



Doi: [10.17933/jppi.v11i2.286](https://doi.org/10.17933/jppi.v11i2.286)

## Desain Sistem Speech Recognition Penerjemah Bahasa Toraja Menggunakan Hidden Markov Model

### *Design System Speech Recognition Translator Toraja Language Using Hidden Markov Modelling*

Rismayani<sup>1</sup>, Sri Wahyuni<sup>2</sup>, Novita Sambo Layuk<sup>3</sup>, Rian Hesron Loly<sup>4</sup>, Ayub Nandito Daud<sup>5</sup>

Program Studi Rekayasa Perangkat Lunak Universitas Dipa Makassar<sup>1</sup>

Program Studi Teknik Informatika universitas Dipa Makassar<sup>2, 3,4,5</sup>

Jl. Perintis Kemerdekaan Km.9 Tamalanrea Makassar Sulawesi Selatan 90245, Indonesia

Email : [maya\\_setya@ymail.com](mailto:maya_setya@ymail.com)

Naskah Diterima: 15 September 2019 ; Direvisi : 27 Desember 2021; Disetujui : 27 Desember 2021

#### Abstrak

Bahasa daerah toraja merupakan salah satu bahasa daerah yang ada di provinsi sulawesi selatan tepatnya di kabupaten Tana Toraja Indonesia. Indonesia memiliki beragam Bahasa daerah, Bahasa daerah merupakan salah satu ciri budaya suatu daerah khususnya Bahasa Toraja. Penelitian ini bertujuan untuk membuat sistem Speech Recognition pada penerjemah Bahasa daerah Toraja dengan menggunakan Hidden Markov Model. Hidden Markov Model merupakan metode klasifikasi yang menggunakan pembelajaran dari Baum-Welch, yang mengestimasi beberapa parameter untuk menemukan model terbaik yang menggambarkan training set. Hidden Markov adalah algoritma berbasis probabilistik yang digunakan untuk menentukan kelas kata terbaik untuk setiap kata dalam data uji. Hasil dari penelitian ini adalah sistem penerjemah bahasa daerah Toraja yang menggunakan metode Hidden Markov dapat membantu siapa saja yang belum mengetahui menggunakan dan arti Bahasa Indonesia dari Bahasa daerah Toraja. Berdasarkan pengujian Black Box untuk masing-masing fungsional pada sistem penerjemah bahasa daerah Toraja diperoleh hasil yang valid; demikian juga, pengujian logika menggunakan metode Hidden Markov bebas dari kesalahan logika. Kemudian berdasarkan pengujian kuesioner yang telah diperoleh dari 20 responden diperoleh hasil 80.2% sangat setuju jika sistem penerjemah tersebut bermanfaat.

**Kata kunci:** Hidden Markov Modelling, Bahasa Daerah Toraja, Speech Recognition

#### Abstract

*The Toraja regional language is one of the regional languages in the province of South Sulawesi, precisely in the Tana Toraja district of Indonesia. Indonesia has a variety of regional languages; regional languages are one of the cultural characteristics of a region, especially the Toraja language. This study aims to create a Speech Recognition system for Toraja regional language translators using the Hidden Markov Model. Hidden Markov Model is a classification method that uses Baum-Welch learning, which estimates several parameters to find the best model that describes the training set. Hidden Markov is a probabilistic algorithm used to determine the best word class for each word in the test data. The results of this study are the Toraja regional language translator system using the Hidden Markov method can help anyone who does not know the use and meaning of Indonesian from the Toraja regional language. Based on the Black Box testing for each function in the Toraja regional language translator system, valid results are obtained; Likewise, logic testing using the Hidden Markov method is free from logical errors. Then based on testing the questionnaire that has been obtained from 20 respondents, the results obtained are 80.2% strongly agree if the translator system is functional.*

**Keywords:** Hidden Markov, Toraja Regional Language, Speech Recognition

## PENDAHULUAN

Tana Toraja merupakan salah satu kabupaten di Sulawesi Selatan yang memiliki luas wilayah 3.203 km<sup>2</sup> dan berpenduduk 221.081 jiwa (2010) serta mendiami daerah pegunungan dan mempertahankan gaya hidup yang unik dan masih menunjukkan gaya hidup asli Austronesia. Toraja merupakan salah satu bahasa daerah yang digunakan di Kabupaten Tana Toraja untuk berkomunikasi. Selain sebagai alat komunikasi, bahasa Toraja juga merupakan salah satu warisan budaya yang perlu dilestarikan agar tetap menjadi bagian dari kekayaan budaya Indonesia. Bahasa Indonesia merupakan bahasa resmi yang digunakan oleh masyarakat daerah Tana Toraja. Tana Toraja merupakan kabupaten yang di jadikan sebagai salah satu ikon wisata di Provinsi Sulawesi Selatan.

Menurut badan Bahasa kemdikbud di bulan Agustus 2021 saat ini jumlah Bahasa daerah di Indonesia sebanyak 718 bahasa, salah satunya terdapat Bahasa daerah Toraja. Banyaknya Bahasa daerah yang dimiliki oleh Indonesia yang mewakili tiap daerah dan suku yang sangat tidak memungkinkan masyarakat umum dapat mengetahui seluruh Bahasa daerah tersebut. Begitu pula jika berkunjung atau berada di suatu daerah yang memiliki Bahasa daerah yang berbeda, karena masih banyak daerah di Indonesia yang masih menggunakan Bahasa daerah sebagai Bahasa sehari-hari mereka.

Oleh karena itu, penelitian ini bertujuan untuk membuat sistem speech recognition penerjemah untuk salah satu Bahasa daerah di Indonesia yaitu bahasa Toraja ke Bahasa Indonesia dan begitu pun sebaliknya yang mengimplementasikan Hidden

Markov Model ke dalam sistem penerjemah yang dirancang.

Dalam penelitian ini peneliti mengusulkan penggunaan Hidden Markov Model untuk sistem penerjemah Bahasa Indonesia ke Toraja dan begitu pula sebaliknya yang dimana model ini merupakan bagian dari *finite state* atau *finite automation*. *Finite automation* sendiri adalah kumpulan state yang transisi antar state-nya dilakukan berdasarkan masukan observasi. Pada rantai markov, setiap transisi antar state berisi probabilitas yang mengindikasikan kemungkinan jalur tersebut Akan diambil. Jumlah probabilitas semua transisi yang keluar dari sebuah simpul Sama dengan satu. Parameter-parameter yang ditentukan kemudian dapat digunakan untuk analisis yang lebih jauh, misalnya untuk aplikasi Pattern Recognition. Sebuah HMM (Hidden Markov Model) dapat dianggap sebagai sebuah Bayesian Network dinamis yang paling sederhana.

### Hidden Markov Model

Model Markov Tersembunyi atau lebih dikenal sebagai Hidden Markov Model (HMM) adalah sebuah model statistik dari sebuah sistem yang diasumsikan sebuah proses Markov dengan parameter yang tak diketahui, dan tantangannya adalah menentukan parameter - parameter tersembunyi (state) dari parameter-parameter yang dapat diamati (observer). Pada model Markov umum (Vanilla/Visible Markov Model), state-nya langsung dapat diamati, oleh karena itu probabilitas transisi state menjadi satu-satunya parameter (Zucchini and L. MacDonald 2009). Didalam Model Markov yang tersembunyi, state-nya tidak dapat diamati secara langsung, Akan tetapi yang dapat diamati adalah variabel-variabel

yang terpengaruh oleh state. Setiap state memiliki distribusi probabilitas atas token-token output yang mungkin muncul. Oleh karena itu rangkaian token yang dihasilkan oleh HMM memberikan sebagian informasi tentang sekuens state-state. Hidden Markov Model sangat populer diaplikasikan di bidang speech recognition dan bioinformatics.

### **Penerjemah**

Penerjemahan adalah komunikasi pesan dari bahasa sumber ke bahasa sasaran dengan menggunakan teks yang ekuivalen. Di mana interpretasi tidak diragukan lebih dulu muncul daripada tulisan, penerjemahan baru muncul setelah kemunculan tulisan (Wuryantoro 2018). Orang yang menerjemahkan disebut penerjemah. Penerjemah saat menerjemahkan harus mempertimbangkan beberapa batasan, termasuk konteks, aturan tata bahasa, konvensi penulisan, dan idiom, serta hal lain antara kedua bahasa. Penerjemah selalu menghadapi risiko untuk tanpa sengaja menyusun gaya bahasa maupun idiom dari bahasa sumber ke dalam bahasa sasaran (Hariyanto 2003). Penerjemahan pada aplikasi yang akan dikembangkan menggunakan metode kata demi kata. Pemilihan metode ini dikarenakan Bahasa Indonesia dan Bahasa Toraja memiliki kesamaan struktur kalimat yaitu terdiri atas subjek (S), predikat (P), dan objek (O).

### **Speech Recognition**

*Speech recognition* adalah kemampuan mesin atau program untuk mengidentifikasi kata – kata dan frase dalam Bahasa lisan dan mengonversinya ke format yang dapat dibaca

mesin. Software pengenalan suara dasar memiliki kosakata yang terbatas kata – kata dan frase dan hanya dapat mengidentifikasi ini jika mereka berbicara dengan sangat jelas (Yu and Deng 2014). Pada teknologi *speech recognition* terdapat beberapa macam kata yang dijadikan input, ada beberapa macam kata yakni *Isolated word*, *connected word*, *Continuous speech*, *spontaneous speech*.

### **Bahasa Daerah Toraja**

Bahasa Toraja adalah bahasa yang digunakan oleh suku Toraja yang tersebar di Kabupaten Tana Toraja, Kabupaten Toraja Utara, dan Kabupaten Mamasa. Bahasa Toraja masih memiliki beberapa dialek di daerah Kabupaten Tana Toraja, yang dibagi atas tiga dialek, yaitu dialek Makale - Rantepao, dialek Saluputti - Bonggakaradeng, dan dialek Sillanan - Gandang batu (Sande and Bahasa 1997). Terdapat beberapa dialek dari bahasa Toraja adalah Makale (*Tallulembangna*), Rantepao (Kesu'), Toraja Barat (Toraja Barat, Mappa-Pana) (Sande and Bahasa 1984).

### **Kajian Terdahulu**

Untuk mendapatkan hasil yang optimal pada penelitian ini maka penulis melakukan kajian – kajian dari beberapa penelitian terkait atau terdahulu dan menjadi referensi dalam pembuatan penelitian ini. Adapun beberapa kajian penelitian yang telah dilakukan oleh penelitian sebelumnya, diantaranya penelitian yang dilakukan oleh Akhmad Wahyu Dani dkk tahun 2016 yaitu *Perancangan Aplikasi Voice Command Recognition Berbasis Android Dan Arduino Uno*, penelitian tersebut menjelaskan perancang system voice command recognition yang digunakan

untuk aplikasi peralatan rumah tangga yang dikendalikan oleh suara pengguna menggunakan teknologi Google Voice Recognition system, Arduino Uno, Bluetooth dan teknologi transistor untuk mengefisiensi dalam hal biaya perancangannya (Dani, Adriansyah, and Hermawan 2016). Penelitian selanjutnya dilakukan oleh Ismail Mohidin tahun 2016 yaitu *Analisis dan Perancangan Speech Recognition Translate Bahasa Inggris-Indonesia-Gorontalo Berbasis Android*, penelitian tersebut menjelaskan pembuatan aplikasi yang dirancang menggunakan sebuah smartphone sebagai perangkat keras untuk menerima input berupa kata yang diucapkan oleh user dengan metode Hidden Markov Model (HMM) yang terdapat pada library Pocketsphinx (Mohidin 2016). Penelitian selanjutnya yang dilakukan oleh Ting Wang dkk tahun 2017 yaitu *Hidden Markov modelling of sparse time series from non-volcanic tremor observations*, Penelitian tersebut menjelaskan proses merancang jenis model Hidden Markov untuk menyelidiki fenomena dimana setiap negara mewakili segmen sumber tremor yang berbeda (Wang et al. 2017). Penelitian selanjutnya yang dilakukan oleh Viney Leos-Barajas tahun 2017 yaitu *Analysis of animal accelerometer data using hidden Markov models*, penelitian tersebut menjelaskan mengenai pendekatan pembelajaran tanpa pengawasan dapat digunakan untuk menyimpulkan aspek-aspek baru dari perilaku hewan ketika variabel-variabel respons yang bermakna secara biologis digunakan, dengan peringatan bahwa negara-negara bagian mungkin tidak memetakan perilaku spesifik dengan menggunakan Hidden Markov Model (Leos-Barajas et al. 2017). Penelitian selanjutnya dilakukan oleh Clement U.Mba tahun

2018 yaitu *Condition monitoring and state classification of gearboxes using stochastic resonance and hidden Markov models*, penelitian tersebut menjelaskan mengenai sistem pendeteksian kesalahan dan klasifikasi baru berdasarkan integrasi resonansi stokastik dan pemodelan Hidden Markov (HMM) dari data getaran yang diuji dengan data getaran gearbox simulasi dan kehidupan nyata (Mba et al. 2018). Penelitian selanjutnya yang dilakukan oleh K. Orfanogiannaki dan D. Karlis tahun 2018 yaitu *Multivariate Poisson hidden Markov models with a case study of modelling seismicity*, penelitian ini menjelaskan mengenai memperluas kelas model Hidden Markov univariat yang terkenal ke kasus multidimensi, dengan memperkenalkan model Hidden Markov Markiv multivariat. Setiap keadaan model perluasan dikaitkan dengan distribusi diskrit multivarian yang berbeda (Orfanogiannaki and Karlis 2018). Penelitian selanjutnya yang dilakukan oleh Qothrun Nada dkk tahun 2019 yaitu *Speech Recognition dengan Hidden Markov Model untuk Pengenalan dan Pelafalan Huruf Hijaiyah*, penelitian tersebut menjelaskan mengenai Pengenalan ucapan dan Model Hidden Markov dilakukan untuk mengembangkan sistem antar muka mesin berbasis suara. Dalam penelitian ini juga menggunakan metode Fast Fourier Transform (FFT) untuk mengekstraksi sifat. *Hidden Markov Model* (HMM) yang digunakan dalam proses pelatihan. Juga, menghasilkan karakteristik khusus untuk setiap huruf Hijaiyah kemudian *Euclidean Distance* (ED) untuk klasifikasi akhir dalam mendeteksi pelafalan huruf Hijaiyah (Nada et al. 2019). Penelitian selanjutnya dilakukan oleh Bagus Robbiyanto dkk tahun 2019 yaitu *Perancangan Dan Analisis Sistem Speech*

*Processing Untuk Tunarungu Menggunakan Metode Hidden Markov Model Dan Mel-frequency Cepstral Coefficients*, penelitian tersebut menjelaskan mengenai mengolah sinyal suara input menjadi suatu text menggunakan metode Mel Frequency Cepstral Coefficient untuk mengekstrasi sinyal suara input dan diklasifikasi menggunakan metode Hidden Markov Model untuk melihat kemiripan antara sinyal suara yang sudah diekstrasi ciri dengan yang di database. Jika terdapat suatu kemiripan maka menghasilkan suatu text, kemudian text tersebut diolah menjadi suatu input baru yang menampilkan video Bahasa Isyarat Indonesia (Robbiyanto, Patmasari, and Magdalena 2019). Penelitian selanjutnya dilakukan oleh Ardi Mardiana dan Tantri Wahyuni tahun 2019 yaitu *Rancang Bangun Aplikasi Android Pengenalan Kosakata Untuk Disabilitas Tunarungu Menggunakan Metode Sistem Isyarat Bahasa Indonesia*, penelitian tersebut menjelaskan mengenai aplikasi media pembelajaran untuk anak tunarungu dengan menggunakan metode SIBI pada smartphone android yang dapat digunakan oleh guru tunarungu yang ada di SLBN Majalengka untuk lebih mempermudah dalam proses mengajar anak yang tunarungu dan memperlihatkan beberapa kosakata umum yang ada dalam beberapa kategori yaitu abjad, angka, benda, buah, hewan, kata ganti orang, mimik wajah, transportasi dan warna yang di sajikan dalam bentuk video dengan menggunakan metode SIBI (Mardiana and Wahyuni 2019). Selanjutnya penelitian yang mendiskusikan mengenai klasifikasi BISINDO menggunakan metode CNN dengan arsitektur mobilenetV2 yang menggunakan tensorflow yang mendapatkan hasil akurasi 54.8%(A.e and Zul

2021). Kemudian penelitian yang mengembangkan sebuah sistem penerjemah ucapan Bahasa Indonesia ke Bahasa isyarat yang menggunakan ASR (Automatic Speech Recognition) dengan akurasi 90.50%(Nasri 2020). Penelitian yang melakukan klasifikasi pada kelas kata Bahasa Madura dengan tingkat akurasi 0.96 menggunakan metode Algoritma Viterbi(Firmansyah, Adikara, and Adinugroho 2021). Kemudian penelitian yang membahas mengenai pencarian kata dalam kamus istilah komputer dan informatika yang menggunakan algoritma brute force(Rismayani et al. 2021). Selanjutnya penelitian yang mendiskusikan mengenai budaya Tana Toraja melalui pakaian adat Toraja yang menerapkan teknologi Augmented Reality(Rismayani and Layuk 2019).

## METODOLOGI

### Jenis Penelitian

Jenis Penelitian yang dilakukan pada penyusunan proposal ini ialah sebagai berikut:

1. Penelitian Lapangan (Field Research), yaitu penelitian yang langsung dilakukan di lapangan atau pada responden untuk mendapatkan data-data yang berkaitan dengan Bahasa daerah Toraja.
2. Penelitian Kepustakaan (Library Research), yaitu penelitian yang dilaksanakan menggunakan literatur (kepuustakaan), berupa buku, catatan, laporan terdahulu maupun artikel dari internet yang bersifat ilmiah yang berhubungan dengan materi pembahasan Bahasa daerah Toraja.

## Alat dan Bahan Penelitian

Adapun alat yang digunakan dalam penelitian yaitu:

- a. Perangkat Keras :
  1. Laptop Acer VX15  
Spesifikasi:
    - a. Processor Intel i7 Gen 7
    - b. GPU Nvidia Gefore GTX 1050
    - c. Ram 16GB DDR4
  2. Mouse Logitech K150
  3. Smartphone Android Xiaomi Mi 5X  
Spesifikasi:
    - a. Qualkomm Snapdragon 625
    - b. Gpu Andreno 506
    - c. Ram 4 Gb Internal 32 GB
- b. Perangkat Lunak :
  1. Windows 10 pro
  2. Microsoft Word 2010
  3. Android Studio 2017
  4. Visual Studio 2016
  5. StarUML
  6. Mozilla Firefox
  7. Os 7.1 Nougat

Adapun bahan penelitian yang digunakan sebagai berikut:

- a. Kamus Buku Bahasa Indonesia
- b. Kamus Buku Bahasa Toraja
- c. Kosa Kata Bahasa Toraja

## Teknik Pengujian

Adapun Teknik Pengujian yang digunakan antara lain:

- a. Teknik uji coba *black box* digunakan untuk menguji fungsi – fungsi khusus dari perangkat lunak yang dirancang. Pada teknik ini, kebenaran perangkat lunak yang diuji hanya dilihat berdasarkan

keluaran yang dihasilkan dari data atau kondisi masukan yang diberikan untuk fungsi yang ada tanpa melihat bagaimana proses untuk mendapatkan keluaran tersebut.

- b. Teknik uji coba *white box* digunakan untuk menguji Algoritma dari Hidden Markov Modelling. Pada teknik ini, kebenaran perangkat lunak yang diuji ialah menjamin operasi-operasi internal sesuai dengan spesifikasi yang telah ditetapkan dengan menggunakan struktur kendali dari prosedur yang dirancang.
- c. Teknik Pengujian Kuesioner digunakan untuk menguji manfaat dari sistem yang dibuat. Pada pengujian kuesioner dilakukan pada responden masyarakat sebanyak 20.

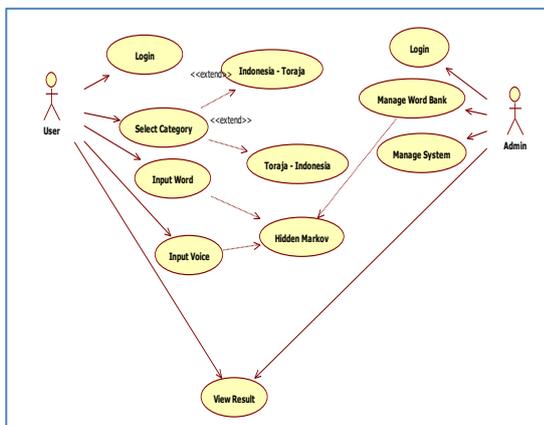
## Tahapan Penelitian

Adapun beberapa tahapan penelitian yang dilakukan penulis sebagai berikut:

- a. Pengumpulan Data: pada tahap ini penulis melakukan pengumpulan data-data yang diperoleh melalui wawancara dan juga studi pustaka yang diperoleh dari jurnal dan internet.
- b. Analisis Data dan Sistem: pada tahap ini penulis melakukan analisis terhadap data-data yang telah dikumpulkan dilapangan kemudian melakukan analisis.
- c. Desain Sistem: pada tahap ini penulis melakukan rancang sistem yaitu menggunakan teknik Unified Modelling Language (UML), setelah itu merancang form pendukung dalam sistem penerjemah yang dibuat.

- d. Coding: pada tahap ini penulis melakukan pengisian coding terhadap sistem yang telah dirancang.
- e. Pengujian Sistem: pada tahap ini ketika semua proses pembuatan sistem telah selesai maka dilakukanlah pengujian terhadap sistem yaitu menguji fungsional sistem, logika dari Hidden Markov Model yang digunakan dan juga menguji seberapa bermanfaat sistem yang dibuat dengan memberikan kuesioner kepada 20 responden.

**Gambaran Umum Sistem**



**Gambar 1.** Gambaran Umum Sistem

Gambar 1 merupakan gambaran umum sistem Penerjemah yang dimana user memilih kategori bahasa yang ingin di terjemahkan. Kemudian memilih inputan berupa teks atau suara. Jika user sudah menginput, user langsung melihat hasil terjemahan dan juga bisa mendengar hasil terjemahan. Kemudian admin melakukan kegiatan login, mengolah bank kata Bahasa daerah Toraja dan mengolah sistem.

**HASIL DAN PEMBAHASAN**

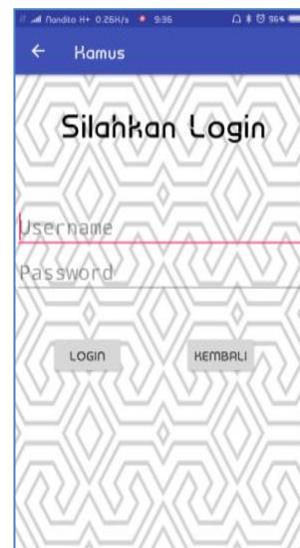
**Analisis Sistem**

Analisis sistem bertujuan untuk memperoleh informasi dan data mengenai sistem aplikasi yang

memanfaatkan *speech recognition* dalam pengiriman pesan atau informasi yang mampu mengenali sebuah kata berbahasa Indonesia yang diucapkan dan menerjemahkannya dalam bahasa Toraja ataupun sebaliknya dalam format text. Kemudian dikirimkan dan diterima sebagai pesan suara pada aplikasi *voice translator*. Dengan menggunakan aplikasi yang Akan dibuat Bertujuan agar user lebih mudah menginput kata tanpa perlu menginput pesan atau informasi secara manual. Selain itu, teknologi *speech recognition* mempermudah dalam proses pengimputan teks dengan memanfaatkan inputan suara yang Akan dikonversi menjadi teks, Sistem memiliki interface atau desain yang sederhana dengan tujuan untuk memudahkan dan dapat dimengerti oleh pengguna aplikasi (*user*).

- a. Sistem harus mampu mengenali sebuah kata berbahasa Indonesia yang diucapkan.
- b. Sistem harus bisa menerjemahkan Bahasa Indonesia yang di inputkan ke dalam Bahasa Toraja ataupun sebaliknya.

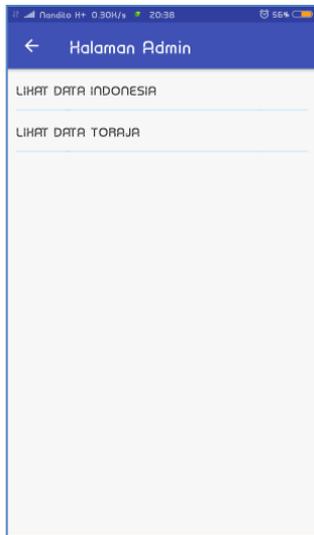
**Login Admin**



**Gambar 2.** Tampilan Login Admin

Gambar 2 merupakan tampilan dari login admin agar dapat masuk ke halaman utama dengan menginputkan username dan password.

**Halaman Utama Admin**



**Gambar 3.** Tampilan Halaman Utama Admin

Gambar 3 merupakan tampilan halaman utama pada admin yang di mana admin dapat melakukan kegiatan lihat data arti Bahasa Indonesia dan lihat data bahasa Toraja.

**Tampilan Melihat Daftar Kata Toraja**



**Gambar 4.** Tampilan Data Kata Toraja

Gambar 4 merupakan tampilan dari melihat daftar kata bahasa Toraja yang tersimpan di dalam database.

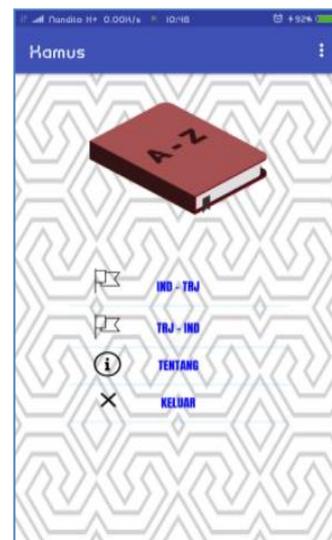
**Melihat Daftar Kata Bahasa Indonesia**



**Gambar 5.** Tampilan Lihat Data Kata Bahasa Indonesia

Gambar 5 merupakan tampilan untuk melihat daftar kata Bahasa Indonesia yang telah tersimpan dalam database.

**Halaman Utama User**



**Gambar 6.** Tampilan Halaman Utama User

Gambar 8 merupakan tampilan halamn utama dari user yang dimana terdapat menu Indonesia Toraja dan Toraja Indonesia, menu tentang sistem dan juga menu keluar dari sistem.

**Tampilan Terjemahan Indonesia-Toraja**



**Gambar 7.** Tampilan Terjemahan Indonesia-Toraja

Gambar 7 menjelaskan user dapat melakukan terjemahan Bahasa Indonesia ke Toraja dengan menginputkan kata pada kolom search yang kemudian akan muncul daftar kata yang diinginkan kemudian melakukan translate.

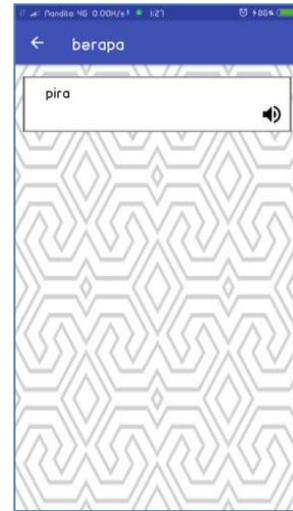
**Tampilan Terjemahan Toraja Indonesia**



**Gambar 8.** Tampilan Terjemahan Toraja Indonesia

Gambar 8 menjelaskan user dapat melakukan terjemahan bahas Toraja ke Indonesia dengan menginputkan kata pada kolom search yang kemudian muncul daftar kata yang diinginkan kemudian melakukan translate.

**Halaman Definisi**



**Gambar 9.** Halaman Definisi

Gambar 9 menjelaskan user dapat melihat arti terjemahan yang telah dilakukan sebelumnya pada halaman terjemahan dalam bentuk suara.

**Pengujian Sistem**

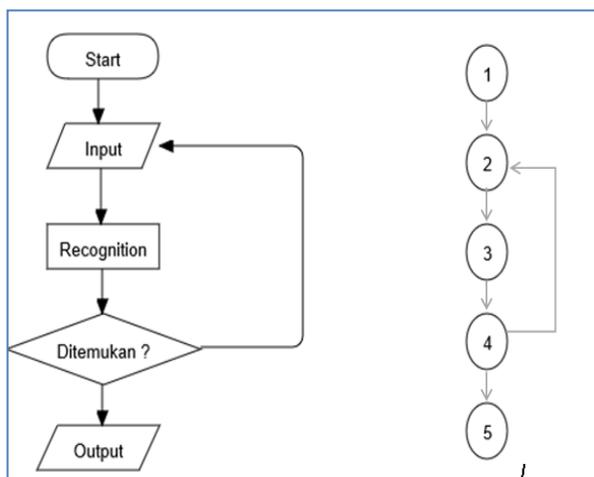
1. Pengujian *Blackbox*, pengujian *blacbox* adalah pengujian yang dilakukan hanya mengamati hasil eksekusi melalui data uji dan memeriksa fungsional dari perangkat lunak(Simarmata 2010).

**Tabel 1.** Pengujian *Blackbox*

No	Test Factor	Skenario	Hasil
1	Login Admin	Berhasil Login admin kedalam Sistem	Valid
2	Melihat Data Kata Indonesia	Berhasil Menampilkan Data Kata Bahasa Indonesia	Valid

No	Test Factor	Skenario	Hasil
3	Melihat Data Kata Toraja	Berhasil Menampilkan Data Kata Bahasa Toraja	Valid
4	Penerjemahan Kata Indonesia - Toraja	Berhasil Menejemahkan Kata Bahasa Indonesia Ke Toraja	Valid
5	Penerjemahan Toraja ke Bahasa Indonesia	Berhasil Menerjemahkan Bahasa Toraja ke Bahasa Indonesia	Valid
6	Proses Defenisi	Berhasil Melakukan proses Defenisi dari kata	Valid
7	Pengujian Audio	Berhasil Mengeluarkan suara hasil terjemahan ke sistem	Valid

2. Pengujian *Whitebox*, pengujian *whitebox* adalah pengujian yang didasarkan pada pengecekan terhadap detail perancangan, menggunakan struktur kontrol dari desain program secara procedural untuk membagi pengujian ke dalam beberapa kasus pengujian. Secara sekilas dapat diambil kesimpulan *white box testing* merupakan petunjuk untuk mendapatkan program yang benar secara 100% (Asnawati 2015).



Gambar 12. *Speech Recognition*

Diketahui :

$$N = 5; E = 5; R = 2$$

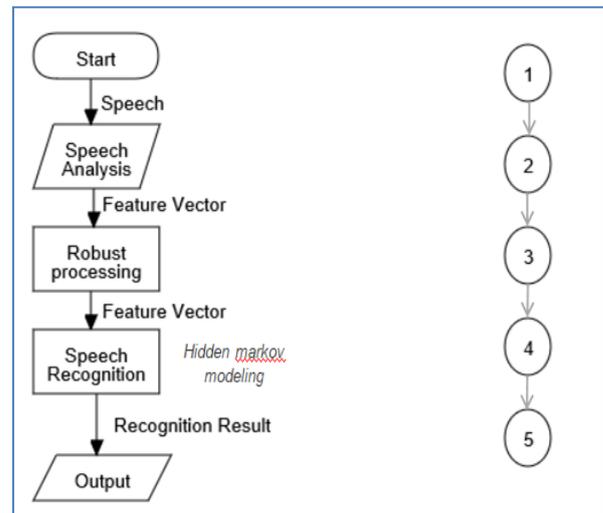
Penyelesaian:

$$CC = (5-5) + 2 = 2$$

$$\text{Indepenth Path} = 2$$

$$\text{Path 1} = 1-2-3-4-5$$

$$\text{Path 2} = 1-2-3-4-2-3-4-5$$



Gambar 13. *Hidden Markov Modelling*

Diketahui :

$$N = 5; E = 4; R = 1$$

Penyelesaian:

$$CC = (5-4) + 2 = 1$$

$$\text{Indepenth Path} = 1$$

$$\text{Path 1} = 1-2-3-4-5$$

3. Kuesioner, Pengujian kelayakan sistem digunakan untuk mengetahui respon pengguna terhadap aplikasi yang dibangun. Pengujian ini dilakukan dengan metode kuisisioner (angket). Teknik kuisisioner digunakan untuk mengumpulkan data yang dibutuhkan dari sejumlah pertanyaan secara tertulis yang diajukan kepada responden yang mendapat bimbingan maupun petunjuk dari peneliti.

Adapun indikator yang menjadi penilaian dalam pengujian ini yakni sebagai berikut:

- a. Kemudahan dalam penggunaan aplikasi
- b. Ketertarikan pengguna terhadap aplikasi
- c. Fungsionalitas aplikasi
- d. Kemanfaatan aplikasi
- e. Rekomendasi pengguna

Instrument yang digunakan dalam penelitian ini berupa angket dengan mengajukan sejumlah pertanyaan kepada responden dengan berpedoman pada indikator yang telah ditetapkan.

Menggunakan skala ordinal pada item-item pertanyaan, dimana setiap alternatif jawaban mengandung perbedaan nilai. Berikut ini adalah hasil kuisisioner yang dibagikan kepada 20 responden dengan 10 pertanyaan.

Adapun perhitungan dari hasil persentase kuisisioner sebagai berikut:

Diketahui:

$$\text{Persentase} = \frac{a}{b} * 100\% \dots \dots \dots (1)$$

Dimana: a = Total nilai Koresponden ke-i, dimana i=1-20

b = Nilai Tetapan 50

Keterangan:

- SS = 5
- S = 4
- CS = 3
- KS = 2
- TS = 1

**Tabel 2.** Nilai Persentase dari Bobot Koresponden

No Koresponden	Jumlah Nilai	Hasil persentase
1	38	76
2	42	84
3	42	84
4	44	88
5	48	96
6	46	92
7	40	80
8	34	68
9	38	76

No Koresponden	Jumlah Nilai	Hasil persentase
10	46	92
11	46	92
12	40	80
13	44	88
14	34	68
15	37	74
16	37	74
17	34	68
18	35	70
19	38	76
20	39	78
<b>Total</b>	<b>802</b>	<b>1604</b>

Jadi total persentase dari nilai bobot koresponden adalah:

$$\text{Total persentase} = \sum \frac{x}{y} * 100 \% \dots \dots \dots (2)$$

Diketahui: x = 1604, y = 20

Dimana: x = Total Hasil Persen

y = Jumlah koresponden

$$\begin{aligned} \text{Total persentase} &= \frac{1604}{20} \% \\ &= 80.2 \% \end{aligned}$$

## PENUTUP

Kesimpulan penelitian adalah dibangun sebuah sistem Speech Recognition penerjemah untuk bahasa daerah Toraja yang dapat digunakan masyarakat umum yang belum mengetahui arti dan pengucapan pada Bahasa daerah Toraja serta sebagai sarana komunikasi. Berdasarkan pengujian fungsional sistem hasil test factor dinyatakan valid dan berdasarkan pengujian logika penggunaan Hidden Markov Modelling dinyatakan bebas dari kesalah logika serta berdasarkan hasil kuisisioner yang diperoleh dari 20 responden yang terdiri dari 10 pertanyaan menyatakan sekitar 80,2 % sangat setuju bahwa sistem tersebut sangat bermanfaat. Adapun saran untuk pengembangan sistem diharapkan

kedepannya dapat menggunakan model lainnya ataupun metode algoritma yang relevan.

## UCAPAN TERIMA KASIH

Terima Kasih penulis ucapkan kepada DRPM Ristek DIKTI yang telah memberikan Dana kepada peneliti sehingga peneliti dapat menyelesaikan penelitian ini dengan baik, Terima kasih juga penulis ucapkan kepada para responden yang bersedia di wawancarai dan juga melakukan pengisian kuesioner untuk penyelesaian penelitian kami dan semua pihak terkait yang telah membantu baik secara langsung maupun tidak langsung.

## DAFTAR PUSTAKA

- A.e, Nasha Hikmatia, and Muhammad Ihsan Zul. 2021. "Aplikasi Penerjemah Bahasa Isyarat Indonesia Menjadi Suara Berbasis Android Menggunakan Tensorflow." *Jurnal Komputer Terapan* 7(1):74–83. doi: 10.35143/jkt.v7i1.4629.
- Asnawati, Feri Hari Utami dan. 2015. *Rekayasa Perangkat Lunak*. Deepublish.
- Dani, Akhmad Wahyu, Andi Adriansyah, and Dodi Hermawan. 2016. "Perancangan Aplikasi Voice Command Recognition Berbasis Android Dan Arduino Uno." *Jurnal Teknologi Elektro* 7(1).
- Firmansyah, Ilham, Putra Pandu Adikara, and Sigit Adinugroho. 2021. "Klasifikasi Kelas Kata (Part-Of-Speech Tagging) untuk Bahasa Madura Menggunakan Algoritme Viterbi." *Jurnal Teknologi Informasi dan Ilmu Komputer* 8(5):1039–48. doi: 10.25126/jtiik.2021854483.
- Hariyanto, Zuchridin Suryawinata &. Sugeng. 2003. *Translation Bahasan Teori & Penuntun Praktis Menerjemah*. Kanisius.
- Leos-Barajas, Vianey, Theoni Photopoulou, Roland Langrock, Toby A. Patterson, Yuuki Y. Watanabe, Megan Murgatroyd, and Yannis P. Papastamatiou. 2017. "Analysis of Animal Accelerometer Data Using Hidden Markov Models." *Methods in Ecology and Evolution* 8(2):161–73. doi: 10.1111/2041-210X.12657.
- Mardiana, Ardi, and Tantri Wahyuni. 2019. "Rancang Bangun Aplikasi Android Pengenalan Kosakata Untuk Disabilitas Tunarungu Menggunakan Metode Sistem Isyarat Bahasa Indonesia." *INFOTECH journal* 5(1):64–68.
- Mba, Clement U., Viliam Makis, Stefano Marchesiello, Alessandro Fasana, and Luigi Garibaldi. 2018. "Condition Monitoring and State Classification of Gearboxes Using Stochastic Resonance and Hidden Markov Models." *Measurement* 126:76–95. doi: 10.1016/j.measurement.2018.05.038.
- Mohidin, Ismail. 2016. "Analisis Dan Perancangan Speech Recognition Translate Bahasa Inggris-Indonesia-Gorontalo Berbasis Android." *ENERGY* 6(1):12–17.
- Nada, Qothrun, Cahya Ridhuandi, Puji Santoso, and Dwi Apriyanto. 2019. "Speech Recognition Dengan Hidden Markov Model Untuk Pengenalan Dan Pelafalan Huruf Hijaiyah." *JURNAL AL-AZHAR INDONESIA SERI SAINS DAN TEKNOLOGI* 5(1):19–26.
- Nasri, Andi. 2020. "Konversi Suara Ucapan Bahasa Indonesia Ke Sistem Bahasa Isyarat Indonesia (Sibi)." *Ainet : Jurnal Informatika* 2(2):7–13. doi: 10.26618/ainet.v2i2.4025.
- Orfanogiannaki, K., and D. Karlis. 2018. "Multivariate Poisson Hidden Markov Models with a Case Study of Modelling Seismicity." *Australian & New Zealand Journal of Statistics* 60(3):301–22. doi: 10.1111/anzs.12242.
- Rismayani, Nfn, and Novita Sambo Layuk. 2019. "Implementation of Augmented Reality Technology for the Mobile-Based

- Introduction of Toraja Traditional Clothes.” *Jurnal Pekommas* 4(2):155–64. doi: 10.30818/jpkm.2019.2040206.
- Rismayani, Rismayani, Novita Sambo Layuk, Sri Wahyuni, Hartana Wali, and Ni Kadek Marselina. 2021. “Pencarian Kata Pada Aplikasi Kamus Istilah Komputer Dan Informatika Menggunakan Algoritma Brute Force Berbasis Android.” *Komputika: Jurnal Sistem Komputer* 10(1):43–52. doi: 10.34010/komputika.v10i1.3644.
- Robbiyanto, Bagus, Raditiana Patmasari, and Rita Magdalena. 2019. “Perancangan Dan Analisis Sistem Speech Processing Untuk Tunarungu Menggunakan Metode Hidden Markov Model Dan Mel-frequency Cepstral Coefficients.” *eProceedings of Engineering* 6(1).
- Sande, J. S., and Pusat Pembinaan dan Pengembangan Bahasa. 1984. *Struktur bahasa Toraja Sa'dan*. Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- Sande, J. S., and Pusat Pembinaan dan Pengembangan Bahasa. 1997. *Tata bahasa Toraja*. Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- Wang, Ting, Jiancang Zhuang, Kazushige Obara, and Hiroshi Tsuruoka. 2017. “Hidden Markov Modelling of Sparse Time Series from Non-Volcanic Tremor Observations.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 66(4):691–715. doi: 10.1111/rssc.12194.
- Wuryantoro, Aris. 2018. *Pengantar Penerjemahan*. Deepublish.
- Yu, Dong, and Li Deng. 2014. *Automatic Speech Recognition: A Deep Learning Approach*. Springer.
- Zucchini, Walter, and Iain L. MacDonald. 2009. *Hidden Markov Models for Time Series: An Introduction Using R* - Walter Zucchini, Iain L. MacDonald - Google Buku. Boca Raton London: CRC Press.

