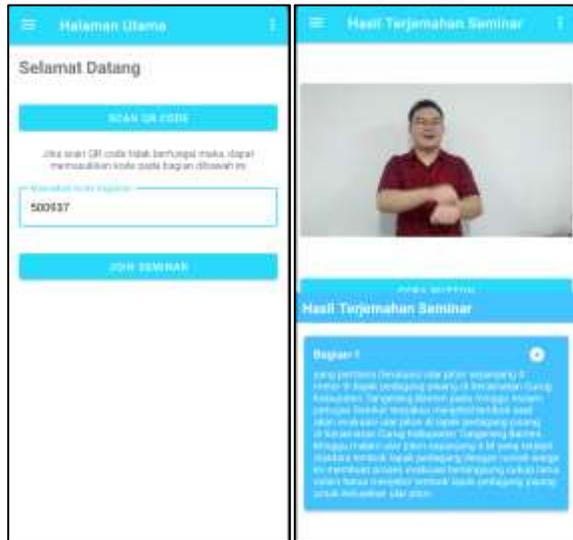
Pada hari acara berlangsung pembicara dapat menekan acara yang sedang berlangsung hari itu pada halaman utama dan dapat menggunakan fitur Speech-To-Text dengan cara menekan switch button pada halaman tersebut dan sesudah dinyalakan fitur Speech-To-Text sudah mulai dapat digunakan seperti pada Gambar 9.



Gambar 9. Halaman Speech-To-Text

4.1.3 Peserta

Pada pihak peserta yang hadir pada acara tersebut setelah berhasil login pada aplikasi dapat masuk ke halaman hasil terjemahan seminar setelah memasukkan kode kegiatan atau scan QR Code yang telah dibagikan oleh penyelenggara acara. Setelah masuk ke halaman tersebut nantinya peserta dapat menerima teks dan video bahasa isyarat dari teks tersebut apabila pembicara sudah mulai menyampaikan materi. Seperti pada Gambar 10.



Gambar 10. Halaman utama peserta memasukkan kode & halaman hasil terjemahan seminar

4.2 Durasi Jeda Antara Terjemahan Suara Pembicara Dikirimkan Ke Peserta

Pengujian ini dilakukan dengan tujuan untuk mengetahui lamanya waktu yang dibutuhkan materi dari pembicara dapat terkirimkan ke perangkat peserta tunarungu yang hadir pada acara dimana perangkat dari tiap peserta memiliki kecepatan internet yang berbeda-beda. Untuk spesifikasi CPU yang akan digunakan pada Tabel 1 akan diwakili dalam bentuk Abjad sebagai berikut:

- A: Octa-core 2.0 GHz Cortex-A53
- B: Octa-core 1.8 GHz Cortex-A53
- C: Octa-core (4x2.35 GHz Cortex-A53 & 4x1.8 GHz Cortex-A53)
- D: Google Play Intel Atom (x86)

Tabel 1. Spesifikasi Handphone Berdasarkan Chipset, CPU, RAM, Versi Android

Narasi Ke-	Durasi Speech-To-Text Berjalan	Banyak Kata yang Didapat	Jumlah Kata Benar	Akurasi (%)	Audio
1	1 Menit 2 Detik	143 Kata	129 Kata	90.21	Marcel
	59 Detik	135 Kata	125 Kata	91.91	Oliv
	58 Detik	146 Kata	131 Kata	89.73	Alek

	59 Detik	124 Kata	97 Kata	78.23	Gerry
	1 Menit	127 Kata	93 Kata	73.23	Gerry Modif
2	1 Menit 2 Detik	113 Kata	111 Kata	98.23	Marcel
	1 Menit	102 Kata	98 Kata	96.08	Oliv
	59 Detik	99 Kata	96 Kata	96.97	Alek
	1 Menit	120 Kata	83 Kata	69.16	Gerry
3	1 Menit	115 Kata	60 Kata	52.17	Gerry Modif
	59 Detik	119 Kata	114 Kata	95.8	Marcel
	58 Detik	131 Kata	126 Kata	96.18	Oliv
	59 Detik	122 Kata	118 Kata	96.72	Alek
	1 Menit	117 Kata	95 Kata	81.2	Gerry
	1 Menit	118 Kata	92 Kata	77.97	Gerry Modif

Dengan 5 kecepatan internet yang memiliki masing-masing download speed dan upload speed sesuai pada Tabel 1, dilakukan uji coba dengan cara mengukur waktu yang dibutuhkan fitur Speech-To-Text untuk menterjemahkan kata pertama yang diterima dengan menggunakan 5 kecepatan internet yang berbeda. Dari pengujian tersebut didapat data seperti pada Tabel 2. Untuk mempermudah penjelasan pada Tabel 2 dan Tabel 3, 5 kecepatan internet tersebut akan diwakili menggunakan abjad sebagai berikut:

- A: Up 7.67 Mbps / Down 1.15 Mbps
- B: Up 2.14 Mbps / Down 3.42 Mbps
- C: Up 1.32 Mbps / Down 2.6 Mbps
- D: Up 0.16 Mbps / Down 6.72 Mbps
- E: Up 0.30 Mbps / Down 14.8 Mbps

Tabel 2. Durasi Speech-To-Text menterjemahkan kata pertama yang diterima

	Kecepatan Internet Pembicara				
	A	B	C	D	E
Durasi Speech-To-Text Menterjemahkan Kata Pertama yang Diterima (Detik)	2.233	2.25	2.355	2.076	2.176

Dari hasil Tabel 2 waktu yang dibutuhkan fitur Speech-To-Text dalam menterjemahkan kata pertama yang diterima dari suara pembicara adalah kurang lebih 2 detik dengan menggunakan 5 kecepatan internet yang berbeda-beda.

Tabel 3. Jeda antara terjemahan suara pembicara dikirimkan ke peserta

Kecepatan Internet Pembicara	Kecepatan Internet Peserta				
	A	B	C	D	E
A		6.39 Detik	4 Detik	4.45 Detik	5.73 Detik
B	3.14 Detik		4.83 Detik	4.16 Detik	5.33 Detik
C	3.13 Detik	5.11 Detik		3.88 Detik	4.39 Detik
D	1 Detik	3.16 Detik	1.88 Detik		3 Detik
E	1.36 Detik	3.44 Detik	2.98 Detik	2.30 Detik	

Dari hasil Tabel 3, pembicara dengan kecepatan internet upload speed 7.67 Mbps dan download speed 1.15 Mbps memiliki jeda waktu pengiriman yang paling lama dimana waktu yang dibutuhkan paling cepat untuk dapat menerima data berkisar pada 4 detik dan jeda waktu paling lama didapat 5.77 detik sedangkan untuk kecepatan internet dengan upload speed 0.16 Mbps dan download speed 6.72 Mbps memiliki waktu jeda yang cukup singkat yaitu 1 detik apabila dan waktu paling lamanya adalah 3.16 detik. Dari Tabel 3 ini dapat disimpulkan dengan adanya perbedaan internet ini baik dari pembicara dan peserta acara mempengaruhi jeda waktu yang dibutuhkan peserta acara agar dapat menerima hasil terjemahan suara pembicara yang dapat disebabkan oleh upload speed dan download speed dari kecepatan internet yang dimiliki.

4.3 Pengujian Fitur Speech-To-Text Dengan Beberapa Jenis Narasi

Tujuan pengujian ini adalah untuk mengetahui jumlah inputan suara yang dapat diterima fitur Speech-To-Text dan mengecek tingkat keakuratan terjemahan suara kedalam bentuk teks dengan menggunakan input suara dari orang yang berbeda-beda. Pengujian ini akan menggunakan 3 narasi yang berbeda dan akan diuji dengan cara diucapkan oleh audio dari 4 orang yang berbeda-beda. Untuk spesifikasi perangkat yang digunakan untuk merekam audio tersebut dapat dilihat pada Tabel 4.

Tabel 4. Spesifikasi handphone pembaca yang digunakan untuk merekam audio

Merk HP	Perangkat yang Digunakan Untuk Merekam	Audio
Xiaomi Redmi 5 Plus	Mic Built-in <i>handphone</i>	Marcel
OPPO A54	Mic Built-in <i>handphone</i>	Oliv
Xiaomi Mi 10t Pro	Mic Built-in <i>handphone</i>	Alek
Poco x3 gt	Mic Built-in <i>handphone</i>	Gerry

Tabel 5. Hasil pengujian fitur Speech-To-Text menggunakan beberapa narasi

Narasi Ke-	Durasi Speech-To-Text Berjalan	Banyak Kata yang Didapat	Jumlah Kata Benar	Akurasi (%)	Audio
1	1 Menit 2 Detik	143 Kata	129 Kata	90.21	Marcel
	59 Detik	135 Kata	125 Kata	91.91	Oliv
	58 Detik	146 Kata	131 Kata	89.73	Alek
	59 Detik	124 Kata	97 Kata	78.23	Gerry
	1 Menit	127 Kata	93 Kata	73.23	Gerry Modif
2	1 Menit 2 Detik	113 Kata	111 Kata	98.23	Marcel
	1 Menit	102 Kata	98 Kata	96.08	Oliv
	59 Detik	99 Kata	96 Kata	96.97	Alek
	1 Menit	120 Kata	83 Kata	69.16	Gerry
3	1 Menit	115 Kata	60 Kata	52.17	Gerry Modif
	59 Detik	119 Kata	114 Kata	95.8	Marcel
	58 Detik	131 Kata	126 Kata	96.18	Oliv
	59 Detik	122 Kata	118 Kata	96.72	Alek
	1 Menit	117 Kata	95 Kata	81.2	Gerry
	1 Menit	118 Kata	92 Kata	77.97	Gerry Modif

Dari hasil Tabel 5, Hasil akurasi yang didapat oleh servis Speech-To-Text Google pada Android memiliki nilai akurasi yang cukup bagus secara keseluruhan dimana memiliki nilai akurasi diatas 80% secara keseluruhan. Tetapi, pada narasi ke-2 yang menggunakan audio dari saudara Gerry memiliki nilai akurasi yang rendah yaitu 69.16%. Oleh karena itu dilakukan modifikasi pada audio Gerry yaitu dengan membesarkan volume suara dari audio rekaman Gerry sehingga ketika dilakukan pengujian kembali didapati hasil seperti pada Tabel 5 dengan sumber audio dari "Gerry Modif". Hasil dari pengulangan uji coba menunjukkan hasil akurasi yang dihasilkan tidak berbeda jauh dan bahkan lebih rendah dari pengujian sebelumnya sehingga dari sini dapat diketahui volume suara meskipun dibesarkan tidak berpengaruh terhadap hasil terjemahan dari Speech-To-Text yang digunakan sehingga hal yang memungkinkan berpengaruh pada hasil terjemahan Speech-To-Text adalah kejelasan dalam pengucapan kata dari pembicara. Oleh karena itu dapat disimpulkan, hasil terjemahan suara dari Speech-To-Text tidak begitu dipengaruhi oleh volume suara dari melainkan ketepatan dan kejelasan pengucapan kata yang diberikan

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan dari hasil penelitian dan pengujian yang telah dilakukan telah didapatkan beberapa kesimpulan sebagai berikut:

1. Kegunaan dari aplikasi yang telah dibuat saat ini belum dapat diimplementasikan pada kegiatan seminar sesungguhnya. Hal ini dikarenakan hasil terjemahan bahasa isyarat terhadap suara yang diterima memiliki waktu delay sekitar 2 detik dan karena terdapatnya perbedaan kecepatan internet yang dimiliki setiap perangkat peserta yang hadir dapat menyebabkan setiap perangkat peserta memiliki durasi yang berbeda-beda dalam menerima hasil terjemahan suara pembicara sehingga informasi yang sedang dibicarakan oleh pembicara dan informasi yang sedang diterima oleh peserta tidak bisa sama.

2. Dari hasil pengujian, ketepatan dan kejelasan volume suara dalam pelafalan suatu kata sangat berpengaruh dalam hasil terjemahan suara menjadi kedalam bentuk teks yang dilakukan oleh speech-to-text

3. Servis dari Speech-To-Text yang digunakan hanya dapat menerima input suara paling lama selama 1 menit kurang lebihnya dan setelah itu servis Speech-To-Text akan berhenti berjalan dan perlu dijalankan kembali agar dapat menerima input suara. Oleh karena itu jumlah kata yang dapat ditampung oleh fitur Speech-To-Text dalam 1 menitnya kurang lebih adalah 100 kata sehingga jumlah kata yang bisa ditampung fitur Speech-To-Text bergantung dengan kecepatan pembicara berbicara.

5.2 Saran

Saran-saran yang dapat diberikan dari hasil penelitian ini dengan harapan untuk mengembangkan penelitian kedepannya diantara lainnya adalah mengembangkan fitur STT yang dapat memproses suara secara real-time sehingga dapat memberikan hasil terjemahan suara disaat yang bersamaan dengan pembicara berbicara, melibatkan komunitas orang yang dapat berbahasa isyarat untuk dapat mengembangkan hasil dari aplikasi yang akan dibuat agar dapat lebih baik, dan untuk mengoptimalkan hasil dari Speech-To-Text pada perangkat pembicara yang digunakan sebagai penerima input suara dapat menggunakan perangkat tambahan yang dapat membantu perangkat dapat menerima suara pembicara lebih dekat.

6. Daftar Pustaka

[1] Gumelar, G., Hafiar, H., & Subekti, P. (2018). KONSTRUKSI MAKNA BISINDO SEBAGAI BUDAYA TULI BAGI

ANGGOTA GERKATIN. *Informasi*, 48(1), 65-78. DOI: 10.21831/informasi.v48i1.17727

- [2] Khan, R. Z., & Ibraheem, N. A. (2012). Survey on Gesture Recognition for Hand Image Postures. *Computer and Information Science*, 5(3). DOI: 10.5539/cis.v5n3p110
- [3] Khilari, P., & Bhope, V. P. (2015). A REVIEW ON SPEECH TO TEXT CONVERSION METHODS. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 4(7). ISSN: 2278 –1323
- [4] Kothadiya, D., Pise, N., & Bedekar, M. (2020). Different Methods Review for Speech to Text and Text to Speech Conversion. *International Journal of Computer Applications*, 975, 8887. DOI: 10.1007/s11042-020-10073-7
- [5] Mitra, S., & Acharya, T. (2007). Gesture Recognition: A Survey. *IEEE Transactions on Systems, Man and Cybernetics, Part c (Applications and Reviews)*, 37(3), 311–324. DOI: 10.1109/TSMCC.2007.893280
- [6] Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 677–695. DOI: 10.1109/34.598226
- [7] Singh, A. P., Nath, R., & Kumar, S. (2018, November). A survey: Speech recognition approaches and techniques. In 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (pp. 1-4). IEEE. DOI: 10.1109/UPCON.2018.8596954
- [8] Trivedi, A., Pant, N., Shah, P., Sonik, S., & Agrawal, S. (2018). Speech to text and text to speech recognition systems-Areview. *IOSR J. Comput. Eng.*, 20(2), 36-43. DOI: 10.9790/0661-2002013643
- [9] Vadwala, A. Y., Suthar, K. A., Karmakar, Y. A., & Thakkar, N. (2017). Survey paper on Different Speech Recognition Algorithm: Challenges and Techniques SAMVAD-Speak Ask Manage Via Android Device View project Survey paper on Different Speech Recognition Algorithm: Challenges and Techniques. *International Journal of Computer Applications*, 175(1). DOI: 10.1109/TSMCC.2007.893280
- [10] Zulpicha, E. (2017). Konflik Kebijakan Penggunaan Sistem Bahasa Isyarat Indonesia di Lingkungan Pendidikan Formal. *Jurnal Analisa Sosiologi*, 6(1), 100–109. DOI: 10.20961/jas.v6i1.18190