APLIKASI KAMUS BAHASA BATAK-INGGRIS (SPEECH TO SPEECH) BERBASIS ANDROID SEBAGAI FITUR PENDUKUNG PARIWISATA DANAU TOBA SEBAGAI DESTINASI WISATA INTERNASIONAL

Aninda Muliani

Program Studi Sistem Informasi, Universitas Islam Negeri Sumatera Utara Medan
Jl. Lap. Golf, Kecamatan Pancur Batu, Kabupaten Deli Serdang
anindamh@uinsu.ac.id

Abstract - Lake Toba in North Sumatra is a tour that is being intensified by President Jokowi to become one of the international tourist destinations. But unfortunately, the residents around the Lake Toba area, which are majority controlled by Toba Batak residents, do not yet have the skills to communicate in English so it is difficult to communicate with foreign tourists. The purpose of this study is to apply the speech recognition method to the voice-based dictionary application in the Batak-English dictionary application. The method used is qualitative research. The use of speech to speech can be used as an analysis of the Batak language voice data which will be translated into English and Indonesian and this application can recognize the Batak language.
Keywords - Speech to speech, dictionary, Batak, lake, Toba.

Abstrak - Wisata Danau Toba yang terdapat di Sumatera Utara ini adalah wisata yang sedang digiatkan oleh Presiden Jokowi untuk menjadi salah satu destinasi wisata internasional. Namun sayangnya, penduduk di sekitar wilayah Danau Toba, yang mayoritas dikuasai oleh warga Batak Toba, belum memiliki kecakapan dalam berkomunikasi dalam Bahasa Inggris sehingga kesulitan berkomunikasi dengan wisatawan asing. Tujuan dari penelitian ini adalah menerapkan metode *speech recognition* pada aplikasi kamus berbasis suara pada aplikasi kamus bahasa Batak-Inggris. Metode yang digunakan adalah penelitian kualitatif. Penggunaan *speech to speech* dapat digunakan sebagai analisis pada data suara Bahasa Batak yang akan di terjemahkan ke Bahasa Inggris dan Indonesia dan aplikasi ini dapat mengenali Bahasa Batak.
Kata Kunci - *Speech to speech*, kamus, Batak, danau, Toba.

I. PENDAHULUAN

Teknologi pengenalan suara mengalami perkembangan yang sangat pesat. Teknologi pengenalan suara atau yang lebih dikenal dengan istilah speech recognition adalah teknologi yang memanfaatkan peralatan dengan sumber masukan suara kemudian diinterpretasikan dengan algoritma tertentu untuk suatu tujuan [1], misalnya untuk menerjemahkan bahasa asing ke bahasa lokal untuk membantu komunikasi di tempat-tempat wisata.

Danau Toba adalah danau terluas se-Asia Tenggara dikenal dengan danau yang dengan pemandangan yang sangat indah sehingga banyak wisatawan mancanegara tertarik untuk mengunjunginya [2]. Wisata Danau Toba yang terdapat di Sumatera Utara ini adalah wisata yang sedang digiatkan oleh Presiden Jokowi untuk menjadi salah satu destinasi wisata internasional. Namun sayangnya, penduduk di sekitar wilayah Danau Toba, yang mayoritas dikuasai oleh warga Batak Toba, belum memiliki kecakapan dalam berkomunikasi dalam Bahasa Inggris sehingga kesulitan berkomunikasi dengan wisatawan asing [3].

Ada beberapa alat terjemahan yang tersedia saat ini. Namun, sebagian besar perangkat lunak terjemahan bahasa berbasis teks; pengguna menulis kata atau kalimat dalam Bahasa Indonesia, dan program kemudian menerjemahkannya ke bahasa lain sebelum menampilkan teks terjemahan di layar. Aplikasi ini tentunya membutuhkan waktu dalam pengetikan masukan dan proses pencarian kata sehingga proses komunikasi antara wisatawan dan masyarakat lokal menjadi kurang lancar, selain itu, tidak semua masyarakat lokal merupakan kaum literasi (bisa membaca dan menulis). Untuk itu diperlukan sebuah aplikasi kamus penerjemah bahasa Batak-Inggris berbasis suara untuk membantu memperlancar komunikasi antara wisatawan mancanegara dengan masyarakat lokal Danau Toba.

A. Bahasa Batak Toba

Orang Batak Toba di Sumatera Utara, Indonesia, kebanyakan mendiami daerah sekitar Danau Toba, Pulau Samosir, dan wilayah Tapanuli bagian utara. Mereka berbicara bahasa Batak Toba (Hata Batak Toba) [4]. Kelompok bahasa Batak termasuk bahasa Batak Toba, yang merupakan anggota rumpun bahasa Austronesia.

Di wilayah barat dan selatan Danau Toba saat ini diperkirakan terdapat 2.000.000 penutur bahasa Batak Toba [5]. Dibandingkan dengan bahasa Batak lainnya

(Simalungun, Karo dan Pakpak), tata bahasa bahasa Batak Toba mirip dengan bahasa Batak Angkola dan Mandailing, sehingga memudahkan komunikasi antara ketiga suku ini [6].

Selain penutur yang bermukim di kawasan Danau Toba, bahasa Batak Toba juga banyak digunakan di kawasan lain. Khususnya para orang tua yang berdomisili di luar Danau Toba sering memanfaatkan bahasa tersebut untuk berkomunikasi sehari-hari [7].

B. Speech Recognition

Aplikasi yang menggunakan *Speech Recognition* dapat mengubah kata yang diucapkan menjadi file audio digital, yang kemudian diubah menjadi teks dan akan dicocokkan dengan kata-kata yang ada pada *database* [8].

Speech Recognition umumnya dikenal sebagai *Automatic Speech Recognition* (ASR) atau *Speech to Text* (STT), adalah cabang interdisipliner linguistik komputasi yang mengeksplorasi pendekatan dan teknologi yang memungkinkan pengenalan computer dan terjemahan bahasa lisan ke dalam teks. Berikut ini adalah algoritma *Speech to Text*:

1. *Speech Signal* masuk melalui mikrofon.
2. *Feature extraction* dari sinyal suara.
3. *Training dataset feature extraction*
4. *Testing dataset feature extraction*
5. Keluaran berupa teks dari sinyal yang diidentifikasi [9]

Dijelaskan Ahmad [10] pada laman Medium.com bahwa *Speech to text* (STT) merupakan metode yang mengubah bahasa lisan menjadi bahasa tulisan. Saat manusia berbicara ke dalam sistem, teknologi mengenali kalimat atau kata yang diucapkan dan menghasilkan teks yang sangat cocok. Untuk mendapatkan skala atau karakteristik yang sederhana untuk diproses pada tahap berikutnya, sinyal ucapan pertama akan diteruskan ke bagian *speech analyzer*. Pola karakteristik yang unik akan dibuat untuk setiap ucapan individu.

Transisi sinyal ucapan dari domain waktu ke domain frekuensi akan dilakukan oleh penganalisa sintaks. Setiap sinyal tampaknya memiliki karakteristik yang berbeda dalam domain frekuensi untuk waktu yang singkat. Pengucapan fonem, bagaimanapun, sering berbeda dari orang ke orang dan dipengaruhi oleh fonem lain, suasana hati seseorang, kebisingan latar belakang, dan variabel lainnya. Setiap unit bunyi (fonem) pembentuk tuturan akan mendapat pengenalan dari sistem *speech to text*, yang kemudian akan dicari kombinasi hasil tuturan yang paling enak didengar. Sistem yang lebih sederhana hanya dapat mengenali sekumpulan kecil kata dengan jumlah terbatas. Meskipun sistem ini lebih tepat dan sederhana untuk dilatih, sistem ini terbatas kemampuannya untuk mengenali kata-kata yang bukan bagian dari kosa kata yang diajarkan [10].

Gambar 1. Diagram Alir *Speech to Text* [10]

Proses konversi ini juga dapat dilakukan untuk mengubah teks menjadi suara (*text to speech*). Sedangkan *speech to speech* merupakan penggabungan dari kedua proses tersebut, yaitu *speech to text* kemudian dilanjutkan dengan *text to speech*.

C. UML (Unified Modeling Language)

Salah satu alat paling efektif di bidang pengembangan sistem berorientasi objek adalah UML (*Unified Modeling Language*). Hal ini karena UML menawarkan bahasa pemodelan visual yang memungkinkan pengembang sistem untuk menyusun *blueprint* ide mereka dalam format standar yang mudah dipahami dan dilengkapi dengan metode yang efisien untuk berbagi dan mendiskusikan ide mereka dengan orang lain [11].

Berikut ini adalah tujuan UML, yaitu:

1. Menawarkan bahasa pemodelan visual ekspresif yang siap pakai untuk memfasilitasi pengembangan dan pertukaran model.
2. Menyediakan bahasa pemodelan yang tidak bergantung pada bahasa pemrograman atau prosedur teknik tertentu.
3. Menyatukan praktik-praktik terbaik yang terdapat pada pemodelan [12].

II. METODE PENELITIAN

Jenis penelitian yang dikenal sebagai penelitian kualitatif atau qualitative research menghasilkan hasil yang tidak dapat diperoleh melalui metode statistik atau teknik kuantitatif lainnya [13].

Untuk mendapatkan hasil yang diinginkan dari suatu penelitian, metode penelitian berfungsi sebagai kerangka kerja. Beberapa teknik digunakan untuk mengumpulkan data untuk penelitian ini, termasuk yang tercantum di bawah ini:

1. Wawancara

Wawancara dilakukan dengan cara yang tidak terstruktur dan pelaksanaannya sama dengan percakapan informal. “Wawancara dalam penelitian kualitatif naturalistik, terutama untuk pemula, biasanya tidak terstruktur,” menurut Nasution [14]. Mencari tahu pengetahuan yang mendalam tentang perspektif orang lain adalah tujuannya.

2. Observasi

Pengamatan langsung nonpartisipatif, atau pengamatan langsung tanpa berpartisipasi aktif dalam kegiatan di tempat penelitian, adalah pendekatan observasi yang digunakan. Pengamatan dilakukan secara tertutup. Nasution [14] menyatakan bahwa “observasi dengan pengamatan tersembunyi bertujuan untuk memperoleh data yang valid dan reliable dan dapat dipercaya karena tidak dibuat-buat”.

Peneliti melakukan observasi tanpa bantuan peralatan umum lainnya. Saat melakukan observasi, banyak hal yang menjadi perhatian peneliti, yaitu:

a. Isi observasi
b. Mencatat observasi

3. Kajian Pustaka

Salah satu cara pengumpulan data yang tidak langsung bersentuhan dengan subjek penelitian adalah Kajian pustaka. Semacam pengumpulan data yang disebut penelitian dokumen melihat berbagai jenis dokumen yang berguna untuk analisis [15]. Berikut dokumen yang digunakan dalam penelitian ini:

a. Kamus Bahasa Batak Toba
b. Kamus Bahasa Inggris
c. Gramatika dan Semantik Bahasa Batak Toba

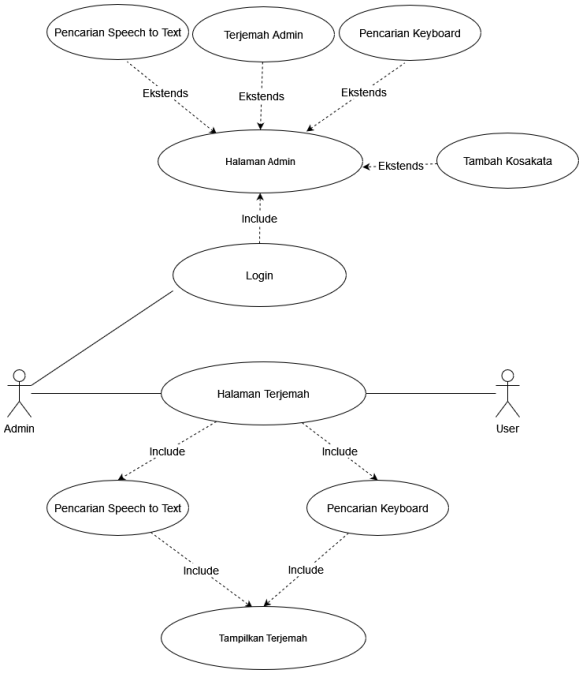
III. HASIL DAN PEMBAHASAN

A. Perancangan Sistem

Untuk mempermudah peneliti dalam membuat aplikasi, proses perancangan sistem yang disarankan pada poin ini dirinci dengan UML (*Unified Modeling Language*), khususnya menggunakan *use case diagram*, *activity diagram*, dan *sequence diagram*.

1. Use Case Diagram

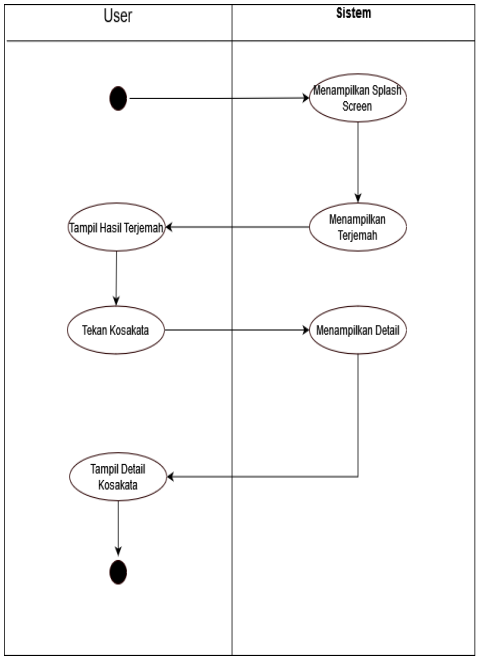
Use case diagram digunakan untuk menggambarkan bagaimana sistem berinteraksi dengan aktor. Jenis interaksi antara pengguna sistem dan sistem dapat digambarkan dengan *use case* [16]. Langkah selanjutnya yaitu merancang *use case diagram* setelah mengidentifikasi aktor dan *use case*. Ada dua pelaku dalam aplikasi yang dapat menggunakannya: admin, yang dapat menambahkan kata dengan masuk dengan *username* dan *password*, dan *user* yang menggunakan aplikasi. Berikut rancangan *use case diagram* aplikasi Kamus Bahasa Batak Toba :



Gambar 2. *Use case Diagram* Kamus Bahasa Batak Toba

2. Activity Diagram

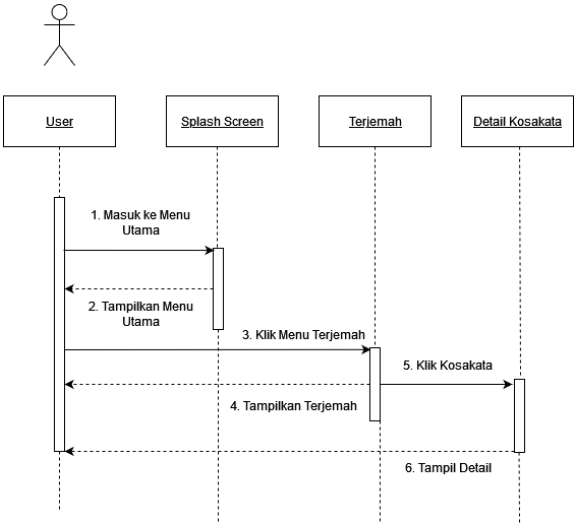
Activity diagram menampilkan aktivitas sistem sebagai serangkaian tindakan, awal setiap tindakan, dan setiap keputusan yang mungkin dibuat sepanjang tindakan. *Activity diagram* dapat digunakan untuk menggambarkan beberapa tindakan simultan [17]. Aktivitas diawali dengan *splash screen*, kemudian dalam menu terjemah akan dapat menerjemahkan bahasa batak toba ke dalam bahasa inggris secara perkata.



Gambar 3. *Activity Diagram* Terjemah Bahasa Batak Toba

3. *Sequence Diagram* Terjemah

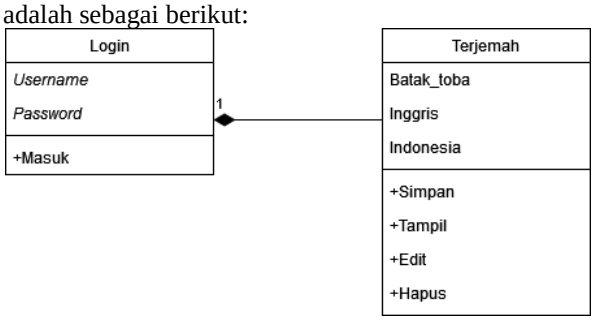
Sequence Diagram menunjukkan bagaimana berbagai objek berinteraksi satu sama lain dari waktu ke waktu. Dimana setiap tindakan diikuti dengan deskripsi hasil atau reaksi yang mengikuti setiap tindakan tersebut [18]. *Sequence Diagram* diawali dengan *splash screen*, kemudian dalam menu terjemah akan dapat menerjemahkan bahasa batak toba ke dalam bahasa inggris secara perkata.



Gambar 4. Squensial Diagram Pengguna

4. Class Diagram

Class diagram representasi visual dari kelas sistem yang ada dan hubungan logisnya. *Class diagram* menjelaskan struktur statis sistem [19]. Perancangan *class diagram* aplikasi yang disarankan adalah sebagai berikut:



Gambar 5. *Class Diagram*

B. Implementasi Sistem

Dibawah ini merupakan implementasi atau penerapan dari aplikasi yang telah dibangun.

1. Splash Screen

Pada gambar 6 merupakan tampilan pada halaman *splash screen* pada awal aplikasi dibuka.

Gambar 6. *Interface Splash Screen*

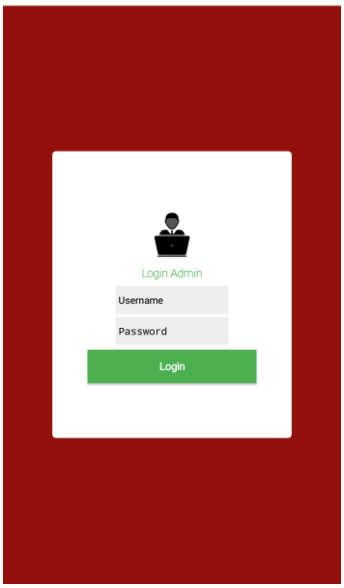
2. Halaman Terjemah

Halaman ini menampilkan kata-kata berbahasa batak beserta terjemahannya dalam Bahasa Inggris

Gambar 7. Halaman Terjemah

3. Halaman *Login* Admin

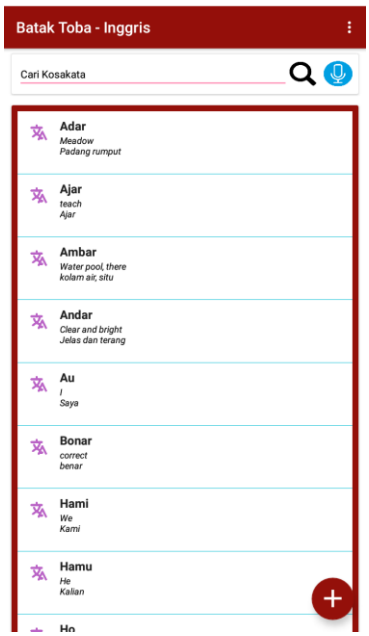
Dibawah ini tampilan *login* untuk admin.



Gambar 8. Halaman *Login*

4. Halaman Terjemah Admin

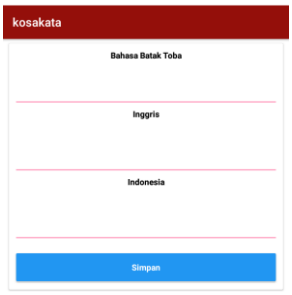
Admin dapat menambahkan kosa kata pada halaman ini.



Gambar 9. Input Nilai

5. Halaman Tambah Kosakata

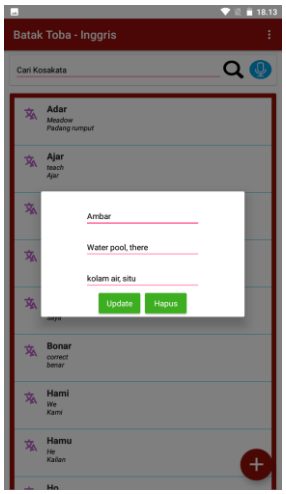
Tampilan tambah kosa kata dapat dilihat pada gambar 10 berikut.



Gambar 10. Halaman Tambah Kosa Kata

6. Edit dan Hapus Kosa Kata

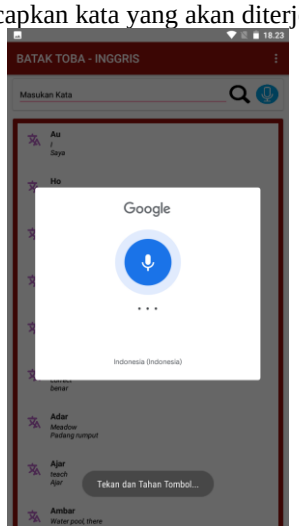
Berikut ini adalah tampilan aplikasi saat mengedit atau menghapus kosa kata.



Gambar 11. Tampilan Edit dan Hapus Kata

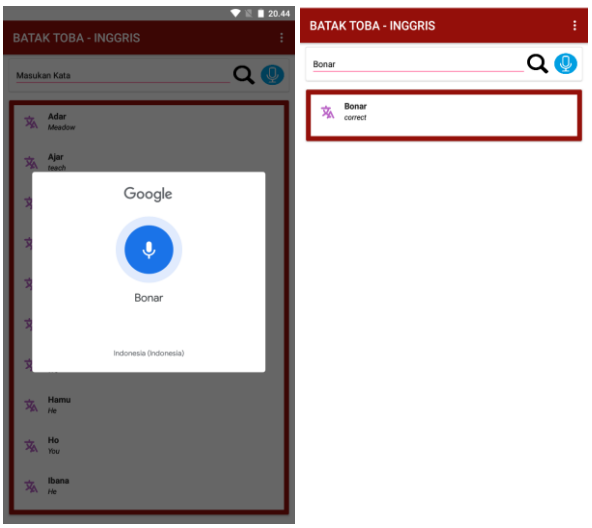
7. Halaman *Speech*

Gambar 12 berikut ini adalah tampilan *speech* untuk mengucapkan kata yang akan diterjemahkan.

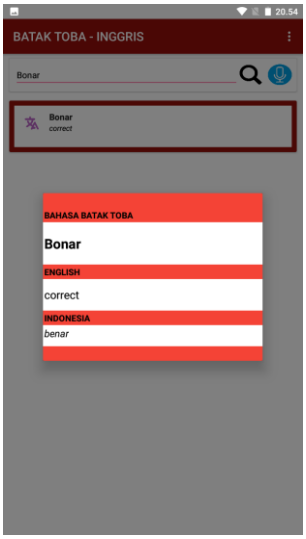


Gambar 12. Tampilan *Speech*

8. Hasil Terjemah *Speech*



Gambar 13. Tampilan Terjemah *Speech*



Gambar 14. Tampilan Detail Terjemahan

C. *Pengujian Sistem*

Pengujian *black-box* dilakukan untuk mengevaluasi teknologi yang digunakan dalam penelitian ini, serta ketepatan deteksi suara. Pendekatan pengujian *black-box* melibatkan penginstalan aplikasi pada *smartphone* untuk mengevaluasi apakah setiap fungsionalitas sistem berjalan dengan benar sesuai dengan analisis perancangan [20]. Ponsel yang digunakan untuk pengujian adalah Realme 5 dengan versi Android 10. Tabel pengujian yang dilakukan pengguna disediakan di bawah ini:

Tabel 1. Pengujian dengan metode *Black Box*

No	Skenario	Hasil yang diharapkan	Hasil
1	Buka Aplikasi	Menampilkan menu splash screen	Ok
2	Halaman Terjemah	Menampilkan terjemah	Ok
3	Klik option menu untuk login	Menampilkan form login (Username dan password)	Ok
4	Melakukan login	Peringatan username dan password salah	Ok
5	Login berhasil	Menuju halaman terjemah admin	Ok
6	Klik Floating Button Tambah	Menuju halaman tambah kosakata	Ok
7	Berhasil memasukkan kosakata baru	Menampilkan daftar kosakata yang tersimpan di halaman terjemah admin	Ok
8	Klik kosakata di halaman terjemah untuk mengedit dan menghapus	Menampilkan form edit dan hapus kosakata	Ok
9	Mengklik option menu untuk logout	Keluar dan kembali ke halaman terjemah	Ok
10	Ketik kata pada kolom pencarian, klik icon pencarian	Menampilkan *speech* terjemahan dan hasil terjemah kata yang diketik	Ok
11	Klik kosakata untuk melihat lebih detail	Menampilkan speech dan detail kosakata yang diklik	Ok
12	Melakukan pencarian dengan *speech to text*	Menampilkan perintah untuk tekan tombol dan ucapkan kata	Ok
13	*Google speech* mendeteksi kata yang diucapkan	Menampilkan suara dan hasil deteksi kosakata yang dicari	Ok

IV. KESIMPULAN

Berdasarkan perancangan, implementasi dan pengujian pada penelitian yang dilakukan, dapat ditarik kesimpulan sebagai berikut:

1. Penggunaan *speech to speech* dapat digunakan sebagai analisis pada data suara Bahasa Batak yang akan di terjemahkan ke Bahasa Inggris dan Indonesia.

2. *Speech to speech* dapat mengenali Bahasa Batak.
3. Akurasi dipengaruhi oleh berbagai parameter dalam sistem pengenalan suara. Akurasi dipengaruhi oleh berbagai faktor, seperti aksen saat mengucapkan kata, perbedaan dialek, pengucapan yang tidak jelas atau lamban, dan jumlah kata yang terdengar mirip.

DAFTAR PUSTAKA

[1] W. Kurniadi dan I. Maliki, "Speech Recognition Menggunakan Elman Recurrent Neural Network Untuk Kata Dalam Bahasa Indonesia," Universitas Komputer Indonesia, 2018.

[2] M. A. B. Butar, "Perancangan Sistem Informasi Geografis Pemetaan Lokasi Objek Wisata di Daerah Danau Toba Berbasis Android," *Kumpul. Karya Ilm. Mhs. Fak. sains dan Tekhnologi*, vol. 1, no. 1, 2019.

[3] E. D. Harianja, R. H. Harahap, dan Z. Lubis, "Budaya Batak Toba dalam Pelayanan Pariwisata Danau Toba di Parapat," *Perspektif*, vol. 10, no. 2, hal. 301–312, 2021, doi: 10.31289/perspektif.v10i2.4306.

[4] N. D. Astri, R. Arista, E. K. B. Ginting, dan S. D. Y. B. Ginting, "Revitalsasi Legenda Danau Toba Melalui Komik," *J. Salaka J. Bahasa, Sastra dan Budaya Indones.*, vol. 3, no. 2, hal. 60–68, 2021.

[5] M. N. M. Sitohang, "Kesetiaan Mahasiswa Asal Sumatera Terhadap Bahasa Batak Toba di Palangka Raya," *Suar Betang*, vol. 11, no. 2, hal. 113–127, 2016.

[6] L. H. Napitupulu dan Y. N. Silaban, "Kajian Leksikostatistik Pada Bahasa Batak Toba Dan Batak Angkola Mandailing," *J. Bhs. Indones. Prima*, vol. 2, no. 2, hal. 82–90, 2020, doi: 10.34012/bip.v2i2.1336.

[7] T. Sibarani, "Pelestarian Bahasa Batak Toba dari Tinjauan Sosiologi dan Struktur Bahasa," *Medan Makna*, vol. XIII, no. 2, hal. 203–214, 2015.

[8] A. Muliani, "Penerapan Teknologi Speech Recognition (Voice To Sign) Untuk Membantu Komunikasi Dengan Penyandang Disabilitas Pendengaran," *J. Teknovasi*, vol. 6, no. 3, hal. 49–53, 2019.

[9] M. F. T. Afkar, B. Irawan, dan S. M. Nasution, "Speech To Text Menggunakan Metode Hidden Markov Model," 2019.

[10] T. Ahmad, "Penjelasan Speech To Text," *medium.com*, 2019. https://medium.com/@toriqahmads/penjelasan-speech-to-text-dc10c5178809.

[11] Suendri, Triase, dan S. Afzalena, "IMPLEMENTASI METODE JOB ORDER COSTING PADA SISTEM INFORMASI PRODUKSI BERBASIS WEB," *J. Sekol.*, vol. 4, no. 2, hal. 97–106, 2020.

[12] I. Zufria, "Pemodelan Berbasis UML (Unified Modeling Language) dengan Strategi Teknik Orientasi Objek User Centered Design (UCD) dalam Sistem Administrasi Pendidikan," *Researchgate*, vol. 1, no. 1, hal. 1–16, 2013.

[13] F. Nugrahani, "Metode Penelitian Kualitatif dalam Penelitian Pendidikan Bahasa," *Cakra Books*, vol. 1, no. 1, 2014.

[14] Nasution, *Metode Penelitian Kualitatif Naturalistik*. Jakarta: SinarGrafika, 1996.

[15] S. G. Pamungkas dan Nurhasan, "Faktor Penyebab Taman Sukowati di Kota Sragen Sepi Pengunjung," in *Seminar Ilmiah Arsitektur*, 2020, hal. 399–405.

[16] A. F. Prasetya, Sintia, dan U. L. D. Putri, "Perancangan Aplikasi Rental Mobil Menggunakan Diagram UML (Unified Modelling Language)," *J. Ilm. Komput. Terap. dan Inf.*, vol. 1, no. 1, hal. 14–18, 2022.

[17] P. B. A. A. Putra, Widiarty, V. H. Pranatawijaya, dan N. N. Sari, "Implementasi Aplikasi Android Untuk Sistem Pendaftaran Dan Antrian Pada Poli Covid RSUD Doris Sylvanus," *J. Teknol. Inf. J. Keilmuan Dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, hal. 81–91, 2022.

[18] Widyawati, R. Budiman, dan H. Robbani, "Rancang Bangun Aplikasi Pusat Oleh Oleh Berbasis Android Di Sate Bandeng Hj.Mariyam Kota Serang," *J. Innov. Futur. Technol.*, vol. 4, no. 2, hal. 56–65, 2022, doi: 10.47080/iftech.v4i2.2227.

[19] N. A. H, Ariyansyah, dan M. A. Dermawan, "Perancangan Aplikasi E-Counseling Kesehatan Berbasis Android Studi kasus Rumah Sakit Umum Daerah Kota Prabumulih," *J. Ilmu Komput.*, vol. 2, no. 2, 2021.

[20] M. Z. Faried, A. Mulwinda, dan Y. Primadiyono, "Pengembangan Aplikasi Android Bimbingan Skripsi dengan Fitur Notifikasi," *J. Tek. Elektro*, vol. 9, no. 2, hal. 1–6, 2017.