

SISTEM PENGENALAN UCAPAN MANUSIA DENGAN MENGGUNAKAN GOOGLE API PADA SISTEM PAKAR KONSELING TRAUMA PADA ANAK DI MASA PANDEMI

Ni Kadek Eva Anggreni¹⁾, Fajar Astuti Hermawati²⁾, IGAA Noviekayati³⁾

Teknik Informatika, Universitas 17 Agustus 1945^{1,2}

E-mail : kadekeva785@gmail.com¹⁾, fajarastuti@untag-sby.ac.id²⁾

ABSTRAK

Pada era modern saat ini perkembangan teknologi semakin canggih dan semakin pesat, seperti pada teknologi informasi dan komunikasi. Kemajuan teknologi saat ini memberi dampak yang luar biasa serta menghasilkan macam-macam teknologi yang dapat menimbulkan sebuah interaksi antara manusia dan komputer. Salah satu kemajuan teknologi saat ini ialah sistem pengenalan ucapan jika diterjemahkan ke dalam bahasa inggris disebut sebagai *Speech Recognition*. Tujuan utama penelitian untuk mendapatkan serta mengidentifikasi suara anak agar bisa mengetahui masalah yang sedang dialami dengan menggunakan suara. Pada penelitian ini memanfaatkan *Speech Recognition* atau pengenalan ucapan untuk membantu anak-anak dalam mengatasi dan mendeteksi dari stress dan trauma yang dialaminya akibat pandemi *covid 19* aplikasi ini dikembangkan menggunakan sebuah teknologi berbasis *Desktop* dengan memanfaatkan *Google Speech API*. Penelitian ini menghasilkan akurasi dengan melakukan 30 kali percobaan yang dibagi 15 anak keadaan sepi dan 15 anak keadaan ramai dengan menggunakan 3 indikator yaitu dengan jarak, suasana sekitar sepi dan juga ramai maka suara yang terdeteksi dalam keadaan sepi dengan keadaan jarak 1 meter hasilnya sebesar 93% sedangkan keadaan ramai dengan jarak 1 meter hasilnya 86% dan pada saat percobaan *speech recognition* menggunakan *chatbot* aplikasi dicoba saat keadaan sepi hasil penangkapan suaranya lebih baik daripada saat keadaan ramai.

Kata kunci : Aplikasi, *Desktop*, *Speech Recognition*, *Google Speech API*, *Chatbot*.

ABSTRACT

In the current modern era, technological developments are increasingly sophisticated and increasingly rapid, such as in information and communication technology. Advances in technology today have a tremendous impact and produce various technologies that can lead to an interaction between humans and computers. One of the current technological advances is a speech recognition system when translated into English is called Speech Recognition. The main purpose of the study was to obtain and identify children's voices so that they could find out the problems they were experiencing using sound. In this study, using Speech Recognition or speech recognition to assist children in overcoming and detecting the stress and trauma they experience due to the covid 19 pandemic, this application was developed using a Desktop-based technology by utilizing the Google Speech API. The experiment was divided into 15 quiet children and 15 busy children using 3 indicators, namely by distance, the atmosphere was quiet and also crowded, the sound detected in a quiet state with distance 1 meter of results was 93% while crowded conditions with a distance of 1 meter resulted in 86 % and during the speech recognition experiment using a chatbot, the application was tried when it was quiet, the results of the sound capture were better than when it was crowded.

Keyword : Application, *Desktop*, *Speech Recognition*, *Google Speech API*, *Chatbot*.

1. Pendahuluan

Pada era modern saat ini perkembangan teknologi semakin canggih dan semakin pesat, seperti pada teknologi informasi dan komunikasi. Kemajuan teknologi saat ini memberi dampak yang luar biasa serta menghasilkan macam-macam teknologi yang dapat menimbulkan sebuah interaksi antara manusia dan komputer. Salah satunya kemajuan teknologi melalui pengenalan ucapan atau bisa disebut sebagai *speech recognition*. Memanfaatkan sebuah fitur-fitur seperti ucapan ke kata bisa disebut dengan *speech-to-text*, lalu seperti kata ke ucapan *text-to-speech*. (Wisdhani, Ontowirjo and Kambey, 2021)

Adanya perkembangan teknologi saat ini seperti pada perkembangan pengguna computer yang paling banyak digunakan di hampir seluruh dunia, banyaknya pengguna computer ini adalah karena mudah digunakan. Contoh fiturnya ialah pengenalan ucapan (*speech recognition*) yang dimana berisi *speech-to-text* dan *text-to-speech* yang bermanfaat untuk memudahkan seseorang dalam memasukkan tanpa menggunakan keyboard ataupun untuk mengoreksi sebuah ucapan atau untuk mencari bagaimana bacaan tersebut diucapkan dengan benar. (Muhammad, Hidayat and Amrullah, 2019)

Pengenalan ucapan (*speech recognition*) dengan menggunakan *API Google Speech* yang dapat digunakan sebagai mencocokkan suara atau pemeriksaan bahasa seseorang dengan menggunakan kata kunci. *Speech recognition* sudah banyak digunakan oleh banyak orang untuk memudahkan seseorang jika tidak ingin menulis teks pada aplikasi. (Akbar *et al.*, 2021) *Google cloud speech API* ialah sebuah aplikasi pengembang mengubah audio menjadi teks atau teks menjadi audio dan merupakan sebuah layanan dari *Machine Learning Model* adanya *API* ini *programmer* atau *developer* bisa dengan mudah mengembangkan serta membuat aplikasi yang dapat melihat, mendengar, dan menerjemahkan (Intan *et al.*, 2017)

Penelitian tentang *speech recognition* yang telah dilakukan sebelumnya dalam penelitiannya berjudul *asmaul husna learning application using the speech recognition* yang dilakukan oleh (Wisdhani, Ontowirjo and Kambey, 2021) pada penelitian ini menggunakan metode *speech recognition* dengan memanfaatkan fitur *speech to text* dan menggunakan algoritma *Hidden Markov Model (HMM)*. Objek penelitian yang digunakan oleh penulis adalah gambar atau foto serta audio dari 99 ayat asmaul husna dan suara user, hasil dari penelitian ini menghasilkan 85,1% dari percobaan 10 kali pengujian. Penelitian berikutnya yaitu *abjad toward interactive learning approach to Arabic reading based on speech recognition* (Alsunaidi *et al.*, 2018). Penelitian ini menggunakan *speech recognition* dengan fitur *speech to text* untuk mencocokkan suara pengguna ke dalam tulisan bahasa arab. Objek penelitian ini menggunakan kalimat-kalimat bahasa arab dan suara dari user, hasil dari penelitian ini dari 25 orang hasil nilai akurasi sebanyak 50,5% untuk orang dewasa untuk anak-anak akurasi sebanyak 35,56%.

Pada penelitian ini pengenalan ucapan digunakan untuk membantu masyarakat di masa pandemi *COVID-19* karena saat ini banyak sekali masyarakat yang mengalami stress akibat lockdown dan juga trauma karena disebabkan anggota keluarga yang telah meninggal dunia akibat terkena penyakit *COVID-19*. Maka dari itu banyak sekali masyarakat yang takut untuk pergi keluar rumah, karena takut tertular penyakit covid. Akibatnya kondisi psikologis mereka menjadi terganggu karena adanya ketakutan dan kecemasan berlebihan tersebut sehingga tidak bisa mengetahui apa yang sebenarnya yang terjadi pada kondisi psikis seseorang tersebut.

Berdasarkan masalah tersebut penggunaan teknologi *Speech Recognition* dapat dijadikan sebuah solusi alternatif guna membantu cek kondisi dini yang dirasakan. Karena aplikasi ini dibangun berbasis desktop dengan memanfaatkan media suara serta menggunakan sebuah teknologi *API Google Speech To Text* dan juga *API Google Text To Speech* yang dapat merekam sebuah suara menjadi teks dan juga teks menjadi suara guna membantu seseorang dalam melakukan konseling. Inputan suara tersebut nantinya akan

dianalisa dan mengeluarkan sesuai dengan keluhan yang diucapkan atau keluhan yang dituliskan.

Diharapkan dengan adanya penelitian ini dapat membantu masyarakat untuk melakukan kegiatan konsultasi tentang psikis yang dirasakan dan diharapkan aplikasi ini dapat menghasilkan sebuah analisa yang benar serta akurat.

2. Metode

Pada penelitian ini dibagi beberapa tahapan yaitu persiapan, implementasi dan evaluasi. Tahapan pertama adalah melakukan pengumpulan data dengan menggunakan studi literatur untuk menganalisa serta mengeksplor teori dan fakta yang mendasari penelitian serta melakukan analisa sistem untuk memenuhi kebutuhan yang akan diperlukan seperti bahan dan perangkat yang digunakan, lalu melakukan desain mockup untuk memberikan gambaran dan mempermudah dalam merancang sebuah sistem.

Pada penelitian ini tahapan pertama yaitu melakukan penentuan obyek penelitian dan pengumpulan data, obyek penelitian ini merupakan dataset berupa suara anak-anak yang terkena trauma akibat pandemic covid 19. Data tersebut didapat dari anak-anak dengan rentang umur 10 sampai 12 tahun. Pada tahapan kedua dilakukan studi literatur yang bertujuan untuk melakukan sebuah kajian dari penelitian terdahulu yang pernah dilakukan sebelumnya, khususnya dalam melakukan penelitian sistem pengenalan ucapan menggunakan *google speech api*. *Google Cloud Speech Api* merupakan sebuah aplikasi pengembang mengubah audio menjadi teks atau teks menjadi audio dan merupakan sebuah layanan dari *Machine Learning Model* yang dimana didalamnya memiliki penggabungan antara *Google Translate API* dan *Cloud Vision API*. Dengan adanya *API* ini *programmer* atau *developer* bisa dengan mudah mengembangkan serta membuat aplikasi yang dapat melihat, mendengar, dan menerjemahkan (Intan *et al.*, 2017)

Tahapan selanjutnya pada penelitian ini melakukan tahapan analisa sistem kebutuhan sistem yang terkait dengan objek penelitian berupa data. Dataset yang akan digunakan adalah suara anak-anak pada rentang usia 10 sampai 12 tahun yang sedang mengalami trauma atau PTSD (*Post Traumatic Stress Disorder*). Berikut ini merupakan tahapan untuk mendapatkan dataset yang dapat dilihat pada gambar flowchart 1 dibawah ini:



Gambar 1 Flowchart mengumpulkan data suara

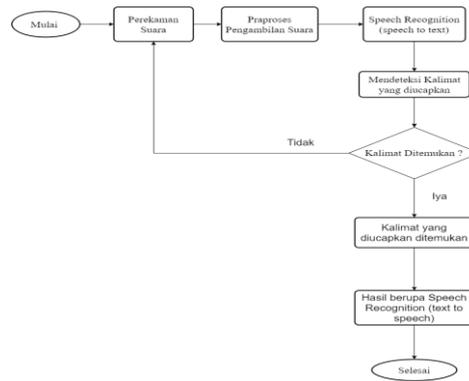
Gambar 1 diatas menunjukkan sebuah alur saat melakukan pengambilan dataset suara menggunakan sebuah mikrofon ponsel yang diambil secara langsung.

Berikut ini merupakan penjelasan dari flowchat mengumpulkan data suara :

1. Tahapan pertama dan kedua melakukan rekaman suara melalui ponsel hp. Perekaman suara ini dibagi menjadi 2 kondisi yaitu sepi dan ramai dengan jarak masing-masing 1 meter.
2. Tahapan selanjutnya setelah melakukan perekaman suara. Suara tersebut disimpan di dalam memori ponsel.

Analisa sistem selanjutnya yaitu menunjukkan sebuah alur dari sebuah aplikasi pengenalan ucapan yang terdapat pada gambar 2 dibawah ini :

SISTEM PENGENALAN UCAPAN MANUSIA DENGAN MENGGUNAKAN...

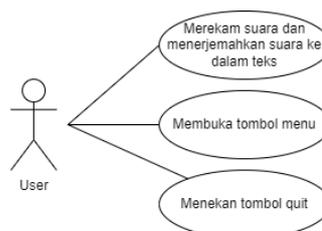


Gambar 2 Alur sistem pengenalan ucapan

Gambar 2 menunjukkan sebuah alur atau flowchart alur aplikasi pengenalan yang akan dijelaskan sebagai berikut:

1. Langkah pertama melakukan perekaman suara dari anak.
2. Melakukan pra proses pengambilan pada tahap ini dilakukan pengecekan pada suara tersebut dapat terdeteksi dengan jelas atau tidak.
3. Langkah ketiga, setelah dilakukan pra proses pengambilan suara tahapan selanjutnya adalah melakukan *speech recognition* yang dimana melakukan perubahan *speech to text*. *Speech recognition* ialah sebuah inovasi tentang pengenalan ucapan dengan memanfaatkan gelombang atau sinyal suara manusia yang diinputkan dan diproses melalui computer. *Speech-to-text* merupakan sebuah fitur dari *google cloud api* yang mengkonversi teknik ke teks dengan cara mengambil input dalam bentuk suara lalu di *convert* ke dalam bentuk tulisan. Keunggulan dari *speech to text* ialah untuk pengurangan kesalahan ejaan yang akan terjadi jika seseorang mengetik menggunakan sebuah *keyboard*. Tujuan dari *speech recognition* ialah untuk mengkonversi sebuah sinyal suara menjadi ke dalam bentuk teks yang sesuai dengan yang diucapkan dan tidak bergantung dengan perangkat yang digunakan dalam merekam suara (Rabiner, 2017; Akbar *et al.*, 2019; Buana, 2020)
4. Langkah keempat setelah kalimat yang diucapkan ditemukan dan sesuai maka system akan memberikan hasilnya berupa *text to speech*. *text-to-speech* adalah sebuah fitur dari *google cloud api* yang bisa menangkap serta memproduksi sinyal ucapan secara otomatis ter transkrip grafem to form untuk sebuah kalimat diucapkan. *Text- to speech* memiliki sebuah tujuan adalah untuk membaca sebuah teks apapun yang dapat dikenali oleh komputer dan mengirimkannya ke sebuah sistem berupa *optical character recognition (OCR)*, sistem harus dapat memahaminya.(Dutoit, 1997; Setiawan, 2016)

Analisa sistem selanjutnya yaitu menunjukkan sebuah *usecase* dari sebuah aplikasi pengenalan ucapan yang terdapat pada gambar 2.3 dibawah ini :



Gambar 2 *Usecase* user pada sistem pengenalan ucapan

Gambar 3 merupakan gambaran dari *usecase* user saat menggunakan aplikasi sistem pengenalan ucapan manusia yang dimana user dapat melakukan perekaman suara yang

diterjemahkan suara ke dalam teks lalu membuka tombol menu dan menekan tombol quit untuk keluar aplikasi.

Berikutnya tahapan penelitian perancangan sistem, dalam tahap ini akan melakukan sebuah perancangan sistem yaitu berupa perancangan *mockup*. Tahap pembuatan sebuah antarmuka pengguna ialah sebuah tahap yang penting dan harus dilakukan untuk memudahkan dalam merancang sebuah sistem pengenalan ucapan manusia, desain *mockup* sistem akan ditampilkan seperti pada gambar dibawah ini.



Gambar 4 Tampilan mockup pengenalan ucapan

Gambar 4 merupakan tampilan mockup pengenalan ucapan berbasis *chatbot*. User dapat melakukan rekaman suara dan nantinya sistem akan merubah suara ke dalam teks dan bot akan mengirimkan teks dan diubah ke suara dan ada kolom pada masukan teks disini untuk mode chat. *Chatbot* merupakan salah satu sistem cerdas yang dihasilkan dari pemrosesan bahasa alami atau *Natural Language Processing* (NLP) yang merupakan salah satu cabang dari kecerdasan buatan atau *Artificial Intelligence*. (Suryani and Amalia, 2017)

3. Hasil dan Pembahasan

Pada subbab ini akan membahas tentang hasil dari penelitian yang sudah dilakukan yaitu pengumpulan data, pengujian *speech to text* dan *pengujian speech to text* pada *chatbot*

3.1 Pengumpulan Data

Datashet didapat dari Sekolah Dasar Negeri Menur Pumpungan IV yang dimana dataset tersebut berupa sebuah rekaman suara anak-anak yang mengalami trauma akibat pandemic covid 19. Suara yang diambil pada penelitian ini adalah anak-anak berumur 10 sampai 12 tahun sebanyak 30 orang.

3.2 Pengujian *speech to text*

Pada subbab ini peneliti melakukan sebuah uji coba *speech to text* dengan menggunakan 3 indikator yaitu jarak saat pengambilan suara diambil dengan jarak 1 meter dan pengambilan suara dengan suasana sekitar keadaan sepi dan ramai. Pada uji coba ini dilakukan bertujuan untuk melihat kecocokan kata yang diucapkan dan hasil yang terbaca di komputer. Pada pengujian ini menggunakan 30 anak dengan rentang umur 10 sampai 12 tahun yaitu 15 anak dalam keadaan sepi dan 15 anak dalam keadaan ramai dengan jarak masing-masing 1 meter.

Hasil dari percobaan ini mendapatkan nilai *confidence* dalam keadaan sepi hasilnya paling tinggi adalah 0.96728677 dan dalam keadaan ramai hasil *confidencennya* paling rendah adalah 0.55728659 terdapat audio yang tidak ditemukan dikarenakan pada saat merekam suara kondisi di sekitar sangat ramai.

Nilai *confidence* merupakan representasi dari tingkat keyakinan akurasi antara data sampel dengan karakteristik pembanding yang sama. (Bengio *et al.*, 2002) Di dalam program dibawah ini nilai *confidence* didapatkan dari *recognize_google*.

SISTEM PENGENALAN UCAPAN MANUSIA DENGAN MENGGUNAKAN...

Karena `recognize_google()` merupakan sebuah api dari google yang berfungsi untuk menerjemahkan ucapan yang ditangkap dari microphone. Hasil dari terjemahan `recognize_google()` ini berisikan sebuah alternatif kemiripan kalimat yang diucapkan berupa transkrip dan memberikan nilai dari confidence atau keyakinan terjemahan kalimat dari suara yang ditangkap.

```
with PATH as source:
    audio = a.record(source)
    print(a.recognize_google(audio, Language='id-ID', show_all=True))
with sr.Microphone() as source:
    print("Ceritakan pengalamamu disini!")
    audio = a.listen(source)

try:
    text = a.recognize_google(audio, Language='id-ID')
    print("Kamu berkata : ", text)
except sr.UnknownValueError:
    print("Audio tidak dapat dipahami")
except sr.RequestError as e:
    print("Hasil tidak dapat diperoleh: {}".format(e))
```

Gambar 5 Source code `recognize_google`

Pada tabel 1 Hasil uji coba *speech to text* dengan keadaan sepi dan ramai dengan jarak perekaman suara yang diambil masing-masing 1 meter dengan menggunakan suara 15 anak yang akan dihitung menggunakan rumus sebagai berikut :

$$\text{Akurasi suara ke dalam teks} = \frac{\sum \text{match}}{\sum \text{tp}} \times 100 \quad (1)$$

Dimana $\sum \text{match}$ adalah jumlah suara rekaman yang terdeteksi dan $\sum \text{tp}$ adalah jumlah data testing. Perhitungan akurasi suara yang didapat dengan rumus diatas menghasilkan nilai akurasi untuk keadaan sepi hasil akurasi yang didapat adalah hasil kecocokan suara yang didapat dari 15 suara anak dengan keadaan sepi maka persentasenya didapatkan sebanyak 93% dengan kepercayaan penerjemahan kualitas suara bagus karena saat pengambilan suara suasana sekitar lebih sepi dan ada beberapa suara anak saat rekaman terlalu cepat dan pelan sehingga tidak dapat tertangkap dengan baik oleh sistem.

Sedangkan untuk kondisi ramai dihasilkan dari 15 suara anak dengan jarak yang sama yaitu 1 meter maka hasil presentase akurasi suara didapatkan sebanyak 86% tingkat kepercayaan terjemahan suara hasilnya lebih rendah karena disebabkan saat pengambilan suara suasana sekitar terlalu ramai dan suara anak yang terlalu pelan dan alat yang digunakan tidak dapat menangkap suara dalam keadaan ramai dengan baik.

Berikut ini merupakan contoh tabel uji coba dari *speech to text* keadaan sepi dan ramai dengan jarak 1 meter :

Tabel 1 Tabel contoh uji coba *speech to text*

No	Input Ucapan	Kondisi Sepi		Kondisi Ramai	
		Output	Confidence	Output	Confidence
1.	Aku sedih karena covid 19 tidak kunjung usai dan aku tidak bisa melakukan aktivitas ku diluar rumah	Aku sedih karena covid 19 tidak kunjung usai dan aku tidak bisa melakukan aktivitas ku diluar rumah	0.84798175	Aku sedih karena covid 19 tidak kunjung usai	0.87250954
2.	Aku susah berkonsentrasi saat belajar di rumah dan di sekolah aku selalu teringat ayah dan ibu	Aku susah berkonsentrasi saat belajar di rumah dan di sekolah aku selalu teringat ayah dan ibu	0.96728683	Aku susah berkonsentrasi saat belajar dirumah di sekolah aku teringat ayah ibu	0.86989975

No	Input Ucapan	Kondisi Sepi		Kondisi Ramai	
		Output	Confidence	Output	Confidence
3.	Aku setiap malam sering susah tidur aku baru bisa tidur pukul 3 pagi	Aku setiap malam sering susah tidur aku baru bisa tidur saat pukul 3 pagi	0.96728677	Setiap malam susah tidur aku baru bisa tidur pukul 3 pagi	0.96728671
4.	Sejak pandemi covid 19 aku selalu melakukan aktivitas di dalam kamar dan melakukan hobiku di dalam kamar	Sejak pandemic covid 19 aku selalu melakukan aktivitas di dalam kamar dan melakukan hobiku di dalam kamar	0.9366901	Audio tidak dapat dipahami karena tidak dapat ditangkap dengan jelas oleh komputer	
5.	Sejak pandemi aku lebih pendiam dan tidak melakukan kegiatan apa-apa	Sejak pandemi aku lebih pendiam dan tidak melakukan kegiatan apa-apa	0.90177041	Audio tidak dapat dipahami karena tidak dapat ditangkap dengan jelas oleh komputer	
6.	Sejak ada pandemic covid aku selalu cemas karena aku takut tertular	Sejak ada pandemic office aku selalu cemas karena aku takut tertular	0.90930402	Aku selalu cemas cemas karena aku takut	0.96728659
7.	Jantungku berdetak kencang aku takut untuk pergi kemana-mana	Jantungku berdetak kencang aku takut untuk pergi kemana-mana	0.87986791	Jantungku berdetak kencang karena aku takut untuk pergi ke mana - mana	0.87796211
8.	Sejak pandemic covid 19 aku tidak pernah bermain diluar Bersama teman-teman	Sejak pandemi covid aku selalu cemas karena aku takut tertular	0.96728671	Ucapkan dan Kaget aku tidak pernah bermain lagi di luar bersama teman-teman	0.91728677
9.	Saat pandemic covid 19 nilaku turun karena susah sekali untuk belajar	Sejak pandemi covid 19 nilaiku turun karena susah sekali untuk belajar	0.96728671	Karena susah sekali untuk keluarga	0.55728659
10.	Sejak pandemic covid 19 aku mudah sekali marah-marrah jika orang disekitarku membuat kesalahan kecil	Sejak pandemi covid 19 aku mudah sekali marah-marrah jika orang disekitar kau membuat kesalahan kecil	0.95253462	Audio tidak dapat dipahami. Audio tidak dapat dipahami karena tidak dapat ditangkap dengan jelas oleh komputer	

3.3 Pengujian *speech to text* pada chatbot

Pada subbab ini peneliti melakukan sebuah uji coba *speech to text* pada *chatbot*. Pada tabel 3 uji coba pada *chatbot* sistem pakar hasil uji *speech recognition* dengan menggunakan *speech to text* dan *text to speech* dapat berjalan dengan baik. Untuk *speech to text* dapat menangkap suara dengan baik dan menerjemahkan ucapan ke dalam bentuk teks hasilnya cukup bagus tetapi ada beberapa di keadaan sepi *speech to text* tidak dapat menangkap suara dengan baik dikarenakan saat rekaman suara, suara anak terlalu cepat dan terlalu banyak noise sedangkan untuk *text to speech* hasilnya bagus sistem dapat menerjemahkan teks ke dalam bentuk suara dengan baik dan terdengar jelas.

SISTEM PENGENALAN UCAPAN MANUSIA DENGAN MENGGUNAKAN...

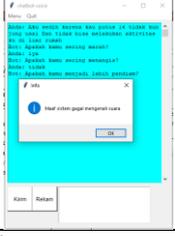
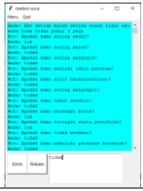
Berikut ini merupakan penjelasan lebih detailnya mengenai hasil uji pada *chatbot* sistem pakar sebagai berikut :

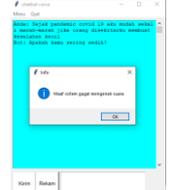
Pada pengujian *chatbot* ini suara yang terdeteksi dari 5 kali percobaan suara pada keadaan sepi dari 15 suara anak dengan masing-masing *delay* 30 detik menghasilkan akurasi sebanyak pada ucapan pertama tercapai 93,3% sebanyak 14 suara anak yang gagal 6,7% sebanyak 1 suara hal ini dikarenakan saat rekaman perekaman suara, suara anak tersebut terlalu kecil, pada percobaan kedua tercapai 86,7% sebanyak 13 suara anak yang gagal 13,3% dikarenakan area sekitar terlalu ramai, suara terlalu cepat dan *aplikasi not responding*, pada percobaan ketiga tercapai 100% dari 15 suara anak, pada percobaan keempat 100% dari 15 suara anak, pada percobaan kelima 100% dari 15 suara anak.

Sedangkan pada pengujian *chatbot* sistem pakar kondisi ramai dengan 5 kali percobaan suara pada keadaan ramai dari 15 suara anak dengan masing-masing *delay* 30 detik untuk menampilkan kesimpulan dan saran dari jenis trauma yaitu pada suara ucapan pertama yang tercapai sebanyak 60% sebanyak 9 suara dan yang gagal mencapai 40% sebanyak 6 suara, pada suara ucapan kedua yang tercapai 26,7% sebanyak 4 suara dan yang gagal sebanyak 73,3% sebanyak 11 suara, pada suara ucapan ketiga tercapai 46,7 % sebanyak 7 suara dan gagal sebanyak 53% sebanyak 8 suara, pada suara ucapan ketiga tercapai 46,7 % sebanyak 7 suara dan gagal sebanyak 53% sebanyak 8 suara, pada suara ucapan kelima tercapai 53% sebanyak 8 suara dan gagal 46,7% sebanyak 7 suara. Hal ini disebabkan oleh beberapa factor yaitu suara terlalu pelan, suara terlalu cepat, microphone tidak dapat menangkap suara karena jaraknya terlalu jauh, dan *chatbot* tidak dapat menyimpulkan dengan jelas.

Berikut ini merupakan contoh dari hasil pengujian *speech to text* pada *chatbot* sistem pakar pada 2 kondisi yaitu sepi dan ramai dengan 5 kali percobaan dan *delay* 30 detik untuk memberikan saran dan kesimpulan pada aplikasi seperti dibawah ini :

Tabel 3 Hasil pengujian *speech to text* pada *chatbot* sistem pakar pada 2 kondisi yaitu sepi dan ramai dengan 5 kali percobaan dan *delay* 30 detik

No	Jenis Kelamin	Kalimat yang diucapkan	Suara terdeteksi pada percobaan
1.	Perempuan	Aku sedih karena covid 19 tidak kunjung usai dan tidak bisa melakukan aktivitas ku diluar rumah (kondisi sepi)	Percobaan 1 : Hasilnya gagal mendeteksi suara karena saat rekaman suara terlalu kecil dan cepat. 
2.	Laki-laki	Aku setiap malam sering susah tidur aku baru bisa tidur pukul 3 pagi (kondisi sepi)	Percobaan 1 : Berhasil aplikasi dapat menangkap suara dengan baik dan juga dapat mendeteksi trauma yang dirasakan. 

			
3.	Laki-laki	Aku marah sekali hari ini karena aku tidak boleh bermain (kondisi ramai)	Percobaan 3 : Gagal karena saat perekaman suara. Suara anak terlalu pelan dan kondisi ramai 
4.	Perempuan	Akuloh sedih karena tidak bisa menolong ayah ibuku saat sedang sakit covid 19 (kondisi sepi)	Percobaan 5 : Gagal, karena suara terlalu cepat dan keadaan sekitar terlalu banyak noise. 

4. Kesimpulan

Berdasarkan pada penelitian Sistem Pengenalan Ucapan Manusia Dengan Menggunakan Google Api Pada Sistem Pakar Konseling Trauma Pada Anak maka dapat ditarik kesimpulan sebagai berikut:

Pada percobaan *speech recognition* dengan menggunakan metode *speech to text* dengan *google api* menggunakan 3 indikator yaitu jarak 1 meter dalam keadaan sepi dan ramai dengan 30 anak yang terbagi menjadi 15 anak keadaan sepi dan 15 anak keadaan ramai. Hasil akurasi untuk keadaan sepi akurasi sebanyak 93% dari 100% karena pada keadaan sepi terdapat suara yang terlalu cepat, sedangkan keadaan ramai akurasi sebanyak 86% dari 100% hasilnya lebih rendah sebanyak 14% karena disebabkan oleh kondisi sekitar dan kecepatan suara pada saat rekaman.

Hasil uji *speech recognition* pada *chatbot* hasilnya sesuai *bot* bisa membuat kesimpulan dan mendeteksi gejala dan dapat mendiagnosa jenis trauma, *bot* dapat menangkap suara dan diubah ke mode *speech to text* dan *bot* dapat melakukan *text to speech* untuk hasil *text to speech bot* dapat menerjemahkan teks ke suara dan hasilnya cukup terdengar jelas.

Hasil uji akurasi dari percobaan suara pada *chatbot* dengan keadaan sepi dari 15 suara anak dengan masing-masing *delay* 30 detik sebagai berikut pada ucapan pertama tercapai pada ucapan pertama tercapai 93,3% sebanyak 14 suara anak yang gagal 6,7% sebanyak 1 suara, pada ucapan kedua tercapai 86,7% sebanyak 13 suara anak yang gagal 13,3%, pada percobaan ketiga tercapai 100% dari 15 suara anak. Pada percobaan keempat 100% dari 15 suara anak, pada percobaan kelima tercapai 100% dari 15 suara anak. Suara yang gagal disebabkan oleh saat rekaman suara terlalu pelan, suara terlalu cepat dan aplikasi terjadi *error not responding*. Sedangkan hasil uji akurasi dari percobaan suara pada *chatbot* dengan keadaan ramai dari 15 suara anak dengan masing-masing *delay* 30 detik sebagai berikut pada suara ucapan pertama yang tercapa sebanyak 60% sebanyak 9 suara dan yang gagal sebanyak 40% sebanyak 6 suara, pada suara ucapan kedua tercapai 26,7% sebanyak 4 suara dan yang gagal sebanyak 73,3% sebanyak 11 suara, pada suara ucapan

ketiga tercapai 46,7% sebanyak 7 suara dan gagal sebanyak 53% sebanyak 8 suara, pada suara ucapan keempat tercapai 46,7% sebanyak 7 suara dan gagal sebanyak 53% sebanyak 8 suara, pada suara ucapan kelima tercapai 53% sebanyak 8 suara dan gagal 46,7% sebanyak 7 suara.

Hasil dalam keadaan ramai kurang baik disebabkan oleh beberapa factor yaitu suara terlalu pelan, terlalu cepat, microphone tidak dapat menangkap suara karena jaraknya terlalu jauh dan *chatbot* tidak dapat menyimpulkan dengan jelas. Maka hasil dari perbandingan pengambilan suara pada *chatbot* sistem pakar suara yang diambil di dua suasana tempat pengambilan rekaman suara di tempat sepi lebih bagus daripada di tempat ramai.

5. Saran

Berdasarkan hasil dari penelitian yang telah dilakukan, untuk penelitian berikutnya adalah sebaiknya aplikasi dapat dikembangkan lebih baik lagi, aplikasi dapat dijalankan di aplikasi android, tampilan aplikasi dipercantik lagi agar lebih ramah dengan anak-anak, dan untuk penelitian selanjutnya menggunakan perangkat yang dapat meminimalisir noise.

Daftar Pustaka

- Akbar, A. *et al.* (2019) 'IMPLEMENTASI GOOGLE SPEECH API PADA APLIKASI KOREKSI HAFALAN AL-QUR ' AN BERBASIS ANDROID (The Implementation of the Google Speech on Qur ' an Recitation Correction', *JTIKA*, 1(1), pp. 1–8.
- Akbar, A. *et al.* (2021) 'Pengembangan Aplikasi Asisten Pintar Pembuka Al Qur'an 30 Juz dengan', *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(10), pp. 1008–1015.
- Alsunaidi, N. *et al.* (2018) 'Abjad: Towards Interactive Learning Approach to Arabic Reading Based on Speech Recognition', *Procedia Computer Science*, 142, pp. 198–205. doi: 10.1016/j.procs.2018.10.476.
- Bengio, S. *et al.* (2002) 'Confidence measures for multimodal identity verification', *Information Fusion*, 3(4), pp. 267–276. doi: 10.1016/S1566-2535(02)00089-1.
- Buana, I. K. S. (2020) 'Implementasi Aplikasi Speech to Text untuk Memudahkan Wartawan Mencatat Wawancara dengan Python', *Jurnal Sistem dan Informatika (JSI)*, 14(2), pp. 135–142. doi: 10.30864/jsi.v14i2.293.
- Dutoit, T. (1997) *An Introduction to Text-to-Speech Synthesis*. Springer Netherlands (Text, Speech and Language Technology). Available at: <https://books.google.co.id/books?id=xjXU3FujWbwC>.
- Intan, D. *et al.* (2017) 'PEMANFAATAN CLOUD SPEECH API UNTUK PENGEMBANGAN MEDIA PEMBELAJARAN BAHASA INGGRIS MENGGUNAKAN', *Jurnal Telematika*, 10(2), pp. 92–105.
- Muhammad, Hidayat, S. and Amrullah, A. Z. (2019) 'Speech Recognition Untuk Aplikasi Kamus Bahasa Indonesia-Sumbawa Berbasis Android', *Jurnal BiTe*, 1(2), pp. 126–137.
- Rabiner, L. R. (2017) 'Repeated interactions versus social ties: Quantifying the economic value of trust, forgiveness, and reputation using a field experiment',

-
- MIS Quarterly: Management Information Systems*, 41(3), pp. 841–866. doi: 10.25300/MISQ/2017/41.3.08.
- Setiawan, R. (2016) ‘Perancangan Arsitektur Enterprise Untuk Perguruan Tinggi Swasta Menggunakan Togaf Adm’, *Jurnal Algoritma*, 12(2), pp. 548–561. doi: 10.33364/algoritma/v.12-2.548.
- Suryani, D. and Amalia, E. L. (2017) ‘Chatbot’, *Smartics*, 3(1), pp. 71–92. doi: 10.21721/p2p.2021v8n1.p71-92.
- Wisdhani, Z., Ontowirjo, A. H. J. and Kambey, F. D. (2021) ‘Asmaul Husna Learning Application Using The Speech recognition’, *Jurnal Teknik Informatika*, 16(3), pp. 283–292.