

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**SPEECH-TO-TEXT BAHASA MINANG (DIALEK AGAM)
MENGUNAKAN CMU SPHINX****TUGAS AKHIR**

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

SYAHRUL YUSMAN

NIM. 12050112675



UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU**

2024

LEMBAR PERSETUJUAN

SPEECH-TO-TEXT BAHASA MINANG (DIALEK AGAM) MENGUNAKAN CMU SPHINX

TUGAS AKHIR

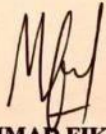
Oleh

SYAHRUL YUSMAN

NIM. 12050112675

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 14 November 2024

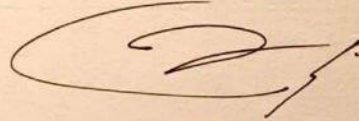
Pembimbing I,



MUHAMMAD FIKRY, S.T., M.SC.

NIP. 198010182007101002

Pembimbing II



YUSRA, ST, M.T

NIP. 198401232015032001

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

**SPEECH-TO-TEXT BAHASA MINANG (DIALEK AGAM)
MENGUNAKAN CMU SPHINX**

Oleh

SYAHRUL YUSMAN

NIM. 12050112675

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 14 November 2024

Mengesahkan,

Ketua Jurusan,



Dr. Hartono, M.Pd.

NIP. 196403011992031003

Iwan Iskandar, ST., MT.

NIP. 198212162015031003

DEWAN PENGUJI

- | | |
|---------------|---------------------------------|
| Ketua | : Dr. Rahmad Abdillah, ST., M.T |
| Pembimbing I | : Muhammad Fikry, S.T., M.Sc |
| Pembimbing II | : Yusra, S.T., M.T. |
| Penguji I | : Muhammad Irsyad, ST., M.T |
| Penguji II | : Muhammad Affandes, ST, M.T. |



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lampiran Surat :
 Nomor : Nomor 25/2021
 Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : Syahrul Yusman
 NIM : 17050112675
 Tempat/Tgl. Lahir : Pekanbaru, 7 November 2000
 Fakultas/Pascasarjana : Sains dan Teknologi
 Prodi : Teknik Informatika
 Judul Disertasi/Thesis (Skripsi) Karya Ilmiah lainnya* :
 Speech - to - text Bahasa ~~Minang~~ Minang Dialek Acah Manggungan
 CAHU SPHINX

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis (Skripsi) Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis (Skripsi) Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis (Skripsi) (Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 14 November 2024
 Yang membuat pernyataan



Syahrul Yusman
 NIM : 17050112675

* pilih salah satu sesuai jenis karya tulis



LEMBAR PERSEMBAHAN

Bismillahirrahmanirrahim

Alhamdulillahirabbil'alamin

Segala puji bagi-Mu Ya Allah, karena berkat Rahmat dan karunia-Mu aku dapat berjuang dalam menyelesaikan Tugas Akhir ini

Untuk kedua orang tuaku

Terima kasih aku ucapkan atas semua doa dan dukungan yang telah kalian berikan sehingga aku dapat semangat dalam menyelesaikan Tugas Akhir ini
jasa kalian tidak akan aku lupakan

Untuk abang dan adikku

Terima kasih juga atas dukungan dan doa selama aku berjuang dalam menyelesaikan tugas akhir ini

Terima kasih kepada dosen pembimbing dan para dosen yang telah membantu saya dalam menyelesaikan tugas akhir ini

Terima kasih juga teman-teman dari kelas TIF'I Angkatan 2020 yang telah mendukungku dalam menyelesaikan tugas akhir ini yang namanya tidak bisa disebutkan satu persatu.

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



ABSTRAK

Indonesia memiliki keragaman bahasa daerah yang sangat luas, tersebar dari Sabang hingga Merauke dengan karakteristik unik. Penelitian ini menyoroti dialek Minang Agam, khususnya subdialek Tilatang Kamang. Fokus utama adalah penggunaan teknologi *Speech-to-Text* (STT) untuk mengkonversi input suara menjadi teks, dengan tujuan mengidentifikasi penulisan yang benar dari input suara. CMU Sphinx dipilih sebagai alat konversi karena kemampuannya menangani variasi dialek. Pengujian akurasi dilakukan dengan menggunakan *Word Error Rate* (WER), sebuah metrik penting untuk mengukur tingkat kesalahan dalam pengenalan kata. Hasilnya menunjukkan WER sebesar 27.2%, yang berarti terdapat 228 kata yang salah dari total 838 kata yang diinputkan, menunjukkan bahwa akurasi yang dihasilkan cukup memuaskan. Meskipun demikian, teknologi STT ini berhasil menjalankan tugasnya dan menampilkan hasil teks dari input suara secara real-time.

Kata kunci : Minang, CMU Sphinx, *Word Error Rate*.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



ABSTRACT

Indonesia has a vast diversity of regional languages, spread from Sabang to Merauke, each with unique characteristics. This study highlights the Minang Agam dialect, specifically the Tilatang Kamang subdialect. The main focus is the use of Speech-to-Text (STT) technology to convert spoken input into text, with the aim of identifying the correct spelling from the spoken input. The CMU Sphinx toolkit was chosen as the conversion tool due to its ability to handle dialect variations. Accuracy testing was conducted using Word Error Rate (WER), an important metric for measuring the error rate in word recognition. The results showed a WER of 27.2%, which means there were 228 incorrect words out of a total of 838 words inputted, indicating that the accuracy produced was quite satisfactory. Nevertheless, this STT technology succeeded in carrying out its task and displaying text results from voice input in real-time.

Keywords: Minang, CMU Sphinx, Word Error Rate.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Assalammu 'alaikum wa rohmatullohi wa barokatuh.

Ahamdulillahi robbil'alamin, tak henti-hentinya kami ucapkan kehadiran Allah *Subhanahu wa ta'ala*, yang dengan rahmat dan hidayah-Nya kami mampu menyelesaikan Tugas Akhir ini dengan baik. Tidak lupa bershalawat kepada Nabi dan Rasul-Nya, Nabi Muhammad *Sholallohu 'alaihi wa salam*, yang telah membimbing kita sebagai umatnya menuju jalan kebaikan.

Tugas Akhir ini disusun sebagai salah satu syarat untuk mendapatkan gelar sarjana pada jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Banyak sekali pihak yang telah membantu kami dalam penyusunan laporan ini, baik berupa bantuan materi ataupun berupa motivasi dan dukungan kepada kami. Semua itu tentu terlalu banyak bagi kami untuk membalasnya, namun pada kesempatan ini kami hanya dapat mengucapkan terima kasih kepada:

1. Bapak Prof Dr Khairunnas Rajab, M.Ag selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Iwan Iskandar, ST., MT. selaku Ketua Prodi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Febi Yanto, S.Kom., M.Kom. selaku Penasehat Akademik. Selama menjalani perkuliahan di Prodi Teknik Informatika.
5. Bapak Muhammad Fikry, S.T., M.Sc. selaku Dosen Pembimbing 1 Tugas Akhir, yang telah meluangkan waktunya, memberikan wawasan, motivasi dan ilmu yang bermanfaat sehingga Tugas Akhir ini dapat diselesaikan dengan baik.
6. Ibuk Yusra, S.T., M.T. selaku Dosen Pembimbing 2 Tugas Akhir, yang telah meluangkan waktunya, memberikan wawasan, motivasi dan ilmu yang bermanfaat sehingga Tugas Akhir ini dapat diselesaikan dengan baik.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

7. Bapak Muhammad Irsyad, MT. selaku Dosen Penguji 1 Tugas Akhir, yang membantu memberikan masukan dalam penyusunan Tugas Akhir ini.
8. Bapak Reski Mai Candra, S.T., M.Sc. selaku Dosen Penguji 2 Tugas Akhir, yang telah banyak meluangkan waktu, memberikan kritik dan saran dalam penulisan dan perbaikan laporan Tugas Akhir ini.
9. Bapak Nursal dan Ibu Yusniar selaku Orang Tua yang telah memberikan doa dan dukungan.
10. Teman-teman TIF'I angkatan 2020
11. Seluruh pihak yang belum kami cantumkan, terima kasih atas dukungannya, baik material maupun spiritual.

Kami menyadari bahwa dalam penulisan laporan ini masih banyak kesalahan dan kekurangan, oleh karena itu kritik dan saran yang sifatnya membangun sangat kami harapkan untuk kesempurnaan laporan ini. Akhirnya kami berharap semoga laporan ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya.

Wassalamu'alaikum wa rohmatullohi wa barokatuh.

Pekanbaru, 14 November 2024

Penulis

UIN SUSKA RIAU

DAFTAR ISI

COVER.....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN	v
LEMBAR PERSEMBAHAN	vi
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL.....	xvi
DAFTAR RUMUS	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB 2 KAJIAN PUSTAKA.....	6
2.1 Landasan Teori.....	6
2.1.1 Bahasa	6
2.1.2 Speech-to-Text (STT)	6
2.1.3 CMU Sphinx	6

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.1.4	<i>Mel Frequency Cepstral Coefficient (MFCC)</i>	7
2.1.5	<i>Hidden Markov Model (HMM)</i>	8
2.1.5	<i>Word Error Rate (WER)</i>	9
2.2	Penelitian Terkait	10
BAB 3	METODOLOGI PENELITIAN	21
3.1	Identifikasi Masalah	21
3.2	Pengumpulan Dataset	21
2.1.6	Pembuatan Teks	22
2.1.7	Perekaman Audio	22
3.3	Penyiapan Dataset	22
3.4	Pelatihan Dataset	23
3.5	Pengujian	23
BAB 4	PEMBAHASAN.....	24
4.1	Pembuatan Dataset	24
4.2	Analisa Metode.....	29
4.3	Implementasi	31
2.1.1	Tahap pelatihan	32
2.1.2	Menjalankan Model Akustik.....	45
4.4	Pengujian	46
4.5	Analisa Hasil Pengujian	51
BAB 5	PENUTUP	52
5.1	Kesimpulan	52
5.2	Saran.....	52
DAFTAR PUSTAKA	54
LAMPIRAN A	58



LAMPIRAN B	79
LAMPIRAN C	88
LAMPIRAN D	109
LAMPIRAN E DOKUMENTASI PENELITIAN	127



UIN SUSKA RIAU

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR GAMBAR

Gambar 2. 1 Flowchart MFCC	7
Gambar 2. 2 Flowchart HMM.....	8
Gambar 3. 1 Metode Penelitian.....	21
Gambar 4. 1 Pengolahan Data Kata Untuk CMU Sphinx.....	25
Gambar 4. 2 File Agam.txt.....	26
Gambar 4. 3 Folder Hasil Rekaman.....	27
Gambar 4. 4 Menghilangkan Noise Audio	27
Gambar 4. 5 Sebelum Dilakukan Penghilangan Noise	28
Gambar 4. 6 Setelah Dilakukan Penghilangan Noise	28
Gambar 4. 7 Proses Pemotongan Audio	29
Gambar 4. 8 Analisa CMU Sphinx	29
Gambar 4. 9 Alur CMU Sphinx	31
Gambar 4. 10 Pembuatan Model Bahasa dan Tata Bahasa.....	32
Gambar 4. 11 Hasil File Agam.txt Yang Telah Diproses	33
Gambar 4. 12 File agam.dic	33
Gambar 4. 13 File agam.lm.DMP	34
Gambar 4. 14 File agam.filler	34
Gambar 4. 15 File agam.phone	35
Gambar 4. 16 File agam_test.fileids.	35
Gambar 4. 17 File agam_test.transcription.	36
Gambar 4. 18 File agam_train.fileids.....	36
Gambar 4. 19 File agam_train.transcription	37
Gambar 4. 20 Menampilkan Python dan Perl yang Telah Diinstal.....	37
Gambar 4. 21 Pocketsphinx	38
Gambar 4. 22 Sphinxbase	38
Gambar 4. 23 Sphinxtrain	39
Gambar 4. 24 Folder Agam.....	39
Gambar 4. 25 Sub folder etc	40
Gambar 4. 26 Sub Folder wav	40
Gambar 4. 27 Setup Database	41

© Hak cipta dilindungi Undang-Undang
 UIN Suska Riau
 State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

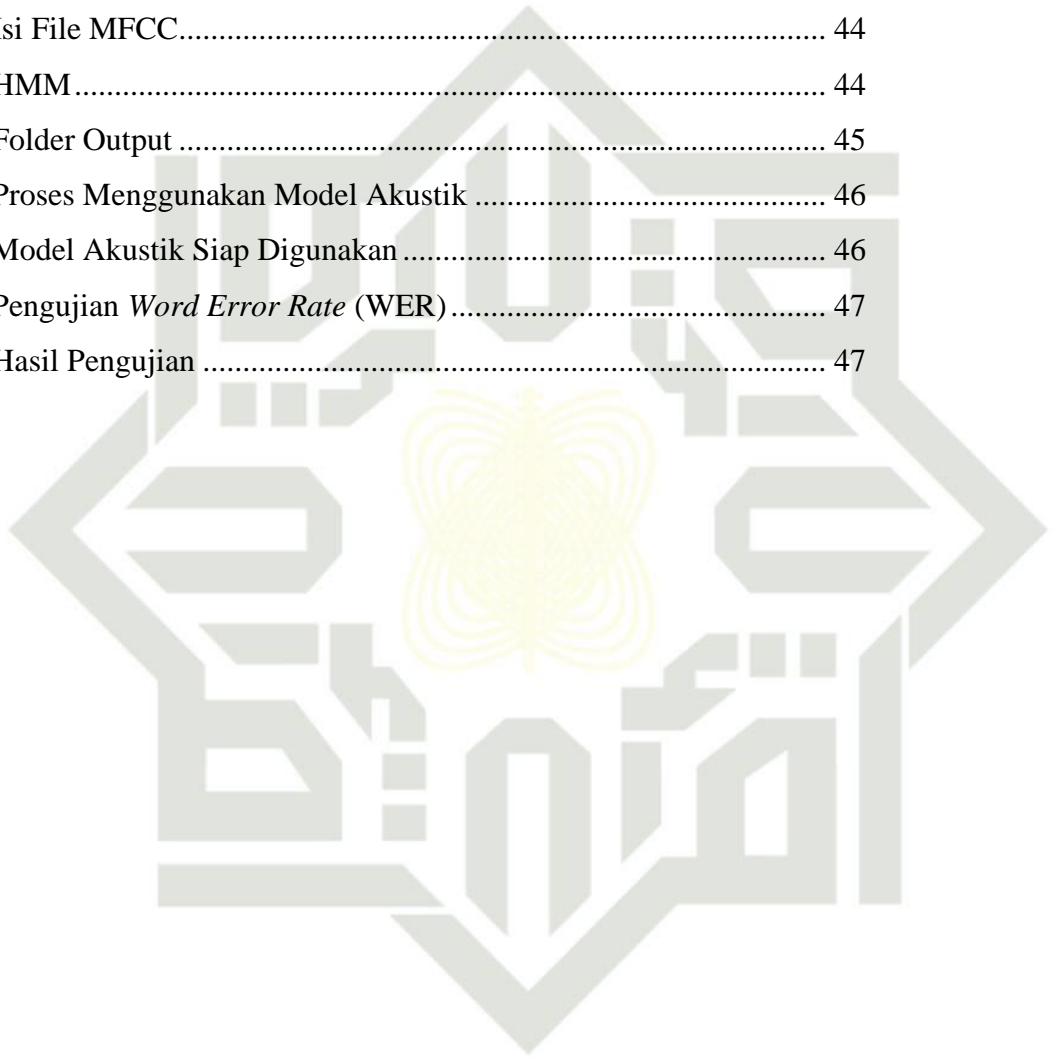
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Gambar 4. 28 Konfigurasi Baris 165 File sphinx_train.cfg.....	41
Gambar 4. 29 Konfigurasi Baris 242 File sphinx_train.cfg.....	42
Gambar 4. 30 Perintah Pelatihan Database.....	42
Gambar 4. 31 000 Computing Feature From Audio Files.....	43
Gambar 4. 32 File Slave_Feat_Pl.....	43
Gambar 4. 33 Isi File MFCC.....	44
Gambar 4. 34 HMM.....	44
Gambar 4. 35 Folder Output.....	45
Gambar 4. 36 Proses Menggunakan Model Akustik.....	46
Gambar 4. 37 Model Akustik Siap Digunakan.....	46
Gambar 4. 38 Pengujian <i>Word Error Rate</i> (WER).....	47
Gambar 4. 39 Hasil Pengujian.....	47



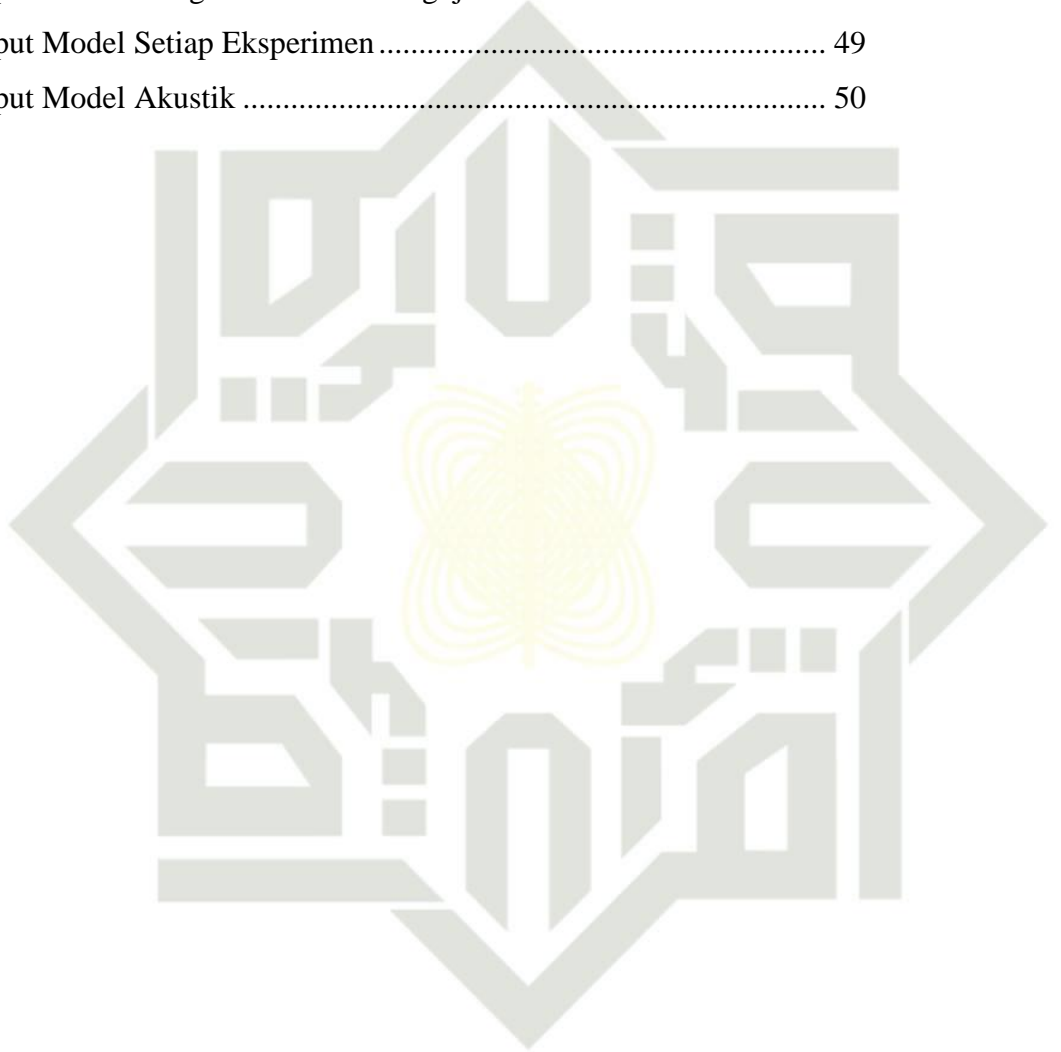


DAFTAR TABEL

Tabel 2. 1 Skala WER.....	10
Tabel 2. 2 Penelitian Terkait	10
Tabel 4. 1 Data teks bahasa Minang dan terjemahan.....	24
Tabel 4. 2 Eksperimen Meningkatkan Hasil Pengujian WER.....	48
Tabel 4. 3 Output Model Setiap Eksperimen.....	49
Tabel 4. 4 Output Model Akustik	50

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





DAFTAR RUMUS

(1) Rumus Satu.....	7
(2) Rumus Dua.....	7
(3) Rumus Tiga	10



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara besar yang kaya akan keberagaman yang tersebar dari sabang sampai merauke. Keberagaman ini meliputi berbagai elemen seperti suku, agama, budaya dan bahasa. Karena keragaman tersebut, Indonesia menjadi salah satu negara dengan bahasa daerah terbanyak di dunia. Menurut (Ratumanan et al., 2022) Indonesia memiliki 746 bahasa daerah selain bahasa nasional, yaitu bahasa Indonesia.

Bahasa daerah adalah bahasa yang digunakan orang di masyarakat suatu daerah untuk berkomunikasi dan berinteraksi dalam kehidupan sehari-hari. Salah satu bahasa daerah tersebut adalah bahasa Minang (Kharisma et al., 2021) yang biasanya digunakan oleh masyarakat di beberapa daerah yang ada di sekitar Sumatera Barat. Meskipun banyak yang berkomunikasi menggunakan bahasa Minang, setiap daerah memiliki karakteristik yang berbeda-beda (Ayub et al., 1993). Secara tradisional, bahasa Minang terbagi menjadi empat macam dialek, yakni: Dialek 50 kota, Dialek Tanah Datar, Dialek Agam, dan Dialek Pesisir.

Pada penelitian ini, penulis mengangkat topik menggunakan Dialek Agam karena dialek agam merupakan dialek yang paling mudah dipahami dibandingkan dialek lainnya. Penyebabnya adalah karena dialek Agam merupakan bentuk standar dari bahasa Minangkabau, dan dialek standar tersebut berperan sebagai penengah antara berbagai dialek yang ada dalam suatu bahasa. Hal ini berhubungan juga dengan dengan fungsi wilayah Agam, sebagai pusat berbagai kegiatan masyarakat yang ditandai dengan adanya pusat pendidikan, pusat perdagangan, pusat kebudayaan, pusat pemerintahan, dan pusat penyebaran agama Islam. Dan juga masyarakat Minangkabau berkepentingan dengan Bukittinggi, dan bahasa Minang yang digunakan adalah dialek Agam (Ayub et al., 1993). Dalam pengucapannya dialek Agam tidak terlalu berirama, yang berbeda dengan dialek Lima Puluh Kota

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

seperti Payakumbuh. Sub-dialek Tilatang Kamang juga memiliki kosakata khas seperti *ano* yang berarti dia, *cako* yang berarti tadi, *miang* yang berarti saja.

Penggunaan *Speech-to-Text* (STT) adalah salah satu kemajuan teknologi saat ini. STT memiliki banyak manfaat, seperti mengubah suara menjadi teks dalam percakapan, pidato, presentasi, dan juga membantu penyandang disabilitas yang mengalami kesulitan mengetik atau mendengar. Selain itu, teknologi STT membantu pelatihan model kecerdasan buatan berbasis data teks. Dalam kehidupan sehari-hari, teknologi ini sangat membantu, baik dalam bekerja, berkomunikasi, atau mengakses informasi. Walaupun pengembangan teknologi *Speech-to-Text* telah berkembang pesat, pengembangan teknologi *Speech-to-Text* menggunakan bahasa daerah seperti Bahasa Minang kurang dukungan untuk perkembangan teknologi yang ada. Oleh karena itu, diperlukan adanya pengembangan dan penelitian lebih lanjut di bidang ini sehingga bermanfaat bagi masyarakat umum.

Salah satu toolkit dalam *Speech-to-Text* (STT) yang cukup populer adalah CMU Sphinx, yang dikembangkan di *Carnegie Mellon University*. CMU Sphinx merupakan toolkit *Speech-to-Text* (STT) yang dapat digunakan dalam berbagai bahasa, seperti bahasa Inggris, Mandarin, dan beberapa bahasa lainnya (Lamere et al., 2003). CMU Sphinx dikenal fleksibel dan cocok untuk model pengenalan suara yang digunakan pada perangkat dengan sumber daya terbatas. Teknologi ini memanfaatkan model akustik untuk mengoptimalkan penggunaan memori dan menggunakan teknik decoding yang efektif dalam mempercepat proses pengenalan suara tanpa memakan banyak memori. CMU Sphinx memiliki tiga komponen utama, yaitu Pocketsphinx untuk pengenalan suara, Sphinxtrain untuk pelatihan model akustik, dan Sphinxbase sebagai pustaka pendukung. Dalam proses pengenalan suara, CMU Sphinx menggunakan *Mel-Frequency Cepstral Coefisien* (MFCC) untuk mengekstrak fitur dari sinyal suara. MFCC membantu dalam mengidentifikasi komponen frekuensi penting dan memberikan representasi dari karakteristik suara, sehingga mempercepat proses pengenalan suara. Selain itu, *Hidden Markov Models* (HMM) digunakan untuk memodelkan urutan suara yang dihasilkan selama pengucapan. HMM menghitung probabilitas transisi fitur suara

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tertentu, memungkinkan sistem untuk menangani variasi dalam pengucapan dan mendekode urutan fitur MFCC untuk menentukan kata atau frasa yang paling mungkin diucapkan.

Penelitian terkait *Speech-to-Text* (STT) telah dilakukan, seperti oleh Channareth Srun, Virbora Ny, Chea Cheat, Sokheang Ching, Piseth Ny (*Srun et al., 2021*) yang berjudul “*Development of Speech Recognition System Based on CMU Sphinx for Khmer Language*” yang mana penelitian ini Penelitian ini membahas *Automatic Speech Recognition* (ASR) untuk bahasa Khmer menggunakan toolkit CMUSphinx, melibatkan rekaman dari 85 pembicara dan 157 kata Khmer, mencapai akurasi Word Error Rate sebesar 10.1% dan akurasi pengenalan kata hingga 10.0%, Penelitian ini juga mencatat tantangan dalam membangun sistem ASR untuk bahasa baru, serta pentingnya kualitas data pelatihan dan akurasi model. Sistem ini dirancang dengan kebutuhan sumber daya yang rendah, sehingga cocok digunakan untuk aplikasi. Penelitian lainnya adalah “*Building CMU Sphinx language model for Quran using simplified Arabic phonemes*” dilakukan oleh (El Amrani et al., 2016) yang menunjukkan bahwa dengan data pelatihan yang cukup, akurasi pengenalan suara dapat mencapai Word Error Rate (WER) sekitar 55,7%. Penelitian lain yang Berjudul “*Eksperimen Pengenalan Ucapan Aksara Lampung Dengan CMU Sphinx 4*” dilakukan oleh Sri Ningsih, Saniati (Ningsih, 2018) yang mana Penelitian ini mengeksplorasi pengenalan ucapan aksara Lampung menggunakan CMU Sphinx 4, dengan fokus pada akurasi pengenalan suku kata, kata, dan frasa. Hasil eksperimen menunjukkan bahwa akurasi tertinggi dicapai pada pengenalan kata (74,03%).

Dari penerapan teknologi *Speech-to-Text* (STT) ini diharapkan penelitian ini memberikan kontribusi terhadap pengembangan teknologi STT untuk Bahasa Minang, khususnya Dialek Agam, sehingga lebih banyak orang yang mendukung pelestarian bahasa daerah di tengah kemajuan teknologi. Selain itu, teknologi ini dapat membantu masyarakat dalam berbagai kegiatan sehari-hari, terutama dalam komunikasi dan akses informasi menjadi lebih mudah.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah yang dapat diambil adalah bagaimana menerapkan model STT untuk bahasa Minang dialek Agam menggunakan CMU Sphinx untuk menghasilkan teks dari suara dengan tepat dan akurat.

1.3 Batasan Masalah

Agar penelitian ini tidak menyimpang dari rumusan masalah, maka penelitian ini mempunyai batasan masalah sebagai berikut:

1. Sumber data didapat dari internet dan hasil wawancara dengan narasumber perempuan yang merupakan warga asli Sumatera Barat yang bisa berbahasa Minang dialek Agam dengan sub dialek Tilatang Kamang.
2. Pengembangan model STT ini hanya berfokus ke dialek Agam dengan sub dialek Tilatang Kamang.
3. Penelitian ini menggunakan dataset yang terbatas.

1.4 Tujuan Penelitian

Tujuan yang dicapai dalam penelitian ini adalah:

1. Membuat dataset suara Bahasa Minang dialek Agam dengan sub dialek Tilatang Kamang.
Menerapkan teknologi *Speech-to-Text* (STT) menggunakan CMU Sphinx sehingga dapat mengenali dan mentranskripsikan ucapan ke dalam teks bahasa Minang dialek Agam dengan sub dialek Tilatang Kamang.
Pengujian pada penelitian ini menggunakan WER (*Word Error Rate*).

1.5 Manfaat Penelitian

Manfaat yang didapat dari penelitian ini sebagai berikut:

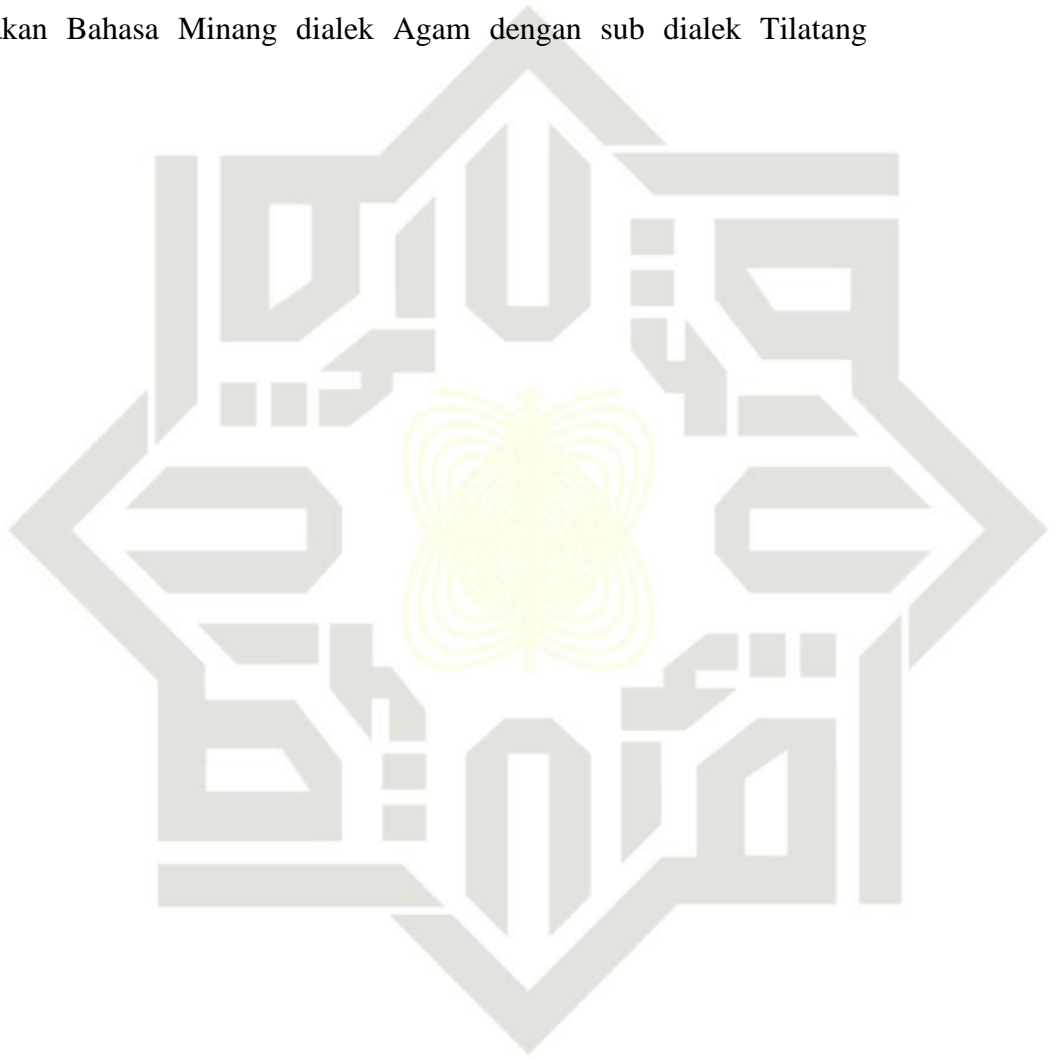
Diketahui bagaimana penulisan yang benar dari ucapan yang disampaikan menggunakan bahasa Minang dialek Agam dengan sub dialek Tilatang Kamang.



Pengembangan model STT ini, dapat mendukung pengembangan teknologi yang ada di kecamatan Tilatang Kamang.

Melestarikan dan mempromosikan Bahasa Minang dialek Agam ke khalayak ramai.

Membantu para penderita disabilitas yang kesulitan dalam berkomunikasi menggunakan Bahasa Minang dialek Agam dengan sub dialek Tilatang Kamang.



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 2

KAJIAN PUSTAKA

2.1 Landasan Teori

Kajian dari penelitian yang menjadi rujukan pada penelitian ini sebagai berikut:

2.1.1 Bahasa

Bahasa tidak hanya bermanfaat untuk berbicara, tetapi juga untuk menyampaikan pendapat, argumen, dan pandangan seseorang kepada orang lain. Bahasa daerah termasuk bahasa yang paling banyak digunakan untuk berkomunikasi. Salah satu bahasa daerah di Indonesia adalah bahasa Minang. Ini adalah salah satu bahasa dari keluarga bahasa Melayu dan rumpun bahasa Austronesia. Bahasa Minang memiliki banyak dialek dan sub dialek, salah satunya dialek Agam (Heni & Suryadi, 2022).

2.1.2 Speech-to-Text (STT)

Speech-to-Text adalah layanan pengenalan suara yang mengubah suara menjadi teks. Mengurangi kesalahan ejaan yang mungkin dilakukan pengguna saat mereka mengetik adalah salah satu keunggulan utama dari layanan ini. Di sisi lain, bagi orang-orang yang mengalami disabilitas, perubahan suara ke dalam teks akan sangat membantu. Keuntungan utama dari teknologi ini adalah orang normal berbicara lebih cepat daripada menulis, pengguna cukup menginputkan suara, sehingga lebih menghemat waktu dan mempermudah pekerjaan dalam mencatat (I Komang Setia Buana, 2020).

2.1.3 CMU Sphinx

CMU merupakan singkatan dari *Carnegie Mellon University* dimana tempat berkembangnya CMU Sphinx yang merupakan sebuah *toolkit* speech recognition *open-source* yang dapat digunakan untuk berbagai bahasa (Lamere et al., 2023). CMU Sphinx dikenal karena fleksibel dan banyak digunakan untuk membuat model pengenalan suara. Hal ini membantu perangkat dengan sumber daya terbatas karena memanfaatkan model akustik untuk mengoptimalkan penggunaan memori. Selain

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ini, teknik decoding yang efektif dalam mempercepat pengenalan suara tanpa menggunakan banyak memori.

Ada tiga komponen utama yang ada dalam CMU Sphinx yaitu

- a. Pocketsphinx – sebagai alat untuk pengenalan suara.
- b. Sphinxtrain – sebagai alat melatih model akustik.
- c. Sphinxbase – sebuah *library* yang diperlukan untuk menjalankan Pocketsphinx.

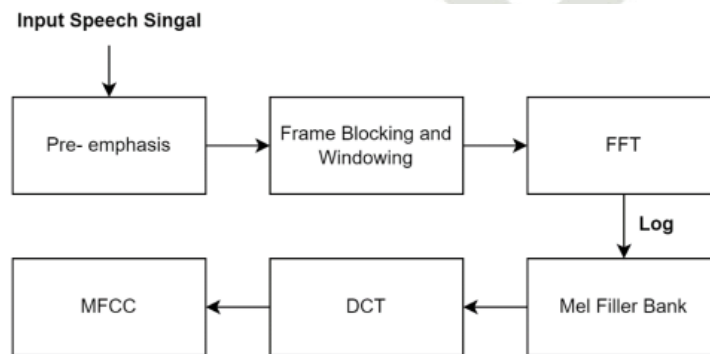
2.1.4 Mel Frequency Cepstral Coefficient (MFCC)

Mel Frequency Cepstral Coefficient (MFCC) merupakan sebuah metode yang digunakan untuk ekstraksi fitur sebuah proses yang mengkonversikan sinyal suara menjadi beberapa parameter. MFCC juga merupakan salah satu metode yang banyak digunakan dalam bidang speech technology. MFCC juga merupakan salah satu metode yang paling optimal dalam ekstraksi fitur pada pengolahan suara. Hal ini dikarenakan metode MFCC menggunakan prinsip-prinsip pendengaran suara manusia (Rahmadani, 2020). MFCC mengubah frekuensi dalam satuan Hz ke dalam skala m (skala mel) dan begitu pula sebaliknya dengan fungsi pada persamaan (1) dan (2).

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \tag{1}$$

$$f = 700 \left(10^{\frac{m}{2595}} - 1 \right) \tag{2}$$

Berikut merupakan alur yang menjelaskan proses ekstraksi sinyal suara pada metode MFCC:



Gambar 2. 1 Flowchart MFCC

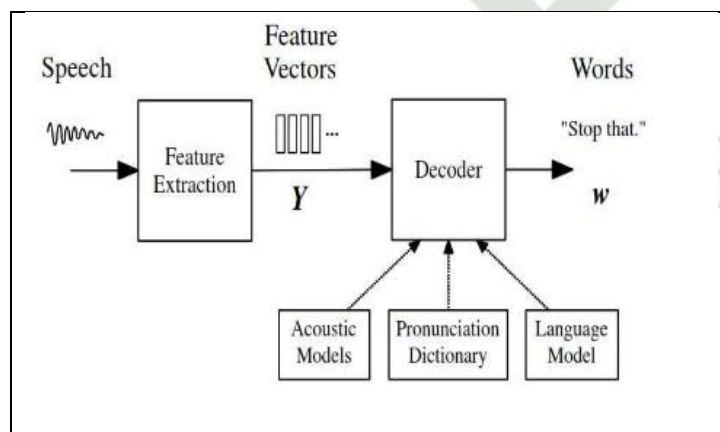
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Pre-emphasis, tahap ini merupakan tahapan proses filtering untuk meningkatkan komponen frekuensi tinggi dalam sinyal. Tujuannya adalah untuk memperkuat informasi penting dalam sinyal ucapan.
2. Frame blocking dan windowing, tahap ini sinyal suara dibagi menjadi frame-frame kecil. Setiap frame kemudian dikalikan dengan jendela (windowing) untuk mengurangi perubahan di ujung-ujung frame.
4. Fast Fourier Transform (FTT), tahap ini digunakan untuk mengubah setiap frame dari domain waktu ke domain frekuensi.
5. *Mel Frequency Bank*, tahap untuk proses pemfilteran Spektrum frekuensi dalam skala logaritmik kemudian difilter menggunakan bank filter berbasis skala Mel. Skala Mel meniru karakteristik pendengaran manusia yang lebih sensitif terhadap frekuensi rendah daripada frekuensi tinggi.
6. *Discrete Cosine Transform (DCT)*, Tahap ini Output dari Mel Filter Bank selanjutnya ditransformasi menggunakan DCT. DCT mengubah domain frekuensi menjadi domain Cepstral, yang akan menghasilkan koefisien MFCC.

2.1.5 Hidden Markov Model (HMM)

Merupakan metode untuk pengenalan suara otomatis (Tirta et al., 2022). Sinyal akustik dimodelkan oleh sekelompok unit akustik kecil yang dapat dianggap sebagai bunyi dasar bahasa. Secara tradisional, satuan yang dipilih adalah fonem, dan karena itu kata dibentuk dengan menggabungkannya. Gambar 2.2 menunjukkan komponen utama dari pengenalan ucapan pada HMM



Gambar 2. 2 Flowchart HMM

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Untuk menghasilkan kata-kata yang sesuai dengan kamus pengucapan, model akustik yang sesuai digabungkan dengan model bahasa serta *pronunciation dictionary*. Data pelatihan, yang terdiri dari bentuk gelombang ucapan dan transkripsi, digunakan untuk menghitung parameter model telepon ini. Dalam model bahasa yang paling umum seperti model N-gram di mana kemungkinan setiap kata hanya dikondisikan pada parameter N-gram dapat diperkirakan dengan menghitung N-gram yang sesuai.

Dekoder beroperasi dengan mencari melalui semua urutan kata yang mungkin untuk menghilangkan hipotesis yang tidak mungkin sehingga menjaga pencarian tetap dapat dilakukan. Ketika akhir ujaran tercapai, urutan kata yang paling mungkin adalah keluaran.

2.1.5 Word Error Rate (WER)

WER (*Word Error Rate*) merupakan Teknik untuk mengukur sejauh mana hasil output model yang dihasilkan oleh model STT mendekati teks referensi atau teks yang seharusnya dihasilkan. Nilai WER dinyatakan dalam persentase, dan semakin rendah nilai WER, maka semakin baik kinerja model STT (Aminudin, 2024). WER kurang dari 10% dianggap memiliki kualitas yang baik dan siap digunakan. WER sebesar 20% dapat diterima, dan WER sebesar 30% atau lebih menunjukkan kualitas yang buruk (Microsoft Learn, n.d.). Rumus WER adalah:

$$WER = \frac{S + D + I}{N} \times 100 \quad (3)$$

Keterangan:

S = adalah jumlah kata yang diganti (substitution),

D = adalah jumlah kata yang dihapus (deletion),

I = adalah jumlah kata yang disisipkan (insertion)

N = adalah jumlah kata dalam teks referensi.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel nilai pengujian menggunakan WER (*Word Error Rate*) sebagai berikut:

Tabel 2. 1 Skala WER

Score	Keterangan
≤10%	Amat Baik
11 – 20%	Baik
21 – 29%	Biasa
≥30%	Buruk

2.2 Penelitian Terkait

Penelitian terkait yang digunakan pada penelitian ini diambil dari beberapa sumber referensi, yakni :

Tabel 2. 2 Penelitian Terkait

No	Peneliti	Judul	Hasil
1.	Channareth Srun, Virbora Ny, Chea Cheat, Sokheang Ching, Piseth Ny	<i>Development of Speech Recognition System Based on CMU Sphinx for Khmer Language</i>	Penelitian ini membahas <i>Automatic Speech Recognition (ASR)</i> untuk bahasa Khmer menggunakan toolkit CMUSphinx, melibatkan rekaman dari 85 pembicara dan 157 kata Khmer, mencapai akurasi <i>Word Error Rate</i> sebesar 10.1% dan akurasi pengenalan kata hingga 10.0%, Penelitian ini juga mencatat tantangan dalam membangun sistem ASR untuk bahasa baru, serta pentingnya kualitas data pelatihan dan akurasi model. Sistem ini dirancang dengan kebutuhan sumber daya yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			rendah, sehingga cocok digunakan untuk aplikasi.
2.	Mohamed Yassine El Amrani, M.M. Hafizur Rahman, Mohamed Ridza Wahiddi, Asadullah Shah.	Building CMU Sphinx language model for Quran using simplified Arabic phonemes	Hasil eksperimen menunjukkan bahwa model mencapai akurasi pengenalan yang baik dengan data pelatihan yang cukup, tetapi kinerjanya menurun signifikan saat jumlah data dikurangi, dengan tingkat kesalahan kata (WER) minimum 55.68% saat 80% file audio digunakan untuk pelatihan. Penggunaan fonem sederhana menghasilkan WER sangat rendah sebesar 1.5% pada dataset kecil, namun meningkat saat diuji dengan data yang tidak terlihat.
3.	Sri Ningsih, Saniati	Eksperimen Pengenalan Ucapan Aksara Lampung Dengan Cmu Sphinx 4	Penelitian ini mengeksplorasi pengenalan ucapan aksara Lampung menggunakan CMU Sphinx 4, dengan fokus pada akurasi pengenalan suku kata, kata, dan frasa. Hasil eksperimen menunjukkan bahwa akurasi tertinggi dicapai pada pengenalan kata (74,03%), sementara akurasi suku kata (52,47%) lebih rendah karena kemiripan pelafalan.
4.	Dunya Yousufzai	Early Stages of Automatic Speech Recognition (ASR) in Non-	Dengan mempersiapkan dataset yang terdiri dari lebih dari 10 jam rekaman dan membuat kamus fonetik menggunakan ARPAbet, sistem ini berhasil mencapai tingkat kesalahan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

		english Speaking Countries and Factors That Affect the Recognition Process	kata (WER) sebesar 3,3% untuk CMU Sphinx dan 1% untuk DeepSpeech. Penelitian ini menyoroti pentingnya kualitas dataset, pemodelan fonetik, dan pemilihan algoritma, serta tantangan yang dihadapi dalam membangun sistem ASR untuk bahasa yang kurang terwakili.
5.	Preben Vangberg, Leena Sarah Farhat, Dewi Bryn Jones, Sean Kinahan	Developing Live Welsh Speech Recognition Models for a Commercial Product	Studi ini berhasil mengembangkan model pengenalan ucapan bahasa Welsh menggunakan Coqui STT dan data pelatihan dari Common Voice, youtube, dan Banc Trawsgrifiadau. Hasilnya menunjukkan peningkatan signifikan dalam akurasi model.
6.	Chi Han, Mingxuan Wang, Heng Ji, Lei Li	Learning Shared Semantic Space for <i>Speech-to-Text</i> Translation	Model Chimera berhasil mencapai hasil yang canggih dalam terjemahan <i>Speech-to-Text</i> pada berbagai benchmark, dengan skor BLEU 27,1 pada must-C EN-DE, yang melampaui hasil terbaik yang pernah dilaporkan hingga +1,9 BLEU. Model ini juga mampu mengungguli teknik lain dalam penerjemahan bahasa dan menunjukkan kemampuannya untuk mempelajari ruang semantik bersama yang terstruktur dengan baik serta berbagi pengetahuan antara machine translation (MT) dan speech translation (ST). Selain itu, model ini juga memanfaatkan data MT



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			eksternal untuk meningkatkan kinerja terjemahan dan visualisasi ablasi pengetahuan dibagikan di seluruh tugas juga menunjukkan potensi model dalam menampung pengetahuan umum yang dibagikan.
7.	Ilnar Salimzianov	A baseline model for computationally inexpensive speech recognition for Kazakh using the Coqui STT framework	Penelitian ini menunjukkan penggunaan model bahasa dapat meningkatkan akurasi dan efisiensi sistem Pengenalan Ucapan Otomatis (ASR) untuk bahasa Kazakh. Model akustik yang dilengkapi dengan model bahasa memiliki tingkat kesalahan kata (WER) dan tingkat kesalahan karakter (CER) yang lebih rendah daripada model akustik saja, dan mereka juga lebih cepat dalam pemrosesan kata. Selain itu, penelitian ini menunjukkan bahwa model Coqui STT dapat digunakan dalam mode streaming pada komputer tanpa GPU khusus, serta potensi ponsel pintar. Penelitian juga menunjukkan sistem ASR pertama untuk bahasa Kazakh yang dapat digunakan dalam mode streaming pada komputer non-GPU.
8.	Jonathan Mukiibi, Andrew Katumba, Joyce	The Makerere Radio Speech Corpus: A Luganda Radio Corpus for	Hasil dari jurnal ini mencakup pengembangan sistem pengenalan ucapan otomatis untuk pemantauan radio menggunakan bahasa Luganda di Uganda. Mereka merilis korpus pidato



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	Nakatumba-Nabende, Ali Hussein, Josh Meyer	Automatic Speech Recognition	radio Luganda berdurasi 155 jam, yang merupakan kumpulan data radio pertama yang tersedia untuk umum di Afrika Sub-Sahara menggunakan toolkit Coqui STT untuk mengembangkan model ASR Luganda dan menunjukkan peningkatan dalam deteksi kata kunci terkait COVID-19. Mereka juga membahas pekerjaan terkait dalam penggunaan ASR untuk pemantauan radio.
9.	Yanan Jia	A Deep Learning System For Domain-Specific Speech Recognition	Penelitian ini menunjukkan bahwa pengembangan sistem Automatic Speech Recognition (ASR) yang disesuaikan dengan domain tertentu dapat mengungguli sistem ASR komersial, bahkan ketika tingkat kesalahan kata transkripsinya lebih tinggi. Metode anotasi pembelajaran semi-supervisi digunakan untuk mengumpulkan data khusus domain, dan hasilnya menunjukkan peningkatan kualitas transkripsi dan ukuran korpus pelatihan.
10.	Jian Wu, Yashesh Gaur, Zhuo Chen, Long Zhou, Yimeng Zhu,	On Decoder-Only Architecture For Speech-to-Text And Large Language	Penelitian ini menunjukkan bahwa pendekatan <i>Speech-to-Text</i> menggunakan metode LLM, dengan menggabungkan informasi akustik ke dalam model bahasa besar berbasis teks, menjanjikan dalam meningkatkan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	Tianrui Wang, Jinyu Li, Shujie Liu, Bo Ren, Linqun Liu, Yu Wu	Model Integration	kinerja model dalam tugas penerjemahan ucapan.
1.	Panji Arisaputra, Amalia Zahra	Indonesian Automatic Speech Recognition with XLSR-53	Penelitian ini menunjukkan bahwa pengembangan Automatic Speech Recognition (ASR) untuk bahasa Indonesia menggunakan model XLSR-53 yang telah dipre-train dapat mencapai tingkat akurasi yang baik, terutama dengan dataset terbatas yang tersedia untuk bahasa Indonesia. Penelitian ini mencapai Word Error Rate (WER) sebesar 20% dengan 24 jam data, yang dapat lebih lanjut dikurangi menjadi 12% dengan penggunaan model bahasa. Model XLSR-53 juga memungkinkan representasi ucapan lintas bahasa dan telah dilatih dalam 53 bahasa yang berbeda.
2.	Dwita Deslianti, Nio Ade Syahputra	Aplikasi <i>Speech-to-Text</i> Bahasa Indonesia Ke Bahasa Bengkulu	Penelitian ini menunjukkan bahwa implementasi Pocketsphinx dalam aplikasi dapat mengenali suara dengan baik. Namun, masih terdapat faktor-faktor yang mempengaruhi akurasi pengenalan suara. Penulis menyarankan untuk menggunakan metode lain yang



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

		Menggunakan Pocketsphinx Berbasis Android	lebih baik dalam pengenalan kata dan untuk mengembangkan aplikasi agar hasil terjemahan dapat didengar, bukan hanya ditampilkan dalam bentuk teks.
3.	Marianne Engen Matre & David Lansing Cameron	Teknologi Speech-to-Text untuk remaja dengan kesulitan dalam pendidikan menengah	Hasil penelitian ini menunjukkan bahwa penggunaan teknologi <i>Speech-to-Text</i> (STT) dapat memberikan manfaat bagi siswa dengan kesulitan belajar di tingkat pendidikan menengah. Beberapa temuan utama termasuk peningkatan keterampilan menulis, penilaian teks, proses menulis, akurasi teknologi, dan pengalaman peserta. Selain itu, teknologi STT juga dapat meningkatkan pengenalan kata, pemahaman membaca, dan kualitas esai bagi siswa dengan kesulitan belajar. Namun, terdapat juga beberapa tantangan yang dihadapi peserta dalam menggunakan teknologi STT, seperti kesalahan pengenalan kata, kurangnya kompetensi dalam menggunakan alat tersebut, dan perasaan malu atau frustrasi.
4.	Ayushi Trivedi, Navya Pant, Pinal Shah, Simran Sonik and	<i>Speech-to-Text</i> and text to speech recognition systems- Areview	Hasil dari penelitian ini mencakup pembahasan tentang sistem pengenalan <i>Speech-to-Text</i> dan teks ke ucapan, serta sistem penerjemahan mesin. Beberapa teknik yang dibahas meliputi penggunaan model statistik, jaringan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	Supriya Agrawal		saraf, dan model markov tersembunyi. Selain itu, terdapat juga pembahasan tentang metode konversi <i>Speech-to-Text</i> dan teks ke pidato, termasuk masalah yang mungkin timbul dalam proses konversi.
5.	Rafael Valle, Kevin Shih, Ryan Prenger, Bryan Catanzaro	Jaringan Generatif berbasis Aliran Autoregresif untuk Sintesis Teks- ke-Ucapan	model sintesis teks ke melspektogram baru, Flowtron, mampu menghasilkan sintesis ucapan yang berkualitas tinggi dengan kontrol atas variasi ucapan dan transfer gaya. Model ini dioptimalkan dengan memaksimalkan kemungkinan data pelatihan, membuat pelatihan menjadi sederhana dan stabil.
16.	Changhan Wang, Yun Tang, Xutai Ma, Anne Wu, Sravya Popuri, Dmytro Okhonko, Juan Pino	FAIRSEQ S2T: Fast <i>Speech-to-Text</i> Modeling with Fairseq	pengembangan FAIRSEQ S2T, sebuah toolkit untuk tugas pemodelan ucapan ke teks seperti pengenalan ucapan dan terjemahan ucapan ke teks. Toolkit ini menyediakan alur kerja end-to-end, model-model terkini, dan resep pelatihan sumber terbuka. FAIRSEQ S2T mendukung pembelajaran multi-task, transfer learning, dan pelatihan self-supervised. Toolkit ini juga mencakup pra-pemrosesan data, komputasi, metrik evaluasi, dan alat visualisasi. Hasil eksperimental menunjukkan efektivitasnya pada dataset benchmark seperti LibriSpeech, MuST-C, dan CoVoST 2 untuk tugas



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			<p>pengenalan ucapan dan terjemahan ucapan ke teks. Model-model yang dikembangkan mencapai hasil yang kompetitif dibandingkan dengan pendekatan terkini tanpa penyetelan khusus untuk tugas tertentu.</p>
7.	<p>Yi Ren, Jinglin Liu, Xu Tan,Chen Zhang, Tao QinZhou Zhaoy</p>	<p>SimulSpeech: End-to-End Simultaneous Speech to Text Translation</p>	<p>pengembangan model SimulSpeech untuk menerjemahkan ucapan secara simultan ke teks dengan menggunakan teknik knowledge distillation. Model ini berhasil mencapai skor BLEU yang cukup baik dan menunjukkan penundaan yang lebih rendah dibandingkan dengan model end-to-end speech to text translation.</p>
18.	<p>Awino, Ebbie; Wanzare, Lilian; Muchemi, Lawrence; Wanjawa, Barack; Ombui, Edward; Indede, Florence; McOnyango, Owen; Okal, Benard</p>	<p>Phonemic Representation and Transcription for Speech to Text Applications for Under- resourced Indigenous African Languages: The Case of Kiswahili</p>	<p>penelitian ini berhasil mengatasi beberapa tantangan dalam pengembangan sistem pengenalan ucapan otomatis (ASR) untuk bahasa-bahasa Afrika yang kurang diresapi, dengan menggunakan Kiswahili sebagai studi kasus. Proses transkripsi dan pengembangan korpus ucapan Kiswahili dilakukan, serta pembuatan kamus fonem yang komprehensif untuk model ASR. Selain itu, penelitian ini juga berhasil mengumpulkan data teks dan ucapan Kiswahili melalui proyek KenCorpus, yang mencakup berbagai jenis data</p>



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			seperti wawancara, percakapan, cerita, dan berita.
9.	Jaeyoung Huha, Sangjoon Parka, Jeong Eun Leeb, Jong Chul Yec	Improving Medical <i>Speech-to-Text</i> Accuracy with Vision-Language Pre-training Model	pengembangan metode Multi-modal Medical Speech Module (MMSM) yang menggunakan Vision-Language Pre-training (VLP) model untuk meningkatkan kinerja sistem pengenalan ucapan (STT) di bidang medis. Eksperimen menunjukkan bahwa metode yang diusulkan memberikan peningkatan signifikan dalam kinerja STT di bidang medis. Model MMSM yang dihasilkan juga menunjukkan kinerja yang lebih baik dalam memperbaiki kesalahan dibandingkan dengan sistem pengenalan ucapan lainnya. Metode ini juga menunjukkan keunggulan dalam menangani kondisi ucapan dunia nyata dan memiliki potensi untuk diterapkan di bidang radiologi. Penelitian ini didukung oleh National Research Foundation of Korea dan memiliki implikasi untuk pengenalan ucapan otomatis di bidang radiologi.
10.	Bogdan IANCU	Evaluating Google <i>Speech-to-Text</i> API's Performance	penelitian ini mengevaluasi kinerja Google Cloud <i>Speech-to-Text</i> API dalam mentranskripsi sumber daya e-learning berbahasa Rumania. Hasilnya menunjukkan bahwa model yang



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

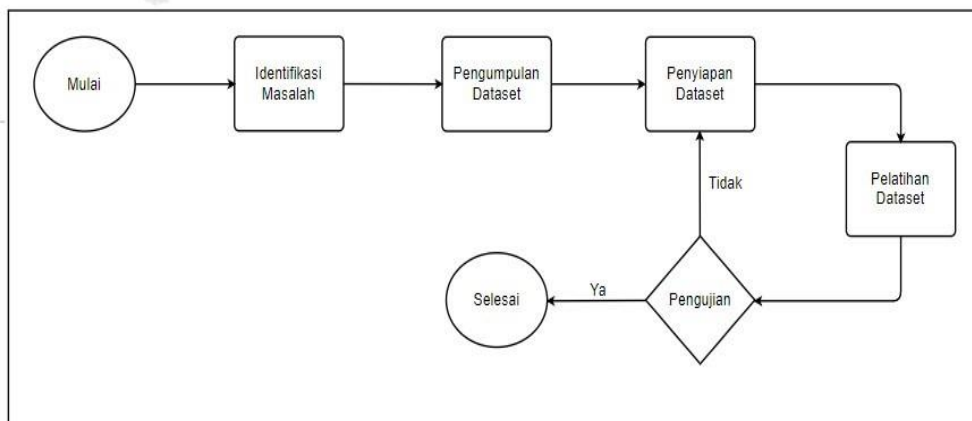
	for Romanian e-Learning Resources	diusulkan dapat mengotomatisasi pengindeksan sumber daya multimedia e-learning. Meskipun Word Error Rate (WER) untuk API tersebut adalah 30.96%, hasil yang lebih baik diperoleh untuk pembicara perempuan dan subjek video tertentu. Penelitian ini bertujuan untuk meningkatkan akurasi model ASR dan menggunakan hasilnya untuk pengindeksan semantik sumber daya e-learning. Penelitian ini dilakukan sebagai bagian dari proyek ATLAS yang didanai oleh UEFISCDI.
--	-----------------------------------	--

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 3

METODOLOGI PENELITIAN

Dalam penelitian ini, metodologi penelitian ini terdiri dari penjelasan tahap-tahapan pada penelitian ini, dengan tahapan sebagai berikut:



Gambar 3. 1 Metodologi Penelitian

3.1 Identifikasi Masalah

Tahap ini, permasalahan yang telah diketahui dari penelitian ini adalah bagaimana cara menerapkan model STT menggunakan CMU Sphinx untuk bahasa Minang dialek Agam. Agar penerapan model STT menggunakan CMU Sphinx dapat berjalan lancar, perlu adanya data dengan kata yang banyak, kecepatan dan kejelasan dalam pengucapan kata yang tepat, serta fonetik yang benar. Jumlah data dan kualitas data dapat mempengaruhi keberagaman kata yang dihasilkan oleh teknologi *Speech-to-Text*.

3.2 Pengumpulan Dataset

Pada tahapan yang dua ini peneliti melakukan pengumpulan data untuk mendapatkan dataset yang sesuai sehingga bisa diterapkan pada teknologi *Speech-to-Text* (STT) menggunakan CMU Sphinx, persiapan yang perlu dilakukan sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.1.6 Pembuatan Teks

Pada tahap ini peneliti mulai membuat transkrip teks bahasa Minang dialek Agam yang mana transkrip teks bersumber dari internet dan hasil dari wawancara dengan Pengisi suara. Transkrip teks yang telah dikumpulkan dilakukan validasi untuk memastikan bahwa dataset ini benar menggunakan bahasa Bahasa Minang Dialek Agam dengan sub dialek Tilatang Kamang. Transkrip teks pada penelitian ini berjumlah 310 kalimat dengan 838 kata.

2.1.7 Perekaman Audio

Dalam proses perekaman audio, audio yang direkam berdurasi pendek. Proses perekaman audio ini menggunakan Pengisi suara wanita dewasa berumur 42 tahun bernama Ibu Delina Feriska yang merupakan masyarakat asli Tilatang Kamang. Proses perekaman audio dilakukan di tempat yang minim kebisingan sehingga mendapatkan rekaman audio dengan kualitas yang baik. Hal yang perlu diperhatikan dalam perekaman audio pada penelitian ini sebagai berikut:

1. Audio dalam format WAV, 16-bit, mono PCM.
2. Data audio masing-masing berdurasi antara 0 sampai 10 detik.
3. Data memiliki kecepatan sampel 16000 Hz.
4. Audio memiliki noise dan distorsi latar belakang yang minimal.
5. Data tidak memiliki jeda hening yang lama di awal, di tengah, dan diakhir.

3.3 Penyiapan Dataset

Pada tahapan ini, dataset yang telah dikumpulkan akan melakukan normalisasi. Pada data teks dilakukan penghapusan kata yang sama dan tanda baca, setelah dilakukan penghapusan kata yang sama dan tanda baca, maka kata tersebut akan diurutkan sesuai abjad. Pada data audio dilakukan normalisasi yang bertujuan untuk menghapus noise yang ada dalam rekaman audio. Setelah dilakukan normalisasi, diperoleh audio yang bersih dari noise dengan jumlah 838 data suara dan 838 data kata dari 310 data kalimat. Data yang telah siap kan maka akan dikelompokkan. Pada data suara akan disimpan dalam format *file* .wav, data kalimat akan disimpan dalam format .xlsx untuk data kata akan disimpan dalam format txt. Dataset yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 5 PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dari penelitian ini, maka dapat ditarik kesimpulan sebagai berikut:

Diperoleh 310 kalimat atau 838 kata, serta 838 rekaman suara wanita dewasa yang merupakan masyarakat asli Minang, kabupaten Agam yang menggunakan sub-dialek Tilatang Kamang

Hasil dari teknologi *Speech-to-Text* Bahasa Minang Dialek Agam Menggunakan CMU Sphinx menunjukkan bahwa model akustik ini berfungsi dan menampilkan hasil teks dari inputan suara secara langsung dengan akurasi dari outputnya cukup memuaskan. Pengujian menggunakan *Word Error Rate* (WER) mendapatkan *score* sebesar 27.2% atau 228 kata mengalami kesalahan pengenalan dari total 838 kata yang diuji.

3. Pemilihan fonetik yang sesuai sangat penting karena dapat meningkatkan kualitas dari model pengenalan suara. Penggunaan fonetik yang sesuai dengan bahasa yang ditargetkan, membuat model lebih mudah mengenali dan memproses suara nantinya.

5.2 Saran

Berdasarkan hasil yang telah didapatkan pada penelitian ini, maka diberikan saran untuk penelitian kedepannya sebagai berikut:

Penelitian selanjutnya dapat menggunakan dataset seperti data suara pria dewasa atau anak-anak yang memiliki gelombang suara yang berbeda, sehingga dataset dapat lebih bervariasi

Penelitian selanjutnya diharapkan dapat menangani kata-kata yang tidak jelas dan kata yang hampir sama dalam penulisan dan pengucapan. Sehingga kata



yang muncul sesuai dengan yang diucapkan dan meningkatkan keakuratan kata yang didapatkan bisa sesuai yang diharapkan.

Penelitian selanjutnya diharapkan memilih fonetik yang sesuai dengan bahasa yang ditargetkan sehingga dapat meningkatkan akurasi dan mengurangi nilai WER.



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Aminudin, I. (2024). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Pengolahan Korpus Dataset Audio Bacaan Al-Qur'an Menggunakan Metode Wav2Vec 2.0.
- Arisaputra, P., & Zahra, A. (2022). Indonesian Automatic Speech Recognition with XLSR-53. *Ingenierie Des Systemes d'Information*, 27(6), 973–982.
- Awino, Wanzare, Lilian, Muchemi, Lawrence, Wanjawa, Barack, Ombui, Edward, Indede, Florence, Mconyango, Owen, Okal, & Benard. (2022). Phonemic Representation and Transcription for Speech to Text Applications for Under-resourced Indigenous African Languages: The Case of Kiswahili.
- Ayub, A., Husin, N., Amir, M., Usman, H., Yasin, A., Pembinaan, P., & Bahasa, P. (1993). *Tata Bahasa Minangkabau*.
- Deslianti, D., & Syahputra, N. A. (2018). Aplikasi Speech To Text Bahasa Indonesia Ke Bahasa Bengkulu Menggunakan Pocketsphinx Berbasis Android.
- El Amrani, M. Y., Rahman, M. M. H., Wahiddin, M. R., & Shah, A. (2016). Building CMU Sphinx language model for the Holy Quran using simplified Arabic phonemes. *Egyptian Informatics Journal*, 17(3), 305–314.
- Han, C., Wang, M., Ji, H., & Li, L. (2021). Learning Shared Semantic Space for Speech-to-Text Translation.
- Heni, A. N., & Suryadi, M. (2022). Variasi Leksikal Bahasa Minangkabau Di Kanagarian Kubang Putih, Kabupaten Agam: Kajian Sosiodialektologi. *Widyaparwa*, 50(1), 151–161.
- Huh, J., Park, S., Lee, J. E., & Ye, J. C. (2023). Improving Medical Speech-to-Text Accuracy with Vision-Language Pre-training Model.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- I Komang Setia Buana. (2020). Implementasi Aplikasi Speech to Text untuk Memudahkan Wartawan Mencatat Wawancara dengan Python. *Jurnal Sistem Dan Informatika (JSI)*, 14(2), 135–142.
- IANCU, B. (2019). Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources. *Informatica Economica*, 23(1/2019), 17–25.
- Jia, Y. (2023). A Deep Learning System for Domain-specific Speech Recognition.
- Kharisma, N., Nadra, N., & Reniwati, R. (2021). Fonologi Bahasa Minangkabau Isolek Sikukur. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, 4(4), 425–440.
- Lamere, P., Kwok, P., Walker, W., Gouvêa, E., Singh, R., Raj, B., & Wolf, P. (2003a). Design of the CMU Sphinx-4 decoder. *EUROSPEECH 2003 - 8th European Conference on Speech Communication and Technology*, 1181–1184.
- Matre, M. E., & Cameron, D. L. (2022). A scoping review on the use of speech-to-text technology for adolescents with learning difficulties in secondary education. In *Disability and Rehabilitation: Assistive Technology*. Taylor and Francis Ltd.
- Menguji akurasi model Ucapan Kustom - Layanan ucapan - Azure AI services | Microsoft Learn. (n.d.). Retrieved January 8, (2024),
- Mukiibi, J., Katumba, A., Nakatumba-Nabende, J., Hussein, A., & Meyer, J. (2022). The Makerere Radio Speech Corpus: A Luganda Radio Corpus for Automatic Speech Recognition.
- Negeri, P., & Rahmadani, L. (2020). Klasifikasi Gender Berdasarkan Suara Dengan Naive Bayes Dan Mel Frequency Cepstral Coefficient. *Technology Journal*, 2(1), 19–26.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ningsih, S. (2018). Eksperimen Pengenalan Ucapan Aksara Lampung Dengan CMU SPHINX 4. In *Jurnal Teknoinfo* (Vol. 12, Issue 1).

Ratumanan, S. D., Rahman, H., Karlina, D. A., Dani, G., Rahayu, S., Anggraini, G. F., Unpatti, P. P., Bahasa, P., Upi, D., Upi, P., Siliwangi, P. I., & Unila, P. (2022). Upaya Pemberdayaan Penggunaan Bahasa Daerah Melalui Budaya Literasi Digital. *Journal of Elementary Education*, 05.

Ren, Y., Liu, J., Tan, X., Zhang, C., Qin, T., Zhao, Z., & Liu, Y. (2020). SimulSpeech: End-to-End Simultaneous Speech to Text Translation.

Salimzianov, I. (2021). A baseline model for computationally inexpensive speech recognition for Kazakh using the Coqui STT framework.

Srun, C., Ny, V., Cheat, C., Ching, S., & Ny, P. (2021). Development of Speech Recognition System Based on CMUSphinx for Khmer Language. In *International Journal of Innovative Science and Research Technology* (Vol. 6, Issue 8).

Tirta, L., Santoso, J., & Setyati, E. (2022). Pengenalan Lirik Lagu Otomatis Pada Video Lagu Indonesia Menggunakan Hidden Markov Model Yang Dilengkapi Music Removal. *Journal of Information System, Graphics, Hospitality and Technology*, 4(2), 86–94.

Trivedi, A., Pant, N., Shah, P., Sonik, S., & Agrawal, S. (2018). Speech to text and text to speech recognition systems-Areview. 20(2), 36–43.

Valle, R., Shih, K., Prenger, R., & Catanzaro, B. (2020). Flowtron: an Autoregressive Flow-based Generative Network for Text-to-Speech

Vangberg, P., Farhat, L. S., Jones, D. B., & Kinahan, S. (2023). Developing Live Welsh Speech Recognition Models for a Commercial Product - a case study. 113–115.



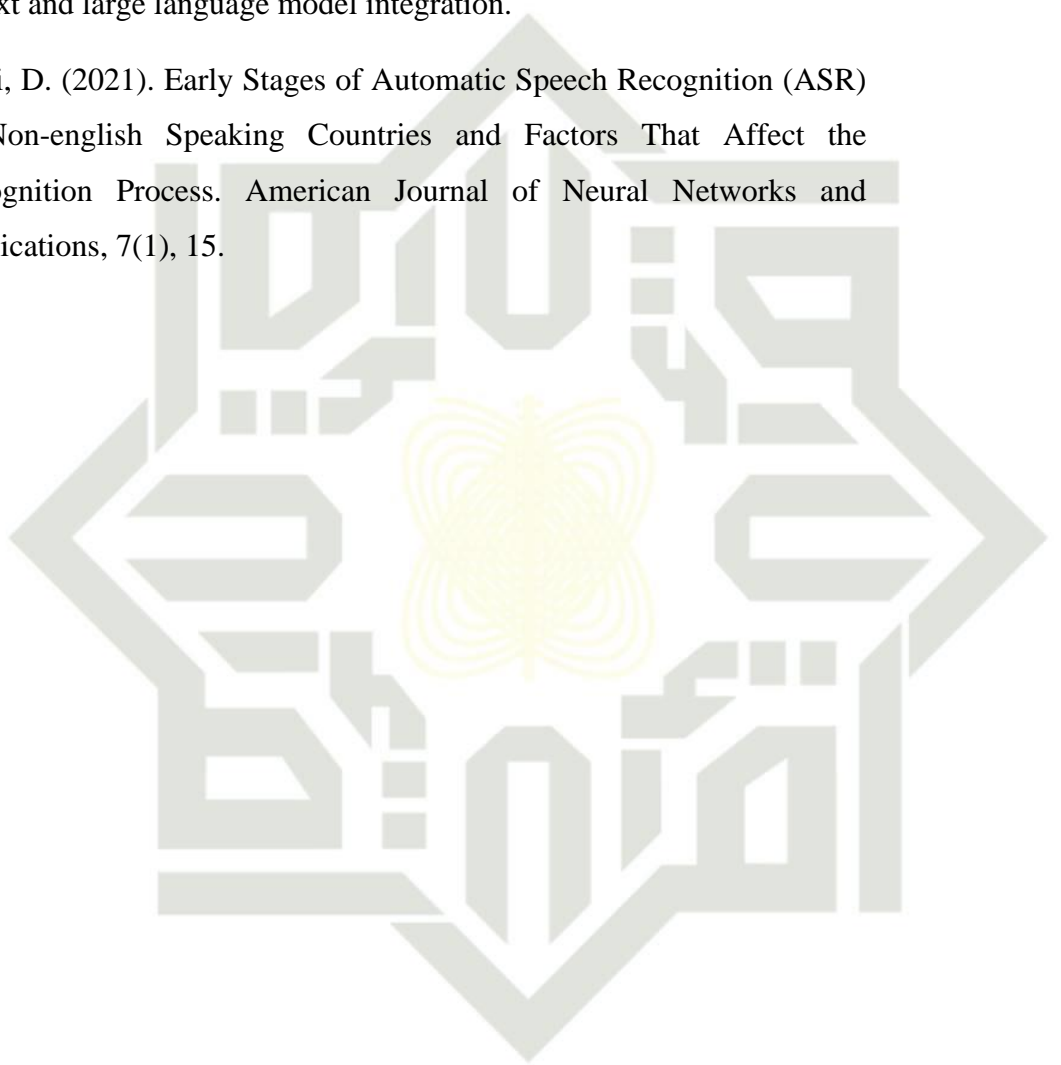
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Wang, C., Tang, Y., Ma, X., Wu, A., Popuri, S., Okhonko, D., & Pino, J. (2020). fairseq S2T: Fast Speech-to-Text Modeling with fairseq.

Wu, J., Gaur, Y., Chen, Z., Zhou, L., Zhu, Y., Wang, T., Li, J., Liu, S., Ren, B., Liu, L., & Wu, Y. (2023). On decoder-only architecture for speech-to-text and large language model integration.

Yousufzai, D. (2021). Early Stages of Automatic Speech Recognition (ASR) in Non-english Speaking Countries and Factors That Affect the Recognition Process. *American Journal of Neural Networks and Applications*, 7(1), 15.



UIN SUSKA RIAU



LAMPIRAN A

Surat pernyataan kebenaran data kalimat bahasa Minang (Dialek Agam) yang bertandatanganinya oleh ketua organisasi Koto Tangan Sepakat (KOTAS) Duri.

SURAT PERNYATAAN KEBENARAN DATA

Saya yang bertandatangan dibawah ini

Nama : HENORI .
 Tempat, Tanggal Lahir : BUKUTINGGI . 18/2 - 1965 .
 No. HP : 0813 6510 6479 .
 Jabatan : KETUA KOTO TANGAN SEPAKAT - DURI .
 Suku : Pili
 Alamat : JL. SAKURA . DURI .

Setelah mendapatkan penjelasan tentang maksud dan tujuan serta memahami penelitian yang dilakukan dengan judul " Penerapan *Speech-to-Text* Bahasa Minang Dialek Agam Menggunakan CMU Sphinx " yang dilakukan oleh:

Nama : Syahrul Yusman
 NIM : 12050112675
 No. HP : 0895618202555
 Jurusan/Fakultas : Teknik Informatika/Sains dan Teknologi
 Universitas : UIN SUSKA RIAU

Dengan ini menyatakan bahwa kalimat dalam bahasa Bahasa Minang Dialek Agam yang disampaikan pada lampiran ini untuk keperluan penelitian dan pendidikan adalah benar merupakan kalimat Bahasa Minang Dialek Agam sub dialek Tilatang Kamang.

Demikian surat pernyataan kebenaran data ini dibuat dengan sebenar-benarnya, tanpa ada paksaan dari pihak manapun, dan untuk digunakan sebagaimana mestinya.

Bengkalis, 16 Agustus 2024

Hormat Saya,

(HENORI)

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.


Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Surat pernyataan kebenaran pengisi suara bahasa Minang (Dialek Agam) yang ditandatangani oleh Ibuk Delina Feriska.

Surat Pernyataan Narasumber

Saya yang bertandatangan dibawah ini :

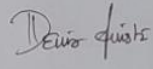


Nama	: DELINA FERISKA
E-mail	: zhanio.radhitya@gmail.com
No. HP	: 081363645033
Pendidikan	: SLTA
Suku	: SIKUMBANG
Alamat	: Jl. Flamboyan III NO. 3A Inkorba. Bukittinggi, Sumatera Barat

Dengan ini menyatakan bahwa data rekaman audio dari penelitian yang berjudul "Speech-to-Text Bahasa Minang (Dialek Agam) Menggunakan CMU Sphinx" merupakan bahasa Minang dialek Agam subdialek Tilatang Kamang serta mengizinkan data rekaman audio tersebut digunakan untuk keperluan pendidikan dan penelitian.

Demikian lembar pernyataan ini dibuat untuk digunakan dalam melengkapi Laporan Tugas Akhir.

Bukittinggi, 23 September 2024
Hormat Saya,



(...DELINA FERISKA...)



Table 1.1 Kalimat Bahasa Minang dan Terjemahannya

Kalimat	Terjemahan
Dalam hiduik sehari-hari, kito baganduah jo banyak suku jo budaya nan bamacam ragam, adaiak dari kabiasoan, adaik, tradisi, pemahaman jo lain sabagainyo.	dalam kehidupan sehari-hari, kita bergaul dengan banyak suku dan budaya yang beraneka ragam baik itu dari kebiasaan, adat, tradisi, pemahaman, dan lain sebagainya.
Salah satu suku nan paliang gadang di Indonesia adolah Minangkabau.	salah satu suku yang paling besar di indonesia adalah Minangkabau
baa mereka bisa sampai tasebar sakitu banyak nyo tapi tatap hiduik damai?	bagaimana mereka bisa sampai tersebar segitu banyaknya tetapi tetap hidup damai?
Ado nan karajonyo manggaleh, wirausaha basakolah, parantau dan masih banyak nan lainnyo.	ada yang kerjanya berdagang, wirausaha, bersekolah, merantau dan masih banyak yang lainnya.
Nyatonyo, sajak ketek mereka alah di agiah baka pituah Minang tantang kahidupan.	nyatanya, sejak kecil mereka telah diberi bekal petuah Minang tentang kehidupan.
Salain tun, pituah pun di bagi dalam babagai golongan.	selain itu, petuah dibagi dalam berbagai golongan.
pituah tu barupo nasihai-kasihai dari niniak mama, urang nan dituokan.	petuah itu berupa nasehat-nasehat dari pemuka adat, orang yang dihormati
Ajaran-ajaran nan elok barupo nasehat, pangaajaran moral jo nilai tamasuak perintah sacaro indak langsung nan di tujuan untuak urang banyak supaya dijadian sabagai pijakan jo padoman dalam manjalani hiduik.	ajaran-ajaran yang baik berupa nasehat, pelajaran moral dengan nilai termasuk perintah secara tidak langsung yang ditujukan untuk orang banyak supaya dijadikan sebagai pijakan dan pedoman dalam menjalani hidup.
Sahinggo apo nan di katoan bisa ba pangaruh untuak urang banyak isuak harinyo.	sehingga apa yang dikatakan bisa berpengaruh untuk orang banyak esok harinya.
Pangartian pituah di Minangkabau adolah	pengertian petuah Minangkabau adalah
Pituah Minang lamo umumnyo dikaluaan dek urang urang alim ulama, cadiak pandai nan punyo ilmu agamo islam nan tinggi.	petuah Minang lama umumnya dikeluarkan oleh orang-orang alim ulama, cerdik pandai yang punya ilmu agama islam yang tinggi.

Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan karya ilmiah, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ado jo untuak parantau, panyamangaik iduik, budi elok jo lain sabagainyo.	ada juga untuk perantau, penyemangat hidup, budi baik dan lain sebagainya.
Arati no mangicek paliharo lidah, ba tingkah laku paliharo sikap.	artinya, berbicara pelihara lidah, bertingkah pelihara sikap.
Bakato paliharo lidah, balaku paliharo budi.	berkata pelihara lidah berperilaku pelihara budi.
Anak padusi manjadi lambang kahormatan jo kemuliaan sarato sabagai bundo kanduang.	anak perempuan menjadi lambang kehormatan dan kemuliaan serta menjadi seorang ibu
Untuak itu niniak mamak, alim ulamo jo cadiak pandai mangaluakan banyak bana pituah Minang untuak padusi sacaro khusus.	untuk itu pemuka adat, alim ulama, serta cerdik pandai mengeluarkan banyak sekali petuah Minang untuk perempuan secara khusus.
Dalam adaik Minangkabau, sajak dulu posisi anak padusi di tengah masyarakaik sangaik lah pantiang dan tinggi.	dalam adat Minangkabau, sejak dulu posisi anak perempuan ditengah masyarakat sangatlah penting dan tinggi.
Kalau laki pulang karajo atau dari bajalan jauh mako sambuik lah baliiau jo muko manih.	kalau suami pulang kerja atau dari perjalanan jauh, maka sambutlah beliau dengan muka yang manis.
pa baiak lakuan, pa elok taratik.	perbaiki perilaku, perbaiki etika.
Karano laki alah latiah mencari pitih untuak keluarga.	karna suami sudah lelah mencari uang untuk keluarga.
kok upiak nak pai ka pakan, ma mintak izin lah dahulu bakeh laki.	kalau wanita akan pergi ke pasar, mintalah izin dahulu ke suami.
Agiah apo nan di mintak no, sanangan hati no, sadioan makan jo minum no.	berikan apa yang dimintanya, senangan hatinya, sediakan makan dan minumannya.
Di amaro pituah Minang untuak anak padusi adolah:	diantara petuah Minang untuk anak perempuan adalah.
Kok upiak naiak bendi, usah lah sa bendi jo urang lain nan bukan dunsanak.	kalau anak perempuan naik bendi, janganlah satu bendi dengan orang lain yang bukan saudara.
Kok padusi indak ba malu, jadi cacek sa umua hiduik.	kalau perempuan tidak ada malu, jadi hinaan seumur hidup.



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mengemukakan sumber dan menyebutkan sumber.

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penerbitan, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

lain pada itu, kalau pai jo kendaraan, jan sampai asing buangan duduak jo urang lain nan indak muhrim.	selain itu, kalau pergi dengan kendaraan, jangan sampai duduk bedempetan dengan orang lain yang bukan muhrim.
Karano sagalo hak ado di tangan laki.	karena segala hak ada ditangan suami
Bak parahu indak bakamudi, biasa sasek dalam alayia	seperti perahu tidak berkemudi, biasa sesat dalam berlayar.
Kok nio kalua rumah, mintak lah izin ka laki.	kalau ingin keluar rumah, mintalah izin kepada suami.
Usahoan kalau pai di kawan jo dunsanak nan tapek	usahakan kalau pergi ditemani dengan saudara yang tepat.
Bak pintu indak ba pasak, mudah rang maliang mamasuak i nyo.	bagai pintu tidak dikunci, mudah orang maling memasukinya
Mangko inyo ka manyasa sa umua hiduik.	maka ia akan menyesal seumur hidup
Ka ta ombang ambiang di tengah lauik.	akan terombang ambing ditengah laut.
Sarupo pintu nan indak ba kunci, sahinggo maliang bisa masuak	serupa pintu yang tidak dikunci, sehingga maling bisa masuk.
Sarupo kapa nan indak ba nakhoda.	seperti kapal tidak bernakhoda
Pituah Minang untuak padusi nan indak punyo raso malu, indak manjago kahormatan sabagai padusi muslimah.	petuah Minang untuk perempuan yang tidak punya rasa malu, tidak menjaga kehormatan sebagai perempuan muslimah.
Handak no mamahami syariat-syariat islam.	hendaknya memahami syariat – syariat islam.
Mangko dari itu, jan jadi padusi nan lemah pakaro akidah jo ilmu agamo.	maka dari itu, jangan jadi perempuan yang lemah tentang akidah dan ilmu agama.
Anak padusi handak no sadar kalau cukup lah rupono sajo nan indak bisa di rubah.	anak perempuan harusnya sadar kalau cukup lah rupa saja yang tidak bisa dirubah
Kok roman nan indak dapek di ubah, tapi kok laku jo parangai lai dapek ba ubah.	kalau bentuk wajah tidak dapat dirubah, tapi kalau perilaku dapat berubah
Tapi kalakuan jo parangai bisa diusahakan barubah ka arah nan labih elok.	tapi kelakuan dan perilaku bisa diusahakan berubah kearah yang lebih baik



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengantumkan sumber;
2. Dilarang mengutip hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
3. Dilarang mengutip dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1	Kok basuo jo urang lain, kok duduak di tengah ami atau di dalam alek jamu, caliak nan usah di atangi, mato usah dipalia.	kalua bertemu dengan orang lain, kalau duduk ditengah orang ramai atau di dalam pesta perjamuan, pandangan jangan ditinggikan, mata jangan liar
2	Pandang sakali lalu sajo, usahlah galak dipabihak.	pandang sekali saja, jangan lah tertawa terbahak-bahak.
3	Lambak nan dari pado itu, sopan jo santun tak ba bateh.	sungguh pun dari pada itu, sopan dan santun tidak ada batas
4	Ijan pulo bagarah-garah sampai galak tabahak jo kaum laki laki nan indak muhrim.	jangan pula bercanda sampai tertawa terbahak-bahak dengan kaum laki-laki yang tidak muhrim.
5	Indak paralu mancaliak lia kian-kamari, cuikuik caliak sakali sajo.	tidak perlu melihat kesana kemari, cukup lihat sekali saja.
6	Bia jo laki awak bana, jan lah ilang baso-basi	biar sama suami kita pun, janganlah hilang basa-basi.
7	Sabagai padusi nan ba gala bini, bia lah jo laki surang, jan ma hilangkan baso-basi.	sebagai perempuan yang bergelar istri, biarlah dengan suami sendiri, jangan menghilangkan basa-basi.
8	Di pamanih muluik sarato kucin nan murah juo handak nyo.	di permanis mulut serta candaan murah juga hendak nya
9	Gadang kan junjungan upiak di tengah rami, muliak inyo di muko rapek.	besarkan junjungan mu (wanita) di tengah keramaian, muliakan dia di tengah perkumpulan
10	Umpamo no, dalam bakumpua sanak family di rumah gadang atau bakumpua jo urang kampung, laki handak nyo tatap dijadian junjungan.	misalnya, dalam berkumpul dengan saudara keluarga di rumah gadang atau berkumpul dengan orang kampung, suami hendaknya tetap dijadikan junjungan.
11	Kok ba rundiang samo gadang, kalamahannyo usah dibukak an.	kalau diskusi sama besar, kelemahanya jangan ditunjukkan
12	Salanjuk no pituah Minang untuak laki laki	selanjutnya petuah Minang untuk laki-laki.
13	Umpamo no, kalamahan atau kakurangan no.	misalnya, kelemahannya atau kekurang nya



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber; 2. Dilarang mengutipnya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1	tu lah mangko no banyak bana laki-laki Minang an ma rantau.	maka dari itu banyak sekali laki-laki Minang yang merantau.
2	abagai bini urang, katiko batamu jo urang lain, aiak itu urang sabalah rumah, mintuo, urang aek, dan lain sabagainyo, jan mambukak an aib laki.	sebagai istri orang, ketika bertemu dengan orang lain, baik itu tetangga, mertua, orang tua, dan lain sebagai nya, jangan membuka aib suami.
3	alain pituah Minang untuak padusi, kaum laki-laki keturunan Minangkabau sangaik tanamo jo kapandaian no dalam manggaleh, wirausaha, membina rumah tangga atau pun batani.	selain petuah Minang untuk perempuan, kaum laki-laki keturunan Minangkabau sangat ternama dengan kepintarannya dalam berdagang, wirausaha, membina rumah tangga ataupun bertani.
4	Salain pado itu, laki-laki Minang alah di aja an ba'a hiduik surang atau mandiri.	selain dari pada itu, laki-laki Minang sudah diajarkan bagaimana hidup sendiri atau mandiri.
5	Jadi indak heran kalau banyak niniak mamak nan mangalua an pituah Minang untuak laki laki.	jadi tidak heran kalau banyak pemuka adat yang mengeluarkan petuah Minang untuk laki-laki
6	Hati lapang paham tak sampik.	hati lapang paham tidak sempit
7	Jan pulo anggan mandanga kalau urang manarangan sesuatu.	jangan enggan mendengar kalau orang menjelaskan sesuatu
8	Mangko no baraja jo manuntuik ilmu sapanuah hati.	makanya belajar dan menuntut ilmu sepenuh hati.
9	Panda manimbang jo manaka.	pandai menimbang dan mengira.
10	Walau batenggang di nan rumik, indak bakisa di nan bana.	walau berhadapan dengan yang rumit, tidak berpaling dari yg benar
11	Baiak itu laki-laki dewasa anak bujang jo anak-anak.	baik itu laki-laki dewasa, anak remaja, dengan anak-anak.
12	Kalau baragiah jo mambari, indak ma harok baleh jaso.	kalau dikasih dan memberi, tidak mengharap balas jasa.
13	Bialah Allah sajo nan manilai kasado nyo.	biarlah allah saja yang menilai semuanya.
14	Jariah nan indak di kana lai, bia nan satu manilai nyo.	jasa yang tidak diingat, biarlah yang satu yang menilainya.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengantarkannya ke sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1	1. Sungguh pun kini awak sedang berang, tatap aliak an senyum jo tutua kato nan elok.	1. sungguh pun sekarang kita sedang marah, tetap perlihatkan senyum dan tutur kata yang baik.
2	2. Jikok suak banyak awak mangarajoan nan elok, tapi indak diharagoi urang, jan ba ketek hati	2. jika besok kita banyak mengerjakan yang baik, tapi tidak dihargai orang, jangan berkecil hati
3	3. Indak barajo ka hati surang.	3. tidak ber raja ke hati sendiri
4	4. Kayo ka tampek rang batenggang.	4. kaya jadi tempat orang minta bantuan
5	5. Kalau nio mamintak tolong pakaro ekonomi, carilah urang nan elok sarato harato no baik.	5. kalau mau meminta tolong perkara ekonomi, carilah orang yang baik serta hartanya baik.
6	6. Alu ta taruang patah tigo, samuik ta pijak indak mati.	6. batang tersandung patah tiga, semut terinjak tidak mati
7	7. Cadiak ka bakeh rang batanyo.	7. pintar untuk tempat orang bertanya.
8	8. Baitu juo wakatu mamintak pandapek atau batanyo, carilah urang nan cadiak dalam bidang no.	8. begitu juga waktu meminta pendapat atau bertanya, carilah orang yang pintar dalam bidangnya.
9	9. Ameh bukan perak pun bukan	9. emas bukan perak pun bukan
10	10. Sarato katiko mamutuihan suatu aturan.	10. serta ketika memutuskan suatu aturan.
11	11. Jikok nio di haragoi urang, mako paliharo lah budi jo laku.	11. jika mau dihargai orang, maka peliharalah budi dan laku
12	12. Ijan sampai babuek musyrik, manyekutukan Allah.	12. jangan sampai berbuat musyrik, menyekutukan allah
13	13. Budi saketek rang haragoi.	13. budi sedikit orang hargai
14	14. Dek kawan jadi panutan.	14. karena kawan jadi panutan
15	15. Karano tolok ukua panilaian indak karano harato pangkaek, ameh atau perak.	15. karena tolak ukur penilaian bukan karena harta, pangkat, emas atau perak
16	16. Di nan dangka nyo tak ba riak.	16. di yang dangkal dia tidak bergelombang
17	17. Jadi lah urang nan disanangi jo disukoi urang banyak.	17. jadilah orang yang disenangi dan disukai orang banyak
18	18. Sarato ijan lupu ka pado Allah subhanahu wa ta ala.	18. serta jangan lupa kepada allah subhanahu wa ta ala
19	19. Ka nan kuaso inyo ma mintak.	19. kepada yang kuasa dia meminta.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengemukakan sumber dan menyebutkan nama penulisannya.
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

01	Baik itu karano awak santiang, elok laku	baik itu karena kita pintar berperilaku baik.
02	Karano hanyo Allah lah tampek mamintak jo nan mangabulkan kandak.	karena hanya allah lah tempat meminta dan mengabulkan keinginan.
03	Katiko bakarajo jo sia sajo, tatap lah bapacik ka dasar kejujuran, manjunjuang aturan agamo, antu ma nan batua jo nan salah.	ketika kerja dengan siapa saja, tetaplah berpegang pada dasar kejujuran, menjunjung aturan agama, tahu mana yang benar dan yang salah
04	Kang kampuang sayang kasadonyo.	orang kampung sayang semuanya.
05	Iman nyo taguah ba pandirian.	imannya teguh berpendirian.
06	Di nan dalam tak ba galombang.	di yang dalam tidak bergelombang
07	Baa bana barek nyo ujian hiduik nan di tamui, tatap lah mamintak tolong kapado Allah subhanahu wa ta ala.	bagaimana pun beratnya hidup yang ditemui, tetaplah meminta tolong kepada allah subhanahu wa ta ala.
08	Karano itu lingkungan bisa manarimo awak.	karena itu, lingkungan bisa menerima kita.
09	Indak mangecek ba hati nan kusuik.	tidak mengatakan dengan hati kusut
10	Walau harimau di dalam paruik, kambing juo nan dikalua kan	walau harimau didalam perut, kambing juga yang dikeluarkan.
11	Muko nan janiah di nampak an.	muka yang jernih diperlihatkan
12	Ba a bana urang lain mampalakukan awak, jadi lah urang nan punyo jiwa panyaba jo lamah lambuk ka sado urang.	bagaimana orang lain memperlakukan kita, jadilah orang yang punyo jiwa penyabar dan lemah lembut kesemua orang.
13	Kok bakiek alah bajajak.	kalau berniat telah ber bekas
14	Maso ketek adolah maso-maso nan banyak pangajaran sopan jo santun.	masa kecil adalah masa-masa yang banyak pelajaran sopan dan santun
15	Di tengah kamajuan zaman nan sangaik capek ko, jadi lah laki-laki nan tatap manjunjuang tinggi pakaro adaik, tradisi jo pegangan hiduik salam bajalan di ateh bumi.	ditengah kemajuan zaman yang sangat cepat ini, jadilah laki-laki yang tetap menjunjung tinggi perkara adat, tradisi dan pegangan hidup selama berjalan diatas bumi.
16	Pituah itu pun bisa dijadikan pedoman, tolak ukua jo pertimbangan dalam sagalo hal nan ka mereka hadok isuak.	petuah pun bisa dijadikan pedoman, tolak ukur dan pertimbangan dalam segala hal yang akan mereka hadapi esok.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengantumkan sumber;

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Habis tahun berganti musim, sandi adat jan di anjak	habis tahun berganti musim. Aturan adat jangan ditinggalkan
Karena itu, pituah Minang untuak anak-anak sangat acok awak danga disampaikan ka anak- anak nan dari urang-urang tuo.	karena itu, petuah Minang untuk anak- anak sangat cocok kita dengar disampaikan ke anak-anak dari orang-orang tua.
itu karena wakatu maso ketek tu, satiok anak punyo kemampuan untuak marekam jo mangana apo nan diliek nyo ataupun nan didanga no.	itu karena masa kecil itu, setiap anak punya kemampuan untuk merekam dan mengingat apa yang dilihatnya ataupun yang didengarnya.
Nan partamo usah manjadi rajo kalah, nan bajalan manakua nakua.	yang pertama jangan menjadi raja kalah, yang berjalan membungkuk-bungkuk.
Jalin kulindan ateh bumi, pa arek buhua jo manusia.	jalin keakraban atas bumi, per erat ikatan dengan manusia
Anak nan baru tumbuh, handak nyo di aja kan baa caro mandakek an diri kapado Tuhan.	anak yang baru tumbuh, hendaknya diajarkan bagaimana cara mendekati diri kepada tuhan
Jadi lah anak nan bisa manarimo budayo lua, namun tatap manjunjuang tinggi nilai-nilai budayo asli.	jadilah anak yang bisa menerima budaya luar, namun tetap menjunjung tinggi nilai-nilai budaya asli.
Mangecek ta tagun-tagun hilang banso punah martabat.	berkata ter tegun – tegun, hilang bangsa punah martabat.
Tapi ado duo nan tak ba hampun, partamo murta, ka duo syirik.	tapi ada dua yang tak diampuni, pertama murta, kedua syirik.
Nan kaduo rantangkan banang banang ka langik, hubungkan diri jo Tuhan	yang kedua rentang kan benang ke langit, hubungkan diri dengan tuhan.
Manuik sipaik jo janji-janji-Nyo Allah itu maha pangampun murah maaf, maha jo dandam.	menurut sifat dan janji-janinya, allah itu maha pengampun murah maaf, mahal dengan dendam
Karena Allah lah satu-satu nyo nan patuik di sambah manusia	karena allah satu-satunya yang patut disembah manusia
Samaso hiduik, jan pernah manyekutukan Allah subhanahu wa ta ala.	semasa hidup, jangan pernah menyekutukan allah subhanahu wa ta ala.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis atau dengan cara lain untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1	Hiduih lah di ateh dunia, ba pacik kapado al-ur'an jo sunnah Rasulullah.	hiduplah diatas dunia, berpegang kepada al-qur'an dan sunnah rasulullah.
1	Arato manyaimbangan jo hubungan nan elok antaro manusia.	serta penyeimbangan dengan hubungan yang baik antar manusia
1	Karajokan apo nan taragak, tapi ingek satiok atamuhan ado papisahan.	kerjakan apa yang diinginkan, tapi ingat setiap pertemuan ada perpisahan
2	Dan juo, jan pernah ta dayo jo apo pun nan bisa manumbuahkan niaik untuak murtad atau pindah dari agama islam.	dan juga, jangan pernah terpedaya dengan apapun yang bisa menumbuhkan niat untuk murtad atau pindah dari agama islam.
2	Sadane di mano kito kini, kamano arah ka di tuju.	sedang dimana kita sekarang, kemana arah yang dituju
2	Cintoi apo nan dititipkan ka awak.	cintai apa yang di titipkan ke kita.
2	Kok indak tau jawabannyo, tuhan manolong manjawabkan.	kalau tidak tahu jawabannya, tuhan menolong menjawabkan.
2	Salamo awak masih ketek, sadari cieik hal nan sangaik pantiang	selama kita masih kecil, sadari satu hal yang sangat penting
2	Satiok manusia nan hiduih di dunia ko, indak diciptakan kecuali untuak manyembah Allah subhanahu wa ta ala.	setiap manusia yang hidup didunia ini, tidak diciptakan kecuali untuk menyembah allah subhanahu wa ta ala.
2	Mako dari itu, jadi lah hamba nan taruih ba sarah diri sa ato mangaduan sagalo nyo, baiak sanang maupun sadiah kapado allah subhanahu wa ta ala	maka dari itu, jadilah hamba yang terus berserah diri dan mengadukan segalanya, baik senang maupun sedih kepada allah subhanahu wa ta ala.
2	Buruak muko, camin di balah.	buruk wajah, cermin dibelah.
28	Bagu kapalang aja, nan bak bungo kembang tak jadi.	berguru kepalang ajar, yang seperti bunga kembang tidak jadi.
29	Suatu ilmu nan ndak salasai di palajari, kurang bisa di pagunoan.	suatu ilmu yang tidak selesai dipelajari, kurang bisa digunakan.
30	Awak nan mambuek kasalahan, tapi manyalahkan urang lain	kita yang membuat kesalahan, tapi menyalahkan orang lain



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ilmu hanya bisa di dapek jo baguru kamuliaan	ilmu hanya bisa didapatkan dengan berguru,
hanyo didapek jo budi nan tinggi	kemuliaan hanya bisa didapatkan dengan budi yang tinggi.
Karano ma utamo an suatu urusan nan kurang saralo sahinggo nan paralu bana ta tingga karano hanyo	karena mengutamakan urusan yang kurang diperlukan sehingga yang sangat diperlukan tertinggal karenanya
Upayo pandai rajin baguru supoyo tinggi naiak an buci	supaya pintar rajib berguru, supaya tinggi naikkan budi
Sifaik urang nan indak suko ba taruih tarang sarato indak bijak dalam ma ambiak kaputusan	sifat orang yang tidak suka berterus-terang serta tidak bijak mengambil keputusan
Samoga pambahasan tentang pituah Minangkabau ko bisa ba manfaik jo ma agiah pangaruah nan elok untuak pajalanan hiduik kito	semoga pembahasan tentang petuah Minangkabau ini bisa bermanfaat serta memberi pengaruh positif untuk perjalanan hidup kita
Sahinggo itulah pambahasan mangenai kumpulan pituah Minang untuak padusi laki laki sarato anak anak ketek	sehingga itulah pembahasan mengenai kumpulan petuah Minang untuk perempuan, laki-laki, serta anak-anak kecil.
Lai sa urang , urang nan tun bansaik bana, ba rumah ba ladang indak.	ada se orang, orang itu miskin betul, punya rumah, punya ladang tidak.
Urang nan tun karajo tiok hari mancukua rambuik urang.	orang itu kerja tiap hari mencukur rambut orang.
Alah sudah inyo mancukua urang dibari malah di urang gak sa uang duo uang bara rilah no.	setelah selesai dia mencukur orang, di beri lah oleh orang agak satu uang dua uang berapa ikhlas nya.
Timbulah musim lapa dalam nagari nan tun.	datang lah musim lpar dalam desa itu.
Jadi, banyak lah urang nagari nan tun habih mati kalaparan.	jadi, banyak lah orang di desa itu habis mati kelaparan.
Hati urang bansaik nan tun susah bana maagaki anak no nan tun.	hati orang miskin itu susah sekali memikirkan anaknya itu.
Ingek pikia lah anak nan tuo, kini bia lah nak den bajalan di nagari nangko mencari pungguang den nan tak basaok, paruik den nan tak makan.	ingat pikir lah anak yang tua, sekarang biar lah mau saya berangkat dari desa ini mencari pungguang



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengacantumkan sumber;
2. Dilarang mengutip hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
3. Dilarang mengutip dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	saya yang tidak bertutup, perut saya yang tidak makan.
4	jadi, dikatakan molah bakeh bapak no, jadi, dikatakan lah kepada bapak nya, sejauh panjang mukasuk no nan tun. maksud nya itu.
4	jadi, dilepehlah anak no nan tun bajalan. jadi, dilepaslah anak nya itu berangkat.
4	Bajalan anak nan tun masuk rimbo kalua rimbo, berangkat anak itu masuk hutan keluar hutan, lah ta sasek io ka rumah urang panyamun. sudah ter sesat iya ke rumah orang perampok.
4	Lah nampak di urang panyamun anak nan tun, sudah tampak oleh orang perampok anak itu, batano lah panyamun nan tun. bertanya lah perampok itu.
4	Bakato pulo panyamun nan tun, “ Isuak malah berkata juga perampok itu,” besok lah dapat yang dapek nan ang cari, singgah ka rumah den .” kamu cari, mampir ke rumah saya .”
4	Jano paja tu, “ Jadih.” kata anak itu ,” baiklah .”
5	Ditahani no lah ba malam di rumah no di ditahan nya lah ber malam di rumah nya oleh panyamun nan tun. perampok itu.
5	Jawek no, “ Awakden ka pai mancarikan paruik jawab nya, “ saya mau pergi mancarikan perut saya den nan tak ba isi, kapalo den nan tak batutuik .” yang tidak ber isi, kepala saya yang tidak bertutup.”
5	Jadi di bari no lah anak nan tun di urang gaek tu jadi di beri nya lah anak itu oleh orang tua itu satangkai baringin ameh. setangkai beringin emas.
5	Baringin ameh nan tun kok dikuncang banyak beringin emas itu kalau digoyang banyak emas ameh jatuh
5	Tengah malam di tuka no dek urang panyamun tengah malam di tukar nya oleh orang perampok itu tu baringin no nan tun. beringin nya itu.
5	Lah tibo no di rumah dihimbau no lah biai no jo sudah tiba dia di rumah di panggil nya lah ibu nya apak o. dan bapak nya.
56	Jano, “Ko to baringin ameh nan den cari.” katanya ,” ini lah beringin emas yang saya cari.”
57	Jano biai no,” Cubo lah kalua kan ameh no.” kata ibu nya,” coba lah keluar kan emas nya .”
58	Usah lai ameh nan ka jatuh daun no lai tidak jangankan lagi emas yang akan jatuh daun nya pun rareh. tidak rontok.”
59	Jadi manangih – nangih io. jadi menangis – nangis ia.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengemukakan sumber:
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

61	Jano, "Bia lah nak den pai pulo, ayah, mencari an bakcando di tuan den nan tun."	katanya ," biar lah akan saya pergi juga, ayah,mencari yang seperti di kakak (laki-laki) saya itu."
62	Tu lah bajalan pulo nan tengah.	dan telah berangkat juga yang tengah."
63	Kato nyo pulo urang panyamun tu, " Ka kama buyuang?"	kata nya pula orang perampok itu," mau kemana nak?"
64	Jawek no, " Awakden ka pai mancarikan pungguang den nan tak basaok, paruik den nan tak barisi."	jawab nya ," saya mau pergi mancarikan punggung saya yang tidak bertutup, perut saya yang tidak berisi."
65	A kato urang panyamun tu pulo, " Isuak malah dapek nan ang cari tu, singgah ang ka rumah den muah."	apa kata orang perampok itu juga ," besok kalau dapat yang kamu cari itu, mampir kamu ke rumah saya ya."
66	Batanyo lah rang gaek nan tun, " Kama ang buyuang?"	bertanya lah orang tua itu, " kemana kamu nak?"
67	Jawek no , " Ambo mancarikan paruik den nan tak makan, pungguang den nan tak basaok."	jawab nya," saya mancarikan perut saya yang tidak makan, punggung saya yang tidak bertutup."
68	Peti nan tun tiok – tiok disingkok barisi nasi langkok jo gulai no.	peti itu tiap – tiap di buka berisi nasi lengkap dengan gulai nya.
69	Jano panyamun nan tun, " Lah suruik mo ang?"	kata perampok itu, " sudah kembali ya kamu?"
70	" Alah" jano nan tengah nan tun.	" sudah ," kata yang tengah itu.
71	Batan o lah panyamun nan tun , Ma to nan ang cari ?	bertanya lah perampok itu , mana kah yang kamu cari ?"
72	Kato nan tengah nan tun, " Iko mah peti, peti nanko tiok – tiok disingkokkan barisi nasi langkok jo gulai no."	kata yang tengah itu, " ini lah peti, peti ini tiap – tiap di buka berisi nasi lengkap dengan gulai nya."
73	Tu lah disingkokkan no peti nan tun.	lalu sudah di buka nya peti itu.
74	Lah di pakalua – kalua kan no nasi jo gulai.	sudah di keluar – keluar kan nya nasi dan gulai.
75	Bih makan lah awakno.	pada makan lah dia.
76	Lah s'lah minum jo makan bih tidua lah io.	sudah selesai minum dan makn pada tidur lah ia.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengantumkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tengah malam lah ditukai no pulo peti paja nan un.	tengah malam sudah di tukar nya pula peti anak itu.
Lah siang hari bajalan pulo paja nan tun mambao peti no pulang.	setelah siang hari berangkat pula anak itu membawa peti nya pulang.
Kato no, “ Iko nan den cari lah dapek sabuah peti peti nanko tiok di singkok barisi nasi jo gulai.”	kata nya ,” ini yang saya cari sudah dapat sebuah peti, peti ini tiap di buka berisi nasi dan gulai.”
Jano biai no, “ Mujua, singkok lah .”	kata ibu nya, “ mujur, buka lah “
Baru disingkokkan no samiang jadi manangih pulo lah anak nan tengah nan tun.	baru dibukakan nya seketika jadi menangis pula lah anak yang tengah itu.
Lah tiba pulo di rumah urang panyamun nan tun juo.	sudah tiba pula di rumah orang perampok itu juga.
Lah tampak pulo di urang panyamun tu.	sudah tampak pula oleh orang perampok itu.
Jawek no, “ ambo pai mancarikan pungguang nan tak basaok, paruik nan tak makan.”	jawab nya, “ saya pergi mancarikan pungguang yang tidak bertutup, perut yang tidak makan.
Jano panyamun tu pulo,” Malah dapek nan ang cari, singgah juo ka rumah den.”	kata perampok itu pula,” kalau dapat yang kamu cari, mampir juga ke rumah saya.”
“ Jadih,” jano nan bungsu tu.	“ baiklah ,” kata anak bungsu itu.
Tungkek nan tun pandai malacuik sandirino.	tongkat itu bisa mencambuk sendiri.
Jadilah babaliak nan bungsu sampai pulang.	jadilah berbalik yang bungsu sampai pulang.
Tu baru ka ditukai no dilacuik no di tungkek nan tun.	lalu baru mau ditukar nya di cambuk nya oleh tongkat itu.
Lah siang hari jago lah si bungsu nan tun, di caliak no urang panyamun lah ta bariang surang di sampiang no.	sudah siang hari, bangun lah si bungsu itu, di lihat nya orang perampok sudah ter baring sendiri di samping nya.
Jadi, takuik lah no, dalam hati no, urang maliang malah ka rumah tadi.	jadi, takut lah dia, dalam hati nya, orang maling sudah ke rumah tadi.
Karan takana di no bana, kok baitu nak manukai tungkek den malah awakno tadi ko, dilacuikno di tungkek den, mati.	karena teringat oleh nya benar, kalau begitu hendak menukar tongkat saya berarti dia tadi ini, di cambuk nya oleh tongkat saya, mati.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip, sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kok iyo bak nan tun tantu lah baringin ameh tuan en, peti tuan den lai siko.	kalau benar seperti itu tentu lah beringin emas kakak (laki – laki) saya, peti kakak (laki – laki) saya ada disini.
Di kuncang no baringin nan tun lah jatuh ameh.	di goyang nya beringin itu sudah jatuh emas.
Di sinkok no peti nan tun lah ado nasi jo gulai.	di buka nya peti itu sudah ada nasi dan gulai.
Udah nan tun lah pulang inyo dibawok an no lah baringin , peti jo tungkek no nan tun.	setelah itu sudah pulang dia di bawa nya lah beringin, peti dan tongkat nya itu.
Lah tibo io di rumah di paragokan no lah bakeh biaino, bakeh apak no jo bakeh rangsanak no.	setelah tiba di rumah di perlihatkan nya lah kepada ibu nya , kepada bapak nya dan kepada saudara nya
Jadi urang bansaik cako kayo lah.	jadi orang miskin tadi kaya lah.
Angko – angko dalam bahaso Minang.	angka – angka dalam bahasa Minang
Ciek, duo, tigo, ampek, Limo, Anam, Tujuh, salapan, sambilan, sapuluah, sabaleh, saratuih	satu, dua, tiga, empat, lima, enam, tujuh, delapan, sembilan, sepuluh, sebelas, seratus.
Sinayan, salasa, Raba ‘ a, kamih, Juma’aik, Satu, akaik	senin, selasa, rabu, kamis, jumat, sabtu, ahad
Urang Minang bisa di tamui di ma sajo di kasado panjuru indonesia.	Orang Minang bisa ditemui diseluruh indonesia
Pituah indak samo jo pepatah.	Petuah tidak sama dengan pepatah.
Contoh pituah bahaso Minang nan paliang tanamo adolah	Contoh petuah Minang yang paling terkenal adalah.
Ado pituah nan untuak padusi, laki – laki dan juo anak-anak.	Ada petuah untuk perempuan, laki-laki, dan juga anak-anak.
Kito bahas nan partamo pituah Minang untuak anak padusi.	Kita bahas yang pertama petuah Minang untuk anak perempuan.
Jikok datang laki dari jauh, sambuik lah jo muko manih, hidangan lah minum jo makan nyo	Jika suami datang dari bepergian jauh, sambutlah dengan muka manis, hidangkanlah minum dan makannya.
Mangko dari itu, inyo harus dilayani.	Karna itu, dia harus dilayani



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kalau padusi di tengah nan rami, apolagi banyak kaum laki – laki di situ, handak no paliharo mato o laki.	Kalau perempuan ditengah orang ramai, apalagi banyak kamu laki-laki disana, hendaknya menjaga mata dan perilaku.
Handak no tatap dalam bateh patuik.	Hendaknya tetap dalam batas wajar
an malu untuak ba tanyo kalau ado nan indak mau.	Jangan malu untuk bertanya kalau ada yang tidak tahu.
Dalam suko jo sadiah.	Dalam suka dan sedih
Salanjuiknyo pituah Minang untuak anak – anak.	Selanjutnya petuah Minang untuk anak-anak.
Di antaro pituah nan acok di disampaikan adolah.	Diantara petuah yang sering disampaikan adalah.
Salain pituah Minang untuak padusi jo laki – laki, ado juo untuak anak ketek.	Selain petuah Minang untuk perempuan dan laki-laki, ada juga untuk anak kecil
Nan kalimo Dari mano datang nyo awak	Yang kelima dari mana dating nya kita
Tapi harus juo di kana, kalau satiok patamuan , pasti ado papisahan.	Tapi harus juga dikenang, kalau setiap pertemuan, pasti ada perpisahan.
Batanyo ka diri, untuak apo waang ko hiduik?.	Bertanya ke diri sendiri, untuk apa kamu ini hidup?
Ka manga datang ka dunia ko?	Mau apa datang ke dunia ini?
Adopun Pituah Minang nan lain antaro no.	Adapun petuah Minang yang lain diantaranya.
Urang nantun baranak tigo, laki – laki ka tigo no.	Orang itu beranak tiga, laki – laki ke tiga nya.
Di cari no urang nan ka di cukua no dalam nagari nantun indak pulo ado.	Di cari nya orang yg akan di cukur nya dalam desa itu, tidak juga ada.
Pikiran apak no, daripado mati kalaparan nampak di mato den bialah nak iyo bajalan.	Pikiran bapak nya, daripada mati kelaparan nampak di mata saya biarlah mau iya berangkat.
Jano, ka kama ang buyuang?”	Kata nya , “ Mau kemana kamu nak?”
Jawek buyuang tu, “ ambo pai mancarikan punggung nan indak basaok, paruik ambo nan indak makan “.	Jawab anak itu, “ Saya pergi mancarikan punggung yang tidak bertutup, perut saya yang tidak makan.”
Ba malam lah paja tu di sinan, di bari no minum makan.	Ber malam lah anak itu di sana, di beri nya minum makan.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengcantumkan sumber;
2. a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penunjang akademik;
2. b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2	Lah siang lah bajalan awakno ka rumah urang gaek.	Dinihari sudah berangkat dia ke rumah orang tua.
2	Kato urang gaek nantun, “ Ka pai kama ang buyuang?”	Kata orang tua itu,” Mau pergi ke mana kamu nak (laki – laki).”
2	Jadi pulang lah no, sampai pulo awakno di urang panyamun tu.	Jadi pulang lah dia, sampai juga dia di orang perampok itu.
2	Lah ba malam awakno di sinan.	Sudah ber malam dia di sana.
3	Lah siang hari bajalan lah awakno pulang.	Sudah siang hari berjalan lah dia pulang.
3	Di kuncang no lah baringin nantun.	Di goyang nya lah beringin itu.
3	Bakato pulo anak no nan tengah ka ayah no.	Berkata pula anak nya yang tengah ke ayah nya.
3	Lah tibo io di rumah urang panyamun nantun.	Sudah tiba ia di rumah orang perampok itu.
3	Lah bajalan paja tu, lah sampai pulo no ka rumah rang gaek nantun.	Sudah berangkat anak itu, sudah sampai juga dia ke rumah orang tua itu.
3	Jadi di bari urang gaek tu awakno sabuah peti.	Jadi di beri orang tua itu dia sebuah peti.
3	Tu lah pulang pulo nan tengah nantun.	Lalu sudah pulang pula yang tengah itu.
3	Lah tibo pulo no di rumah urang panyamun cako.	Sudah tiba pula dia di rumah orang perampok tadi.
3	Lah bamalam pulo nan tengah nantun di sinan.	Sudah bermalam pula yang tengah itu di sana.
3	Lah tibo no di rumah lah dihimbau no biai no jo apak no .	Sudah tiba dia di rumah sudah di panggil nya ibu nya dan bapak nya.
4	Jadi pai pulo lah anak no nan bungsu.	Jadi pergi pula lah anak yang bungsu.
4	Jano panyamun tu, “ Ka kama ang buyuang?”	Kata perampok itu, “ mau kemana kamu nak?”
4	Lah bajalan pulo io , lah sampai pulo bakeh urang gaek nan tun.	Sudah berangkat pula ia, sudah sampai pula kepada orang tua itu.
43	Batanio pulo urang tuo tu,” Ka kama ang buyuang?”	Bertanya pula orang tua itu,” mau kemana kamu nak?”
44	“ Indah, ambo mencari nan tidak di ambo .”	“ Tidak , saya mencari yang tidak di saya .”
45	Tu lah di bari no pulo dek urang gaek nantun awakno sabuah tungkek.	Lalu sudah di beri nya pula oleh orang tua itu dia sebuah tongkat.
46	Lah sampai di rumah panyamun nantun lah bamalamlah io di sinan.	Setelah sampai di rumah perampok itu sudah bermalam lah ia di sana.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengizinkan penggunaan atau seluruhnya dalam bentuk apapun tanpa izin UIN Suska Riau.
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penerbitan, penulisan, penyusunan laporan, penyusunan sumber.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4	Tengah malam lah dituka i pulo di panyamun antun.	Tengah malam sudah di tukar pula oleh perampok itu.
4	Mati lah panyamun nantun.	Mati lah perampok itu
4	Jadi dicari no lah, lah basuo di no baringin jo peti antun.	Jadi dicari nya lah, sudah bertemu oleh nya beringin dan peti itu.
5	Nama hari dalam bahaso Minang	Nama hari dalam bahasa Minang
5	Ukua bara awak pai ?	Jam berapa kita pergi ?
5	Alah makan jo minum bagai .	Sudah makan dan minum juga.
5	Mo lah makan awak lai.	Ayo lah makan kita lagi.
5	Paruik ambo litak bana.	Perut saya lapar sekali.
5	Taraso auih badan ambo kini .	Terasa haus tubuh saya sekarang.
5	Sia namo sanak ?	Siapa nama saudara ?
5	Dima angku tingga ?	Di mana kamu tinggal ?
5	Apo suku angku ?	Apa suku Kamu ?
5	Nio pai kama sanak ?	Mau pergi kemana saudara?
6	Apo nan sedang angku karajoan kini ?	Apa yang sedang kamu kerjakan sekarang?
6	Lai rancak baju ko mak?	Ada bagus baju ini bu ?
6	Baa kaba Amak, lai sehat sajo?	Bagaimana kabar ibu, ada sehat saja?
6	Ka sina miang wak.	Ke sana saja kita.
6	Kami turun di siko sajo.	Kami turun di sini saja.
6	Kau kamek, rancak, manih.	Kamu imut, cantik, manis.
6	Angku gagah.	Kamu ganteng.
6	Sadang manga Ayah kini ?	Sedang ngapain ayah sekarang?
68	Mandeh wak suko mambuek kue ko.	Ibu saya suka membuat kue ini.
69	Etek bilo tibo?	Tante kapan datang?
70	Mamak suko bana naiak onda ko.	Om suka sekali naik motor ini.
71	Etek ko elok bana.	Tante ini baik sekali.
72	inyo nan ambo nan paliang tuo.	Dia kakak (perempuan) saya yang paling tua.
73	Uda ambo bakarajo di sinan.	Kakak (laki – laki) saya bekerja di sana.
74	Adiak padusi ambo masih ketek.	Adik perempuan saya masih kecil.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip, sebarkan, atau menyalin seluruh atau sebagian isi tanpa mengizinkan sumber:
 - a. Pengutipan hanya untuk kepentingan penelitian, penulisan, penerbitan, atau penyusunan laporan, atau untuk tujuan pendidikan, penelitian, atau penyusunan laporan, atau untuk tujuan pendidikan, penelitian, atau penyusunan laporan.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

77	Bukak lah pintu ko.	Buka lah pintu ini .
77	tingkok lah tuduang tu.	Buka lah payung itu.
77	Apak no mambali oto di siko cako.	Bapak nya membeli mobil di sini tadi.
77	Alah lah lantai no.	Sudah bersih lantai nya.
77	amba alah ta latak di ateh meja.	Lauk sudah ter letak di atas meja.
88	etek, agiah ambo pituluik ciek.	Tante, beri saya pena satu.
88	Ambo pai mauruik kaki nan sakik ko.	Saya pergi memijit kaki yang sakit ini.
88	Ano manulih di karateh tun.	Dia menulis di kertas itu.
88	Tas no warna kulabu.	Tas itu warna abu – abu.
88	Iko ubek untuak sakik kapalo.	Ini obat untuk sakit kepala.
88	Tolong ambiak kan karateh tu.	Tolong ambil kan kertas itu.
88	Etek mamakai tilakuang ka masajik cako.	Tante memakai mukena ke mesjid tadi.
88	Kurisi no alah ba susun.	Kursi nya sudah di susun.
88	Inyiak ka pakan jo kareta angin.	Kakek ke pasar dengan sepeda.
88	Ano sedang mamakan kacang abuih.	Dia sedang memakan kacang rebus.
99	Tarompa tu alah sampik.	Sandal itu sudah sempit.
99	Muko nyo layua sajo.	Wajah nya lesu saja.
99	Tabek nan ba paga sirah.	Kolam yang ber pagar merah.
99	Inyo samangaik bana nak bakarajo.	Dia semangat sekali mau bekerja.
99	Ubek sagalo panyakik adolah hati nan senang.	Obat segala penyakit adalah hati yang senang.
99	Sadiah mandeh, sadiah denai juo.	sedih ibu, sedih saya juga.
99	Ambo sedang balanjo di pakan akaik.	Saya sedang belanja di pasar minggu.
99	Mandeh sedang pai baralek.	Ibu sedang pergi pesta.
98	Bisuaak tibo lah ka rumah ambo.	Besok datang lah ke rumah saya.
99	Badan nyo rampiang, rancak di caliak.	Tubuh nya ramping, cantik di lihat.
00	Ambo takantuak bana dek batanggung.	Saya mengantuk sekali karena begadang.
01	Lah senang hiduik angku nampak di ambo.	Sudah senang hidup kamu terlihat oleh saya.
02	Randang nyo sabana lamak.	Rendang nya sangat enak.
03	Buan lah sarok di tampek nyo.	Buang lah sampah di tempat nya.
04	Bisuaak nyo barangkek ka Banduang.	Besok dia berangkat ke Bandung.



LAMPIRAN B

Dataset Bahasa Minang

Abuiah	Acok	Adaik	Adat
Adiak	Ado	Adolah	Adopun
Agak	Agamo	Agiah	Agiahan
Aiah	Aja	Ajaran	Akaik
Akidah	Alah	Alek	Alim
Allah	Alu	Amak	Ambiak
Ambo	Ameh	Ampek	Ampiang
An	Anak	Anam	Ang
Anggan	Angin	Angko	Angku
Anjak	Ano	Antaro	Apak
Apo	Apolai	Arah	Arati
Area	Asli	Atau	Ataupun
Aten	Aturan	Auih	Awak
Awakden	Awakno	Ayah	Ba
Baa	Babagai	Babaliak	Babuek
Baan	Bagai	Baganduah	Baganti
Bagarah	Bagi	Baguru	Bahas
Bahaso	Baiak	Baitu	Bajajak
Bajalan	Baju	Bak	Baka
Bakamudi	Bakato	Bakcando	Bakeh

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Bateh	Bakisa	Bakumpua	Balah
Batuaku	Balanjo	Balayia	Baleh
Baliau	Bamacam	Bamalam	Bamalamlah
Banang	Banang	Banduang	Bansaik
Banso	Banyak	Bapacik	Bapak
Baraja	Baragiah	Baraja	Barajo
Baralek	Baranak	Barangkek	Barek
Bariang	Bariang	Baringin	Barisi
Barubah	Barubah	Barupo	Basakolah
Basaok	Basi	Basingguangan	Baso
Basuo	Batamu	Batani	Batanyo
Bateh	Batenggang	Batua	Batunggang
Batutuik	Bendi	Berang	Bia
Biai	Biaino	Bialah	Biaso
Bidang	Bijak	Bilo	Bini
Bisa	Bisuak	Buang	Budaya
Buayo	Budi	Buhua	Bujang
Bukak	Bukan	Buliah	Bumi
Bundo	Bungo	Bungsu	Buruak
Buyuang	Cacek	Cadiak	Cako
Calak	Camin	Capek	Cari
Carilah	Caro	Ciek	Cindua
Cisoi	Contoh	Cubo	Cukua
Cukrik	Dahulu	Dalam	Damai



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Danda	Dandam	Danga	Dangka
Dapek	Dari	Dasar	Datang
Daun	Dayo	Dek	Den
Denai	Dewasa	Di	Dibaok
Dibari	Dibukak	Dicari	Diciptakan
Dianga	Didapek	Diharagoi	Dihimbau
Dindian	Dikaluaan	Dikatokan	Dikuncang
Dilcuik	Dilapehlah	Dilayani	Dilieik
Dina	Dimaafkan	Dipabanyak	Dipalia
Diri	Disampaikan	Disanangi	Disiko
Disingkok	Disingkokkan	Disukoi	Ditahani
Dititipkan	Ditukai	Dituokan	Diusahoan
Duduak	Dulu	Dunia	Dunsanak
Duo	Ekonomi	Elok	Etek
Famili	Gadang	Gagah	Gak
Gambang	Garah	Golongan	Gulai
Hadiah	Hadok	Hak	Hal
Hadpa	Hampun	Handak	Hanyo
Haragoi	Harato	Hari	Harimau
Harinyo	Harok	Harus	Hati
Heran	Hidangkanlah	Hiduik	Hilang
Hilangkan	Hubuangan	Hubuangkan	Iduik
Ijan	Iko	Ilang	Ilimu
Ille	Iman	Indak	Indonesia

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ingek	Inyiak	Inyo	Io
Issam	Isti	Isuak	Itu
Itulah	Iyo	Izin	Jadi
Jadiah	Jadilah	Jago	Jalin
Jambi	Jan	Janiah	Janji
Jambak	Jariah	Jaso	Jatuah
Jauh	Jawabannyo	Jawek	Jikok
Jilid	Jiwa	Jo	Jumaaik
Jujur	Juo	Ka	Kaba
Kabiasaan	Kacang	Kaduo	Kahidupan
Kahormatan	Kajujuran	Kaki	Kakurangan
Kalah	Kalakuan	Kalamahan	Kalamahannyo
Kalapanan	Kalau	Kalimo	Kalua
Kama	Kamajuan	Kamampuan	Kamano
Kamari	Kambang	Kambiang	Kamek
Kami	Kamih	Kampung	Kamuliaan
Kapas	Kana	Kandak	Kandaraan
Kapaduang	Kapa	Kapado	Kapalang
Kapalo	Kapandaian	Kaputusan	Karajo
Karajoan	Karajokan	Karajonyo	Karano
Karateh	Kareta	Kasado	Kasadonyo
Kasalahan	Katan	Katiko	Kato
Kasaban	Katurunan	Kau	Kaum
Kasayan	Kayo	Kecuali	Keluarga

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ketek	Kian	Kini	Kito
Koko	Kok	Kuaso	Kucindan
Kue	Kulabu	Kulindan	Kumpulan
Kumbang	Kunci	Kurang	Kurisi
Kusuk	Labiah	Ladang	Lah
Laka	Lain	Lainnyo	Laki
Lakua	Lakuan	Lalu	Lamah
Lambak	Lambak	Lambang	Lambuik
Langik	Langik	Langkok	Langsuang
Lantai	Lapa	Lapang	Latak
Latiah	Lauik	Layua	Lia
Lidah	Limo	Lingkuangan	Litak
Lua	Lupo	Ma	Maaf
Mah	Maha	Mak	Makan
Makanannyo	Mako	Malacuik	Malah
Malam	Maliang	Malu	Mamahami
Mamak	Mamakai	Mambali	Mambao
Mambari	Mambina	Mambuek	Mambukak
Mamintak	Mampalakukan	Mamutuihan	Manaka
Manakua	Manangih	Manarangan	Manarimo
Mancaliak	Mancari	Mancarikan	Mancukua
Mandakek	Mandanga	Mandeh	Mandiri
Mandaaik	Manga	Mangabulkan	Mangaduan
Mangaluakan	Mangana	Mangarajoan	Mangecek



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Mangnai	Manggaleh	Mangicek	Mangko
Manilai	Manilai	Manimbang	Manjadi
Manjalago	Manjalani	Manjawabkan	Manjunjuang
Manlong	Manolong	Manukai	Manulih
Manumbuhkan	Manuntuik	Manuruik	Manusia
Manyimbangan	Manyalahkan	Manyambah	Manyasa
Manyekutkan	Manyeso	Marekam	Martabat
Masjik	Masih	Maso	Masuak
Masyarakaik	Mati	Mato	Maupun
Meja	Mereka	Miang	Minang
Minangkabau	Mintak	Mintuo	Minum
Minuman	Mo	Molah	Moral
Muah	Mudah	Muhrim	Mujua
Muko	Muliakan	Muluik	Murah
Murtad	Musim	Muslimah	Musyrik
Nagari	Naiak	Nak	Nakhoda
Nakua	Namo	Nampak	Namun
Nangih	Nangih	Nanko	Nasehat
Nasihaik	Nasihaik	Ndak	Niaik
Niniak	Niniak	Nio	No
Nyonyo	Nyo	Ombang	Onda
Oto	Pa	Pabukoan	Pacik
Pado	Padoman	Padusi	Paga
Pagangan	Pagunoan	Paham	Pai



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pajak	Pajalanan	Pakalua	Pakan
Pakaro	Palajari	Paliang	Paliharo
Pamahaman	Pamanih	Pandai	Pandang
Pandapek	Pandirian	Pangajaran	Pangampun
Pangaratian	Pangaruh	Pangkaek	Panilaian
Panguru	Pantiang	Panutan	Panyaba
Panyakik	Panyamangaik	Panyamun	Papatah
Pangolahan	Paragokan	Parahu	Parak
Papalu	Parangai	Parantau	Parintah
Pariuak	Partamo	Paruik	Pasak
Pasti	Patah	Patamuan	Patimbangan
Patinggi	Patuik	Perak	Pernah
Peti	Pijak	Pijakan	Pikia
Pikiran	Pindah	Pintu	Pitih
Pituah	Pituluik	Posisi	Pukua
Pulang	Pulo	Pun	Punah
Pungguang	Punyo	Quran	Rabaa
Ragam	Rajin	Rajo	Rambuik
Rami	Rapiang	Rancak	Randang
Rang	Rangsanak	Rantangkan	Rantau
Rapek	Rareh	Raso	Rasulullah
Riak	Rilah	Rimbo	Roman
Rubah	Rumah	Rumik	Rundiang
Rupanyo	Sa	Sabagai	Sabagainyo



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Sabalalah	Sabaleh	Sabana	Sabuah
Saparo	Sadang	Sadar	Sadari
Sadiah	Sadioan	Sado	Sagalo
Sahari	Sahinggo	Sajak	Sajo
Sakali	Saketek	Sakik	Sakitu
Salah	Salain	Salaintun	Salamo
Salanjuiknyo	Salapan	Salasa	Salasai
Salamo	Samaso	Samba	Sambah
Sambilan	Sambuik	Samiang	Samo
Samoga	Sampai	Sampaikan	Sampiang
Sampik	Samuik	Sanak	Sanang
Sanangan	Sandi	Sandirino	Sangaik
Sangko	Santiang	Santun	Sapanjang
Sapanuah	Sapeda	Sapuluah	Sarah
Sarato	Saratuih	Sarok	Sarupo
Sasok	Satangkal	Satiok	Satu
Sayang	Sehat	Senyum	Sesuatu
Sipaik	Siang	Sikap	Siko
Sipin	Sinayan	Singgah	Singkok
Sipaik	Sirah	Situ	Sopan
Suatu	Subhanahu	Sudah	Suko
Suki	Sumangaik	Sungguah	Sunnah
Suayo	Surang	Suruik	Susah
Susan	Syariat	Syirik	Ta

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Talihak	Tabek	Tadi	Tagaduah
Taman	Tahun	Tak	Takana
Takantuak	Takuik	Talok	Tamasuak
Tampak	Tampek	Tamui	Tanamo
Tangah	Tangan	Tanggo	Tantang
Tatau	Tanyo	Tapek	Tapi
Taragak	Tarang	Taraso	Taratik
Tarumpa	Taruang	Taruih	Tas
Tasebar	Tatap	Tau	Tentang
Tibo	Tidak	Tidua	Tigo
Timbua	Tingga	Tinggi	Tingkah
Tiok	Tolak	Tolong	Tradisi
Tu	Tuan	Tuduang	Tuhan
Tuju	Tujuh	Tujuan	Tuka
Tumbuah	Tun	Tungkek	Turun
Tudia	Uang	Ubah	Ubek
Uda	Ujian	Ukua	Ulama
Uleho	Umpamo	Umua	Umuanyo
Uni	Untuak	Upiak	Urang
Urusan	Usah	Usahlah	Usahoan
Utano	Wakatu	Walau	Warna
Wirasaho	Zaman		



LAMPIRAN C

Tabel C. 1 Hasil Pengujian Model Bahasa

Kata	Hasil Pengujian	Benar	Salah	Akurasi (%)
ABUIH	TARUIH	0	1	0.00
Acok	Acok	1	0	100.00
Adaik	Adaik	1	0	10.00
ADAT	HANDAK	0	1	0.00
Adiak	Adiak	1	0	100.00
ADO	PADO	0	1	0.00
Adolah	Adolah	1	0	100.00
Adopun	Adopun	1	0	10.00
Agak	HANDAK	0	1	0.00
Agama	Agama	1	0	100.00
Agiah	GAGAH	0	1	0.00
Agiahan	Agiahan	1	0	100.00
Aib	AUIH	0	1	0.00
Aja	PAJA	0	1	0.00
Ajaran	Ajaran	1	0	100.00
Akaik	PANGKAEK	0	1	0.00
Akidah	Akidah	1	0	100.00
Alah	Alah	1	0	100.00
Alek	Alek	1	0	100.00
Alim	PALIANG	0	1	0.00
Allah	ALAH	0	1	0.00
Alu	TANYO	0	1	0.00
Amak	LAMAH	0	1	0.00
Ambiak	SAMUIK	0	1	0.00
Ambo	Ambo	1	0	100.00
Ameh	Ameh	1	0	100.00
Ampek	Ampek	1	0	100.00
Ampiang	RAMPIANG	0	1	0.00
An	An	1	0	100.00
Anak	KALAH	0	1	0.00
Anam	TANTANG	0	1	0.00
Ang	Ang	1	0	100.00
Anggan	Anggan	1	0	100.00
Angin	Angin	1	0	100.00

© Hak cipta dimiliki oleh Universitas Islam Sumatera Utara. © Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Angko	Angko	1	0	100.00
Angku	ANGKO	0	1	0.00
Anjak	Anjak	1	0	100.00
Ano	TANYO	0	1	0.00
Antaro	Antaro	1	0	100.00
Apak	PA	0	1	0.00
Apo	PA KO	0	2	-100.00
Apolai	Apolai	1	0	100.00
Arah	BALAH	0	1	0.00
Arati	Arati	1	0	100.00
Arek	Arek TAS	0	2	-100.00
Asli	Asli	1	0	100.00
Atau	Atau	1	0	100.00
Ataupun	Ataupun	1	0	100.00
Ateh	AKIDAH	0	1	0.00
Aturan	Aturan	1	0	100.00
Auih	Auih	1	0	100.00
Awak	AWAK awak	0	1	0.00
Awakden	KO DEN	0	2	-100.00
Awakno	JUO	0	1	0.00
Ayah	CARILAH	0	1	0.00
Ba	BAA	0	1	0.00
Baa	Baa	1	0	100.00
Babagai	Babagai	1	0	100.00
Babaliak	Babaliak	1	0	100.00
Babuek	Babuek	1	0	100.00
Badan	DAN	0	1	0.00
Bagai	Bagai	1	0	100.00
Baganduah	Baganduah	1	0	100.00
Baganti	Baganti	1	0	100.00
Bagarah	Bagarah	1	0	100.00
Bagi	Bagi	1	0	100.00
Baguru	BAKCANDO	0	1	0.00
Bahas	Bahas	1	0	100.00
Bahaso	BAKA SUKO	0	2	-100.00
Baiak	Baiak	1	0	100.00
Baitu	Baitu	1	0	100.00
Bajajak	Bajajak	1	0	100.00
Bajalan	Bajalan	1	0	100.00
Baju	Baju	1	0	100.00
Bak	UBAH	0	1	0.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Baka	BAK KAN	0	2	-100.00
Bakamudi	Bakamudi	1	0	100.00
Bakato	Bakato	1	0	100.00
Bakcando	Bakcando	1	0	100.00
Bakeh	BATEH	0	1	0.00
Bakieh	Bakieh	1	0	100.00
Bakisa	Bakisa	1	0	100.00
Bakumpua	BAGANDUAH	0	1	0.00
Balah	Balah	1	0	100.00
Balaku	BALAH KO	0	2	-100.00
Balanjo	Balanjo	1	0	100.00
Balayia	Balayia	1	0	100.00
Baleh	Baleh	1	0	100.00
Baliau	Baliau	1	0	100.00
Bamacam	Bamacam	1	0	100.00
Bamalam	Bamalam	1	0	100.00
Bamalamlah	Bamalamlah	1	0	100.00
Bana	BANANG	0	1	0.00
Banang	Banang	1	0	100.00
Banduang	Banduang	1	0	100.00
Bansaik	Bansaik	1	0	100.00
Banso	Banso	1	0	100.00
Banyak	Banyak	1	0	100.00
Bapacik	Bapacik	1	0	100.00
Bapak	BA PATAH	0	2	-100.00
Bara	BALAH	0	1	0.00
Baragiah	Baragiah	1	0	100.00
Baraja	Baraja	1	0	100.00
Barajo	Barajo	1	0	100.00
Baralek	Baralek	1	0	100.00
Baranak	Baranak	1	0	100.00
Barangkek	Barangkek	1	0	100.00
Barek	Barek	1	0	100.00
Bari	Bari	1	0	100.00
Bariang	Bariang	1	0	100.00
Baringin	Baringin	1	0	100.00
Barisi	Barisi	1	0	100.00
Baru	Baru	1	0	100.00
Barubah	Barubah	1	0	100.00
Barupo	Barupo	1	0	100.00
Basakolah	Basakolah	1	0	100.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Basaok	Basaok	1	0	100.00
Basi	Basi	1	0	100.00
Basingguangan	Basingguangan	1	0	100.00
Baso	Baso	1	0	100.00
Basuo	Basuo	1	0	100.00
Batamu	BATANYO	0	1	0.00
Batani	Batani	1	0	100.00
Batanyo	Batanyo	1	0	100.00
Bateh	Bateh	1	0	100.00
Batenggang	Batenggang	1	0	100.00
Batua	BAK TUAN	0	2	-100.00
Batunggang	Batunggang	1	0	100.00
Batutuik	Batutuik	1	0	100.00
Bendi	Bendi	1	0	100.00
Berang	Berang	1	0	100.00
Bia	DIMA	0	1	0.00
Biai	DICARI	0	1	0.00
Biaino	DI LAINNYO	0	2	-100.00
Bialah	DI ALAH	0	2	-100.00
Biaso	DI RASO	0	2	-100.00
Bidang	DIDANGA	0	1	0.00
Bijak	PIJAK	0	1	0.00
Bilo	DI PULO	0	2	-100.00
Bini	KINI	0	1	0.00
Bisa	DI SALAH	0	2	-100.00
Bisuak	DISUKOI	0	1	0.00
Buang	UANG	0	1	0.00
Budaya	Budaya	1	0	100.00
Budayo	Budayo	1	0	100.00
Budi	Budi	1	0	100.00
Buhua	PUKUA	0	1	0.00
Bujang	Bujang	1	0	100.00
Bukak	DUO TAK	0	2	-100.00
Bukan	OTO KAN	0	2	-100.00
Buliah	Buliah	1	0	100.00
Bumi	Bumi	1	0	100.00
Bundo	Bundo	1	0	100.00
Bungo	Bungo	1	0	100.00
Bungsu	Bungsu	1	0	100.00
Buruak	DUDUAK	0	1	0.00
Buyuang	TU ILANG	0	2	-100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Golongan	Golongan	1	0	100.00
Gulai	Gulai	1	0	100.00
Habih	HARI	0	1	0.00
Hadok	Hadok	1	0	100.00
Hak	Hak	1	0	100.00
Hal	HAROK	0	1	0.00
Hamba	Hamba	1	0	100.00
Hampun	Hampun	1	0	100.00
Handak	Handak	1	0	100.00
Hanyo	Hanyo	1	0	100.00
Haragoi	Haragoi	1	0	100.00
Harato	Harato	1	0	100.00
Hari	Hari	1	0	100.00
Harimau	Harimau	1	0	100.00
Harinyo	Harinyo	1	0	100.00
Harok	Harok	1	0	100.00
Harus	Harus	1	0	100.00
Hati	Hati	1	0	100.00
Heran	QURAN	0	1	0.00
Hidangkanlah	Hidangkanlah	1	0	100.00
Hiduik	IDUIK	0	1	0.00
Hilang	Hilang	1	0	100.00
Hilangkan	Hilangkan	1	0	100.00
Hubuangan	JUNJUANGAN	0	1	0.00
Hubuangkan	SUKU ANGGAN	0	2	-100.00
Iduik	Iduik	1	0	100.00
Ijan	Ijan	1	0	100.00
Iko	DI KO	0	2	-100.00
Ilang	Ilang	1	0	100.00
Ilmu	LIMO	0	1	0.00
Ilmu	LAMO	0	1	0.00
Iman	Iman	1	0	100.00
Indak	Indak	1	0	100.00
Indonesia	Indonesia	1	0	100.00
Ingek	Ingek	1	0	100.00
Inyiak	DIMA	0	1	0.00
Inyo	Inyo	1	0	100.00
Iyo	IYO	0	1	0.00
Islam	Islam	1	0	100.00
Istri	Istri	1	0	100.00
Isuak	Isuak	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Itu	DI TUN	0	2	-100.00
Itulah	Itulah	1	0	100.00
Iyo	Iyo	1	0	100.00
Izin	Izin	1	0	100.00
Jadi	Jadi	1	0	100.00
Jadiah	Jadiah	1	0	100.00
Jadilah	Jadilah	1	0	100.00
Jago	Jago	1	0	100.00
Jalin	Jalin	1	0	100.00
Jamu	Jamu	1	0	100.00
Jan	Jan	1	0	100.00
Janiah	Janiah	1	0	100.00
Janji	Janji	1	0	100.00
Jano	Jano	1	0	100.00
Jariah	Jariah	1	0	100.00
Jaso	Jaso	1	0	100.00
Jatuah	Jatuah	1	0	100.00
Jauah	Jauah	1	0	100.00
Jawabannyo	Jawabannyo	1	0	100.00
Jawek	Jawek	1	0	100.00
Jikok	Jikok	1	0	100.00
Jilah	Jilah	1	0	100.00
Jiwa	Jiwa	1	0	100.00
Jo	Jo	1	0	100.00
Jumaaik	RUMAH PAI	0	2	-100.00
Junjuangan	Junjuangan	1	0	100.00
Juo	DUO	0	1	0.00
Ka	KALAH	0	1	0.00
Kaba	Kaba	1	0	100.00
Kabiasoan	Kabiasoan	1	0	100.00
Kacang	Kacang	1	0	100.00
Kaduo	Kaduo	1	0	100.00
Kahidupan	Kahidupan	1	0	100.00
Kahormatan	Kahormatan	1	0	100.00
Kajujuran	Kajujuran	1	0	100.00
Kaki	Kaki	1	0	100.00
Kakurangan	Kakurangan	1	0	100.00
Kalah	Kalah	1	0	100.00
Kalakuan	Kalakuan	1	0	100.00
Kalamahan	PA LAMAH KA NO	0	4	-300.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kalamahannyo	Kalamahannyo	1	0	100.00
Kalaparan	Kalaparan	1	0	100.00
Kalau	Kalau	1	0	100.00
Kalimo	Kalimo	1	0	100.00
Kalua	KANDUANG	0	1	0.00
Kama	Kama	1	0	100.00
Kamajuan	Kamajuan	1	0	100.00
Kamampuan	Kamampuan	1	0	100.00
Kamano	Kamano	1	0	100.00
Kamari	Kamari	1	0	100.00
Kambang	Kambang	1	0	100.00
Kambiang	SAMPIANG	0	1	0.00
Kamek	Kamek	1	0	100.00
Kami	Kami	1	0	100.00
Kamih	Kamih	1	0	100.00
Kampuang	Kampuang	1	0	100.00
Kamuliaan	Kamuliaan	1	0	100.00
Kan	Kan	1	0	100.00
Kana	Kana	1	0	100.00
Kandak	Kandak	1	0	100.00
Kandaraan	Kandaraan	1	0	100.00
Kanduang	Kanduang	1	0	100.00
Kapa	KA PA	0	1	0.00
Kapado	Kapado	1	0	100.00
Kapalang	Kapalang	1	0	100.00
Kapalo	Kapalo	1	0	100.00
Kapandaian	Kapandaian	1	0	100.00
Kaputusan	Kaputusan	1	0	100.00
Karajo	Karajo	1	0	100.00
Karajoan	Karajoan	1	0	100.00
Karajokan	Karajokan KA	0	1	0.00
Karajonyo	Karajonyo	1	0	100.00
Karano	Karano	1	0	100.00
Karateh	Karateh	1	0	100.00
Kareta	Kareta	1	0	100.00
Kasado	Kasado	1	0	100.00
Kasadonyo	Kasadonyo	1	0	100.00
Kasalahan	Kasalahan	1	0	100.00
Katan	Katan	1	0	100.00
Katiko	Katiko	1	0	100.00
Kato	Kato	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Katoan	Katoan	1	0	100.00
Katurunan	Katurunan	1	0	100.00
Kau	Kau	1	0	100.00
Kaum	DAUN	0	1	0.00
Kawan	Kawan	1	0	100.00
Kayo	Kayo	1	0	100.00
Kecuali	Kecuali	1	0	100.00
Keluarga	Keluarga	1	0	100.00
Ketek	Ketek	1	0	100.00
Kian	Kian	1	0	100.00
Kini	CAMIN	0	1	0.00
Kito	Kito	1	0	100.00
Ko	Ko	1	0	100.00
Kok	PA	0	1	0.00
Kuaso	Kuaso	1	0	100.00
Kucindan	Kucindan	1	0	100.00
Kue	Kue	1	0	100.00
Kulabu	Kulabu	1	0	100.00
Kulindan	Kulindan	1	0	100.00
Kumpulan	Kumpulan	1	0	100.00
Kuncang	Kuncang	1	0	100.00
Kunci	Kunci	1	0	100.00
Kurang	Kurang	1	0	100.00
Kurisi	Kurisi	1	0	100.00
Kusuik	PADUSI	0	1	0.00
Labiah	LATIAH	0	1	0.00
Ladang	Ladang	1	0	100.00
Lah	Lah	1	0	100.00
Lai	LANGIK	0	1	0.00
Lain	Lain	1	0	100.00
Lainnyo	Lainnyo	1	0	100.00
Laki	Laki	1	0	100.00
Laku	LAH KO	0	2	-100.00
Lakuan	LAH TUAN	0	2	-100.00
Lalu	Lalu	1	0	100.00
Lamah	Lamah	1	0	100.00
Lamak	Lamak	1	0	100.00
Lambak	Lambak	1	0	100.00
Lambang	Lambang	1	0	100.00
Lambuik	Lambuik	1	0	100.00
Lamo	Lamo	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Langik	Langik	1	0	100.00
Langkok	Langkok	1	0	100.00
Langsuang	Langsuang	1	0	100.00
Lantai	Lantai	1	0	100.00
Lapa	LAPANG	0	1	0.00
Lapang	Lapang	1	0	100.00
Latak	Latak	1	0	100.00
Latiah	Latiah	1	0	100.00
Lauik	Lauik	1	0	100.00
Layua	LAI MUAH	0	2	-100.00
Lia	RILAH	0	1	0.00
Lidah	Lidah	1	0	100.00
Limo	Limo	1	0	100.00
Lingkuangan	Lingkuangan	1	0	100.00
Litak	Litak	1	0	100.00
Lua	RUBAH	0	1	0.00
Lupo	Lupo	1	0	100.00
Ma	MAH	0	1	0.00
Maaf	Maaf	1	0	100.00
Mah	Mah	1	0	100.00
Maha	MAKAN	0	1	0.00
Mak	mak	1	0	100.00
Makan	MA KAN	0	2	-100.00
Makanannyo	Makanannyo	1	0	100.00
Mako	Mako	1	0	100.00
Malacuik	MALAH CARI MA	0	3	-200.00
Malah	Malah	1	0	100.00
Malam	Malam	1	0	100.00
Maliang	Maliang	1	0	100.00
Malu	KALIMO	0	1	0.00
Mamahami	Mamahami	1	0	100.00
Mamak	Mamak	1	0	100.00
Mamakai	Mamakai	1	0	100.00
Mambali	Mambali	1	0	100.00
Mambao	Mambao	1	0	100.00
Mambari	Mambari	1	0	100.00
Mambina	MAMINTAK	0	1	0.00
Mambuek	Mambuek	1	0	100.00
Mambukak	MANGKO KA	0	2	-100.00
Mamintak	Mamintak	1	0	100.00

Mampalakukan	Mampalakukan	1	0	100.00
Mamutuihan	MALU TU IJAN	0	3	-200.00
Manaka	MAMAK BAA	0	2	-100.00
Manakua	MAMAK KURANG	0	2	-100.00
Manangih	Manangih	1	0	100.00
Manarangan	Manarangan	1	0	100.00
Manarimo	Manarimo	1	0	100.00
Mancaliak	Mancaliak	1	0	100.00
Mancari	Mancari	1	0	100.00
Mancarikan	MANCARI KAN	0	2	-100.00
Mancukua	Mancukua	1	0	100.00
Mandakek	MAMAK DEK	0	2	-100.00
Mandanga	Mandanga	1	0	100.00
Mandeh	Mandeh	1	0	100.00
Mandiri	Mandiri	1	0	100.00
Manfaaik	Manfaaik	1	0	100.00
Manga	Manga	1	0	100.00
Mangabulkan	MANGADUAN	0	1	0.00
Mangaduan	Mangaduan MA	0	1	0.00
Mangaluakan	Mangaluakan MA	0	1	0.00
Mangana	Mangana	1	0	100.00
Mangarajoan	KARAJOAN	0	1	0.00
Mangecek	Mangecek	1	0	100.00
Mangenai	Mangenai	1	0	100.00
Manggaleh	Manggaleh	1	0	100.00
Mangicek	Mangicek	1	0	100.00
Mangko	Mangko	1	0	100.00
Manih	Manih	1	0	100.00
Manilai	Manilai	1	0	100.00
Manimbang	Manimbang	1	0	100.00
Manjadi	Manjadi	1	0	100.00
Manjago	Manjago	1	0	100.00
Manjalani	Manjalani	1	0	100.00
Manjawabkan	Manjawabkan	1	0	100.00
Manjunjuang	Manjunjuang	1	0	100.00
Mano	Mano	1	0	100.00
Manolong	Manolong MALU	0	1	0.00
Manukai	Manukai	1	0	100.00
Manulih	Manulih MALU	0	1	0.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Manumbuhkan	MANANGIH KAN	0	2	-100.00
Manuntuik	Manuntuik	1	0	100.00
Manuruik	Manuruik	1	0	100.00
Manusia	Manusia MA	0	1	0.00
Manyaimbangan	MA MANIMBANG KAN MA	0	4	-300.00
Manyalahkan	Manyalahkan	1	0	100.00
Manyambah	Manyambah	1	0	100.00
Manyasa	Manyasa	1	0	100.00
Manyekutukan	Manyekutukan	1	0	100.00
Manyeso	Manyeso	1	0	100.00
Marekam	Marekam	1	0	100.00
Martabat	Martabat	1	0	100.00
Masajik	Masajik	1	0	100.00
Masih	Masih	1	0	100.00
Maso	Maso	1	0	100.00
Masuak	Masuak	1	0	100.00
Masyarakaik	Masyarakaik	1	0	100.00
Mati	Mati	1	0	100.00
Mato	MA KO	0	2	-100.00
Maupun	Maupun	1	0	100.00
Meja	Meja	1	0	100.00
Mereka	Mereka	1	0	100.00
Miang	Miang	1	0	100.00
Minang	Minang	1	0	100.00
Minangkabau	Minangkabau	1	0	100.00
Mintak	Mintak	1	0	100.00
Mintuo	Mintuo	1	0	100.00
Minum	MINANG	0	1	0.00
Minuman	Minuman	1	0	100.00
Mo	Mo	1	0	100.00
Molah	Molah	1	0	100.00
Moral	MURAH	0	1	0.00
Muah	MAH	0	1	0.00
Mudah	Mudah	1	0	100.00
Muhrim	MO TURUN	0	2	-100.00
Mujua	Mujua	1	0	100.00
Muko	MA KO	0	2	-100.00
Muliakan	Muliakan	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Muluik	Muluik	1	0	100.00
Murah	Murah	1	0	100.00
Murtad	MUAH TAK	0	2	-100.00
Musim	Musim	1	0	100.00
Muslimah	Muslimah	1	0	100.00
Musyrik	Musyrik	1	0	100.00
Nagari	Nagari	1	0	100.00
Naiak	MANIH	0	1	0.00
Nak	Nak	1	0	100.00
Nakhoda	NAK SUDAH	0	2	-100.00
Nakua	LAH QURAN	0	2	-100.00
Namo	Namo	1	0	100.00
Nampak	Nampak	1	0	100.00
Namun	Namun	1	0	100.00
Nan	Nan	1	0	100.00
Nangih	Nangih	1	0	100.00
Nanko	Nanko	1	0	100.00
Nasehat	MASIH	0	1	0.00
Nasi	Nasi	1	0	100.00
Nasihaik	Nasihaik	1	0	100.00
Ndak	Ndak	1	0	100.00
Niaik	MOLAH	0	1	0.00
Nilai	Nilai	1	0	100.00
Niniak	Niniak	1	0	100.00
Nio	MIANG	0	1	0.00
No	NYO	0	1	0.00
Nyatonyo	Nyatonyo	1	0	100.00
Nyo	Nyo	1	0	100.00
Ombang	Ombang	1	0	100.00
Onda	PUN BAA	0	2	-100.00
Oto	PUN KO	0	2	-100.00
Pa	Pa	1	0	100.00
Pabukoan	Pabukoan	1	0	100.00
Pacik	Pacik	1	0	100.00
Pado	Pado	1	0	100.00
Padoman	Padoman	1	0	100.00
Padusi	Padusi	1	0	100.00
Paga	Paga	1	0	100.00
Pagangan	Pagangan	1	0	100.00
Pagunoan	Pagunoan	1	0	100.00
Paham	PAKAN	0	1	0.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pai	BARI	0	1	0.00
Paja	Paja	1	0	100.00
Pajalanan	Pajalanan	1	0	100.00
Pakalua	PA TARUANG	0	2	-100.00
Pakan	KAN	0	1	0.00
Pakaro	Pakaro	1	0	100.00
Palajari	BABAGAI	0	1	0.00
Paliang	Paliang	1	0	100.00
Paliharo	Paliharo	1	0	100.00
Pamahaman	Pamahaman	1	0	100.00
Pamanih	Pamanih	1	0	100.00
Pandai	Pandai	1	0	100.00
Pandang	Pandang	1	0	100.00
Pandapek	Pandapek PA	0	1	0.00
Pandirian	Pandirian	1	0	100.00
Pangajaran	Pangajaran	1	0	100.00
Pangampun	Pangampun	1	0	100.00
Pangaratian	Pangaratian	1	0	100.00
Pangaruah	Pangaruah	1	0	100.00
Pangkaek	Pangkaek	1	0	100.00
Panilaian	Panilaian	1	0	100.00
Panjuru	Panjuru	1	0	100.00
Pantiang	Pantiang	1	0	100.00
Panutan	Panutan	1	0	100.00
Panyaba	Panyaba	1	0	100.00
Panyakik	PAGA DI	0	2	-100.00
Panyamangaik	Panyamangaik	1	0	100.00
Panyamun	Panyamun	1	0	100.00
Papatah	Papatah	1	0	100.00
Papisahan	Papisahan	1	0	100.00
Paragokan	Paragokan	1	0	100.00
Parahu	BALAKU	0	1	0.00
Parak	Parak	1	0	100.00
Paralu	PARTAMO	0	1	0.00
Parangai	Parangai	1	0	100.00
Parantau	Parantau	1	0	100.00
Parintah	Parintah	1	0	100.00
Pariuak	Pariuak	1	0	100.00
Partamo	Partamo	1	0	100.00
Paruik	Paruik	1	0	100.00
Pasak	Pasak	1	0	100.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pasti	Pasti	1	0	100.00
Patah	Patah	1	0	100.00
Patamuan	Patamuan	1	0	100.00
Patimbangan	Patimbangan	1	0	100.00
Patinggi	Patinggi	1	0	100.00
Patuik	PA TU	0	2	-100.00
Perak	Perak	1	0	100.00
Pernah	Pernah	1	0	100.00
Peti	Peti	1	0	100.00
Pijak	Pijak	1	0	100.00
Pijakan	Pijakan	1	0	100.00
Pikia	DI JILAH	0	2	-100.00
Pikiran	Pikiran	1	0	100.00
Pindah	Pindah	1	0	100.00
Pintu	Pintu	1	0	100.00
Pitih	Pitih	1	0	100.00
Pituah	ITULAH	0	1	0.00
Pituluik	DI TUNGKEK	0	2	-100.00
Posisi	Posisi	1	0	100.00
Pukua	KURANG	0.00	0.00	0.00
Pulang	Pulang	1	0	100.00
Pulo	PUNYO	0	1	0.00
Pun	Pun	1	0	100.00
Punah	Punah	1	0	100.00
Pungguang	Pungguang	1	0	100.00
Punyo	Punyo	1	0	100.00
Quran	Quran	1	0	100.00
Rabaa	BAA	0	1	0.00
Ragam	Ragam	1	0	100.00
Rajin	Rajin	1	0	100.00
Rajo	Rajo	1	0	100.00
Rambuik	TAMUI	0	1	0.00
Rami	Rami	1	0	100.00
Rampiang	SANTIANG	0	1	0.00
Rancak	Rancak	1	0	100.00
Randang	Randang	1	0	100.00
Rang	Rang	1	0	100.00
Rangsanak	Rangsanak	1	0	100.00
Rantangkan	Rantangkan	1	0	100.00
Rantau	Rantau	1	0	100.00
Rapek	RANG DEK	0	2	-100.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Rareh	Rareh	1	0	100.00
Raso	Raso	1	0	100.00
Rasulullah	Rasulullah	1	0	100.00
Riak	Riak	1	0	100.00
Rilah	Rilah	1	0	100.00
Rimbo	Rimbo	1	0	100.00
Roman	RUMAH	0	1	0.00
Rubah	Rubah	1	0	100.00
Rumah	Rumah	1	0	100.00
Rumik	Rumik	1	0	100.00
Rundiang	Rundiang	1	0	100.00
Ruponyo	Ruponyo	1	0	100.00
Sa	SALAH	0	1	0.00
Sabagai	SA BAGAI	0	2	-100.00
Sabagainyo	Sabagainyo	1	0	100.00
Sabalah	Sabalah	1	0	100.00
Sabaleh	Sabaleh	1	0	100.00
Sabana	Sabana	1	0	100.00
Sabuah	Sabuah	1	0	100.00
Sacaro	Sacaro	1	0	100.00
Sadang	Sadang	1	0	100.00
Sadar	SADIAH	0	2	-100.00
Sadari	Sadari	1	0	100.00
Sadiah	Sadiah	1	0	100.00
Sadioan	SANDI TUAN	0	2	-100.00
Sado	Sado	1	0	100.00
Sagalo	Sagalo	1	0	100.00
Sahari	Sahari	1	0	100.00
Sahinggo	Sahinggo	1	0	100.00
Sajak	Sajak	1	0	100.00
Sajo	Sajo	1	0	100.00
Sakali	Sakali	1	0	100.00
Saketek	Saketek	1	0	100.00
Sakik	SAHARI	0	1	0.00
Sakitu	SAKETEK TU	0	2	-100.00
Salah	Salah	1	0	100.00
Salain	Salain	1	0	100.00
Salaintun	Salaintun	1	0	100.00
Salamo	Salamo	1	0	100.00
Salanjuiknyo	Salanjuiknyo	1	0	100.00
Salapan	Salapan	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Salasa	Salasa	1	0	100.00
Salasai	Salasai	1	0	100.00
Salemo	Salemo	1	0	100.00
Samaso	Samaso	1	0	100.00
Samba	SAMBAH	0	1	0.00
Sambah	Sambah	1	0	100.00
Sambilan	Sambilan	1	0	100.00
Sambuik	SAMUIK	0	1	0.00
Samiang	SANTIANG	0	1	0.00
Samo	Samo	1	0	100.00
Samoga	Samoga	1	0	100.00
Sampai	SAMO PAI	0	2	-100.00
Sampaikan	Sampaikan	1	0	100.00
Sampiang	SANTIANG	0	1	0.00
Sampik	Sampik	1	0	100.00
Samuik	Samuik	1	0	100.00
Sanak	SALAH	0	1	0.00
Sanang	Sanang	1	0	100.00
Sanangan	Sanangan	1	0	100.00
Sandi	Sandi	1	0	100.00
Sandirino	Sandirino	1	0	100.00
Sangaik	SALAH	0	1	0.00
Sangko	Sangko	1	0	100.00
Santiang	Santiang	1	0	100.00
Santun	Santun	1	0	100.00
Sapanjang	Sapanjang	1	0	100.00
Sapanuah	Sapanuah	1	0	100.00
Sapeda	SA TIDAK	0	2	-100.00
Sapuluah	Sapuluah	1	0	100.00
Sarah	Sarah	1	0	100.00
Sarato	Sarato	1	0	100.00
Saratuih	Saratuih	1	0	100.00
Sarok	Sarok	1	0	100.00
Sarupo	Sarupo	1	0	100.00
Sasek	Sasek	1	0	100.00
Satangkai	Satangkai	1	0	100.00
Satiok	Satiok	1	0	100.00
Satu	SANTUN	0	1	0.00
Sayang	CARI ANG	0	2	-100.00
Sehat	Sehat	1	0	100.00
Senyum	Senyum	1	0	100.00

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Sesuatu	Sesuatu	1	0	100.00
Sia	SIANG	0	1	0.00
Siang	Siang	1	0	100.00
Sikap	SINGGAH	0	1	0.00
Siko	SINGKOK	0	1	0.00
Sinan	Sinan	1	0	100.00
Sinayan	Sinayan	1	0	100.00
Singgah	Singgah	1	0	100.00
Singkok	Singkok	1	0	100.00
Sipaik	Sipaik	1	0	100.00
Sirah	Sirah	1	0	100.00
Situ	KITO	0	2	-100.00
Sopan	Sopan	1	0	100.00
Suatu	TU WAKATU	0	2	-100.00
Subhanahu	Subhanahu	1	0	100.00
SUDAH	MUDAH	0	1	0.00
SUKO	TU KO	0	2	-100.00
SUKU	SUKO	0	1	0.00
Sumangaik	Sumangaik	1	0	100.00
Sungguah	Sungguah	1	0	100.00
Sunnah	Sunnah	1	0	100.00
Supayo	Supayo	1	0	100.00
Surang	Surang	1	0	100.00
Suruik	Suruik	1	0	100.00
Susah	Susah	1	0	100.00
Susun	Susun	1	0	100.00
Syariat	SADIAH	0	1	0.00
Syirik	Syirik	1	0	100.00
Ta	TANGAH	0	1	0.00
Tabahak	Tabahak	1	0	100.00
Tabek	BATEH	0	1	0.00
Tadi	Tadi	1	0	100.00
Tagaduah	Tagaduah	1	0	100.00
Tagun	TAHUN	0	1	0.00
Tahun	TAGUN	0	1	0.00
Tak	Tak	1	0	100.00
Takana	Takana	1	0	100.00
Takantuak	Takantuak	1	0	100.00
Takuik	Takuik	1	0	100.00
Talok	Talok	1	0	100.00
Tamasuak	Tamasuak	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tampak	Tampak	1	0	100.00
Tampek	Tampek	1	0	100.00
Tamui	Tamui	1	0	100.00
Tanamo	Tanamo	1	0	100.00
Tangah	Tangah	1	0	100.00
Tangan	Tangan	1	0	100.00
Tanggo	Tanggo	1	0	100.00
Tantang	Tantang	1	0	100.00
Tantu	Tantu	1	0	100.00
Tanyo	Tanyo	1	0	100.00
Tapek	Tapek	1	0	100.00
Tapi	TAK DI	0	2	-100.00
Taragak	Taragak	1	0	100.00
Tarang	Tarang	1	0	100.00
Taraso	Taraso	1	0	100.00
Taratik	Taratik	1	0	100.00
Tarompa	Tarompa	1	0	100.00
Taruang	Taruang	1	0	100.00
Taruih	Taruih	1	0	100.00
Tas	SALAH	0	1	0.00
Tasebar	Tasebar	1	0	100.00
Tatap	AKIDAH	0	1	0.00
Tau	DAUN	0	1	0.00
Tentang	TANTANG	0	1	0.00
Tibo	TIGO	0	1	0.00
Tidak	Tidak	1	0	100.00
Tidua	Tidua	1	0	100.00
Tigo	Tigo	1	0	100.00
Timbua	Timbua	1	0	100.00
Tingga	Tingga	1	0	100.00
Tinggi	Tinggi	1	0	100.00
Tingkah	Tingkah	1	0	100.00
Tiok	PIJAK	0	1	0.00
Tolak	Tolak	1	0	100.00
Tolong	Tolong	1	0	100.00
Tradisi	Tradisi	1	0	100.00
Tu	TOLONG	0	1	0.00
Tuan	Tuan	1	0	100.00
Tuduang	Tuduang	1	0	100.00
Tuhan	Tuhan	1	0	100.00
Tuju	Tuju	1	0	100.00



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tujuh	Tujuh	1	0	100.00
Tujuan	Tujuan	1	0	100.00
Tuka	Tuka	1	0	100.00
Tumbuah	TU MUAH	0	2	-100.00
Tun	Tun	1	0	100.00
Tungkek	Tungkek	1	0	100.00
Turun	Turun	1	0	100.00
Tutua	Tutua	1	0	100.00
Uang	Uang	1	0	100.00
Ubah	Ubah	1	0	100.00
Ubek	KUE	0	1	0.00
Uda	MUDAH	0	1	0.00
Ujian	Ujian	1	0	100.00
Ukua	KAWAN	0	1	0.00
Ulama	Ulama	1	0	100.00
Ulamo	PULANG	0	1	0.00
Umpamo	Umpamo	1	0	100.00
Umua	RUMAH	0	1	0.00
Umuanyo	Umuanyo	1	0	100.00
Uni	Uni	1	0	100.00
Untuak	Untuak	1	0	100.00
Upiak	Upiak	1	0	100.00
Urang	KURANG	0	1	0.00
Urusan	Urusan	1	0	100.00
Usah	Usah	1	0	100.00
Usahlah	Usahlah	1	0	100.00
Usahoan	Usahoan	1	0	100.00
Utamo	TU KAMEK	0	2	-100.00
Wakatu	Wakatu	1	0	100.00
Walau	Walau	1	0	100.00
Warna	Warna	1	0	100.00
Wirusaho	Wirusaho	1	0	100.00
Zaman	Zaman	1	0	100.00

LAMPIRAN D

Perhitungan Output Model Eksperimen 1a

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\agam\output\agam.cl_cont" -dict "C:\CMUSPHIN...
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 153 nodes, 393 links
INFO: ps_lattice.c(1376): Bestpath score: -11335
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:297:334) = -794552
INFO: ps_lattice.c(1437): Joint P(0,S) = -867700 P(5|0) = -73228
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAKA KABA MAH LAI SIKO JO
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 34.05 -6.76 -5.04 -13.39 -10.47 -16.90 -3.72 -2.67 -11.04 -5.07 -4.31 -7.75 -0.12 >
INFO: cmn_live.c(105): Update to < 32.41 -6.83 -5.70 -12.65 -8.64 -15.41 -3.19 -2.88 -10.03 -4.54 -3.83 -7.10 -0.47 >
INFO: cmn_live.c(120): Update from < 32.41 -6.83 -5.70 -12.65 -8.64 -15.41 -3.19 -2.88 -10.03 -4.54 -3.83 -7.10 -0.47 >
INFO: cmn_live.c(138): Update to < 28.63 -6.89 -5.23 -11.05 -6.31 -12.29 -2.45 -2.64 -8.24 -3.72 -2.84 -5.78 -0.57 >
INFO: ngram_search_fwdtree.c(1550): 568 words recognized (3/fr)
INFO: ngram_search_fwdtree.c(1552): 13680 senones evaluated (72/fr)
INFO: ngram_search_fwdtree.c(1556): 43900 channels searched (232/fr), 18917 1st, 726 last
INFO: ngram_search_fwdtree.c(1559): 726 words for which last channels evaluated (3/fr)
INFO: ngram_search_fwdtree.c(1561): 733 candidate words for entering last phone (3/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 2.69 wall 1.422 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 9 words
INFO: ngram_search_fwdflat.c(948): 627 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 2901 senones evaluated (15/fr)
INFO: ngram_search_fwdflat.c(952): 2023 channels searched (10/fr)
INFO: ngram_search_fwdflat.c(954): 1217 words searched (6/fr)
INFO: ngram_search_fwdflat.c(957): 343 word transitions (1/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.>163
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
  
```

- Kalimat awal : ba'a kaba amak, lai sehat sajo (6 kata)
- Output model : baka kaba mah lai siko jo (6 kata)
- Analisis:

Substitusi (S):) : “ba’a” digantikan oleh “patuik” (S = 1)

“amak” digantikan oleh “mah” (S = 1)

“sehat” digantikan oleh “siko” (S = 1)

“sajo” digantikan oleh “jo” (S = 1)

Total substitusi = 4

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 6 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{4 + 0 + 0}{6} \times 100\% = 66.7\%$$

Jadi, score WER untuk contoh ini sebesar 66.7%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Perhitungan Output Model Eksperimen 1b

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\agam\output\agam.ci_cont" -dict "C:\CMUSPHINX...
TARASO PATUIK BADAN BARUPO KINI
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 34.87 -6.68 1.53 -4.69 -14.42 -9.92 -1.94 -4.80 -13.41 -4.76 -2.72 -5.34 3.31 >
INFO: cmn_live.c(105): Update to < 36.49 -6.56 -0.16 -7.43 -15.72 -11.34 -3.57 -4.57 -13.58 -5.15 -2.64 -6.34 3.37 >
INFO: cmn_live.c(120): Update from < 36.49 -6.56 -0.16 -7.43 -15.72 -11.34 -3.57 -4.57 -13.58 -5.15 -2.64 -6.34 3.37 >
INFO: cmn_live.c(138): Update to < 34.16 -7.39 1.20 -4.49 -15.28 -10.35 -2.17 -5.58 -14.49 -5.55 -3.35 -5.56 2.74 >
INFO: ngram_search_fwdtree.c(1550): 1707 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 16824 senones evaluated (44/fr)
INFO: ngram_search_fwdtree.c(1556): 43044 channels searched (113/fr), 16708 1st, 2743 last
INFO: ngram_search_fwdtree.c(1559): 2743 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1152 candidate words for entering last phone (3/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.004 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 7.33 wall 1.929 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 63 words
INFO: ngram_search_fwdflat.c(948): 1645 words recognized (4/fr)
INFO: ngram_search_fwdflat.c(950): 10540 senones evaluated (28/fr)
INFO: ngram_search_fwdflat.c(952): 15923 channels searched (41/fr)
INFO: ngram_search_fwdflat.c(954): 7472 words searched (19/fr)
INFO: ngram_search_fwdflat.c(957): 4134 word transitions (10/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.342
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 134 nodes, 140 links
INFO: ps_lattice.c(1376): Bestpath score: -10315
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:342:378) = -724521
INFO: ps_lattice.c(1437): Joint P(0,S) = -770596 P(S|0) = -46075
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
TARASO PATUIK BADAN BARUPO KINI
    
```

- Kalimat awal : Taraso auih badan ambo kini (5 kata)
- Output model : Taraso patuik badan barupo kini (5 kata)
- Analisis:

Substitusi (S): : “auih” digantikan oleh “patuik” (S = 1)

“ambo” digantikan oleh “barupo” (S = 1)

Total substitusi = 2

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 0 + 0}{5} \times 100\% = 40\%$$

Jadi, score WER untuk contoh ini sebesar 40%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Perhitungan Output Model Eksperimen 2a

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\SPHINX2\AGAM\output\agam.ci_cont" -dict "C:\SPHINX2\AGA...
INFO: cmn_live.c(88): Update from < 36.88 -2.15 -0.07 -6.73 -6.79 -11.01 -1.11 -1.67 -6.79 -1.35 -1.33 -3.79 2.94 >
INFO: cmn_live.c(105): Update to < 38.15 -2.53 -3.61 -9.78 -8.10 -13.23 -2.17 -2.83 -7.56 -1.46 -2.27 -5.27 0.91 >
INFO: cmn_live.c(120): Update from < 38.15 -2.53 -3.61 -9.78 -8.10 -13.23 -2.17 -2.83 -7.56 -1.46 -2.27 -5.27 0.91 >
INFO: cmn_live.c(138): Update to < 37.75 -4.48 -2.77 -9.41 -8.26 -13.70 -2.75 -2.31 -8.60 -3.05 -3.10 -5.99 1.79 >
INFO: ngram_search_fwdtree.c(1550): 1496 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 16107 senones evaluated (45/fr)
INFO: ngram_search_fwdtree.c(1556): 44286 channels searched (123/fr), 16857 1st, 2814 last
INFO: ngram_search_fwdtree.c(1559): 2814 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1766 candidate words for entering last phone (4/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 7.46 wall 2.083 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 80 words
INFO: ngram_search_fwdflat.c(948): 1585 words recognized (4/fr)
INFO: ngram_search_fwdflat.c(950): 10743 senones evaluated (30/fr)
INFO: ngram_search_fwdflat.c(952): 21547 channels searched (60/fr)
INFO: ngram_search_fwdflat.c(954): 8942 words searched (24/fr)
INFO: ngram_search_fwdflat.c(957): 4558 word transitions (12/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.326
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 142 nodes, 150 links
INFO: ps_lattice.c(1376): Bestpath score: -12712
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:326:356) = -874793
INFO: ps_lattice.c(1437): Joint P(O,S) = -935006 P(S|O) = -60213
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAKA KABA MA LAI SIKO JO
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : ba’a kaba amak, lai sehat sajo (6 kata)
- Output model : baka kaba ma lai siko jo (6 kata)
- Analisis:

Substitusi (S): : “ba’a” digantikan oleh “patuik” (S = 1)
 “amak” digantikan oleh “ma” (S = 1)
 “sehat” digantikan oleh “siko” (S = 1)
 “sajo” digantikan oleh “jo” (S = 1)
 Total substitusi = 4

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 6 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{4 + 0 + 0}{6} \times 100\% = 66.7\%$$

Jadi, score WER untuk contoh ini sebesar 66.7%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Perhitungan Output Model Eksperimen 2b

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\SPHINX2\AGAM\output\agam.ci_cont" -dict "C:\SPHINX2\AGA...
INFO: cmn_live.c(88): Update from < 31.00 -4.15 2.40 -3.36 -8.94 -7.62 -0.73 -4.95 -9.18 -2.57 -2.39 -4.52 3.24 >
INFO: cmn_live.c(105): Update to < 32.80 -5.69 2.49 -4.74 -11.91 -9.29 -1.65 -5.84 -12.25 -4.71 -2.93 -5.91 3.37 >
INFO: cmn_live.c(120): Update from < 32.80 -5.69 2.49 -4.74 -11.91 -9.29 -1.65 -5.84 -12.25 -4.71 -2.93 -5.91 3.37 >
INFO: cmn_live.c(138): Update to < 31.27 -6.04 2.57 -4.22 -11.15 -8.90 -1.72 -5.44 -11.39 -4.41 -2.71 -5.30 3.25 >
INFO: ngram_search_fwdtree.c(1550): 1655 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 17829 senones evaluated (47/fr)
INFO: ngram_search_fwdtree.c(1556): 45413 channels searched (119/fr), 18289 1st, 2681 last
INFO: ngram_search_fwdtree.c(1559): 2681 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1192 candidate words for entering last phone (3/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.05 CPU 0.012 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 8.88 wall 2.344 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 58 words
INFO: ngram_search_fwdflat.c(948): 1616 words recognized (4/fr)
INFO: ngram_search_fwdflat.c(950): 10380 senones evaluated (27/fr)
INFO: ngram_search_fwdflat.c(952): 14831 channels searched (39/fr)
INFO: ngram_search_fwdflat.c(954): 7062 words searched (18/fr)
INFO: ngram_search_fwdflat.c(957): 3794 word transitions (10/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.342
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 118 nodes, 113 links
INFO: ps_lattice.c(1376): Bestpath score: -10785
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>.342:377) = -753968
INFO: ps_lattice.c(1437): Joint P(0,S) = -792816 P(S|0) = -38848
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
TARASO PATUIK BADAN PADO KINI
INFO: continuous.c(275): Ready....

```

- Kalimat awal : Taraso auih badan ambo kini (5 kata)
- Output model : Taraso patuik badan pado kini (5 kata)
- Analisis:

Substitusi (S):): “auih” digantikan oleh “patuik” (S = 1)

“ambo” digantikan oleh “pado” (S = 1)

Total substitusi = 2

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 0 + 0}{5} \times 100\% = 40\%$$

Jadi, score WER untuk contoh ini sebesar 40%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Perhitungan Output Model Eksperimen 3a

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\SPHINX2\AGAM1\output\AGAM1.ci_cont" -dict "C:\SPHINX2\A...
INFO: ngram_search_fwdflat.c(948): 1554 words recognized (5/fr)
INFO: ngram_search_fwdflat.c(950): 10713 senones evaluated (31/fr)
INFO: ngram_search_fwdflat.c(952): 24028 channels searched (70/fr)
INFO: ngram_search_fwdflat.c(954): 9136 words searched (26/fr)
INFO: ngram_search_fwdflat.c(957): 4572 word transitions (13/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.306
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 143 nodes, 345 links
INFO: ps_lattice.c(1376): Bestpath score: -10265
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:306:340) = -736436
INFO: ps_lattice.c(1437): Joint P(0,S) = -767730 P(S|0) = -31294
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAGARAH KAMA LAI SEHAT SAJO
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 36.63 -6.77 -5.67 -13.54 -10.33 -15.73 -3.88 -1.84 -12.65 -3.40 -4.91 -7.31 -1.62 >
INFO: cmn_live.c(105): Update to < 37.52 -5.63 -7.16 -15.37 -11.26 -16.17 -4.29 -2.02 -12.49 -3.00 -4.30 -7.50 -2.77 >
INFO: cmn_live.c(120): Update from < 37.52 -5.63 -7.16 -15.37 -11.26 -16.17 -4.29 -2.02 -12.49 -3.00 -4.30 -7.50 -2.77 >
INFO: cmn_live.c(138): Update to < 37.29 -6.79 -5.78 -13.99 -10.93 -16.69 -3.58 -1.70 -13.46 -3.39 -5.20 -7.86 -1.39 >
INFO: ngram_search_fwdtree.c(1550): 1479 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 16842 senones evaluated (49/fr)
INFO: ngram_search_fwdtree.c(1556): 53539 channels searched (156/fr), 18481 1st, 2992 last
INFO: ngram_search_fwdtree.c(1559): 2992 words for which last channels evaluated (8/fr)
INFO: ngram_search_fwdtree.c(1561): 2270 candidate words for entering last phone (6/fr)
  
```

- Kalimat awal : ba'a kaba amak, lai sehat sajo (6 kata)
- Output model : bagarah kama lai sehat sajo (6 kata)
- Analisis:

Substitusi (S): : “ba’a” digantikan oleh “bagarah” (S = 1)
 “kaba” digantikan oleh “kama” (S = 1)
 Total substitusi = 2

Deletion (D): “amak” dihapus dalam transkripsi model (D = 1).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 6 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 1 + 0}{6} \times 100\% = 50\%$$

Jadi, score WER untuk contoh ini sebesar 50%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Perhitungan Output Model Eksperimen 3b

```

C:\> Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\SPHINX2\AGAM1\output\AGAM1.ci_cont" -dict "C:\SPHINX2\A...
INFO: ngram_search_fwdflat.c(957): 3566 word transitions (9/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.343
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 114 nodes, 135 links
INFO: ps_lattice.c(1376): Bestpath score: -10075
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:343:376) = -706947
INFO: ps_lattice.c(1437): Joint P(0,S) = -754620 P(S|0) = -47673
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
TARASO PATUIK BADAN AMBO KINI
INFO: continuous.c(275): Ready....
INFO: cmn_live.c(88): Update from < 32.91 -7.44 -0.09 -5.80 -13.56 -11.40 -2.45 -5.41 -15.18 -5.62 -4.02 -6.02 1.03 >
INFO: cmn_live.c(105): Update to < 32.49 -7.38 -0.28 -5.51 -12.94 -11.16 -2.11 -5.26 -14.70 -5.45 -3.82 -5.84 1.04 >
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(120): Update from < 32.49 -7.38 -0.28 -5.51 -12.94 -11.16 -2.11 -5.26 -14.70 -5.45 -3.82 -5.84 1.04 >
INFO: cmn_live.c(138): Update to < 29.31 -7.72 -0.33 -5.37 -9.37 -9.66 -0.80 -3.94 -10.61 -4.41 -3.47 -4.99 0.52 >
INFO: ngram_search_fwdtree.c(1550): 1025 words recognized (5/fr)
INFO: ngram_search_fwdtree.c(1552): 12681 senones evaluated (66/fr)
INFO: ngram_search_fwdtree.c(1556): 55911 channels searched (289/fr), 17056 1st, 1543 last
INFO: ngram_search_fwdtree.c(1559): 1543 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 2587 candidate words for entering last phone (13/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 4.41 wall 2.283 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 21 words
INFO: ngram_search_fwdflat.c(948): 939 words recognized (5/fr)
INFO: ngram_search_fwdflat.c(950): 3450 senones evaluated (18/fr)
INFO: ngram_search_fwdflat.c(952): 5028 channels searched (26/fr)
INFO: ngram_search_fwdflat.c(954): 2464 words searched (12/fr)
  
```

- Kalimat awal : Taraso auh badan ambo kini (5 kata)
- Output model : Taraso patuik badan ambo kini (5 kata)
- Analisis:
 - Substitusi (S): : “auh” digantikan oleh “patuik” (S = 1)
 - Deletion (D): Tidak ada kata yang dihapus (D = 0).
 - Insertion (I): Tidak ada kata yang ditambahkan (I = 0).
 - Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal
- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{1 + 0 + 0}{5} \times 100\% = 20\%$$

Jadi, score WER untuk contoh ini sebesar 20%

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Perhitungan Hasil Eksperimen 4a

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang.ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1381): Lattice has 133 nodes, 249 links
INFO: ps_lattice.c(1376): Bestpath score: -10679
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:310:340) = -783212
INFO: ps_lattice.c(1437): Joint P(0,S) = -831275 P(S|0) = -48063
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
BAKA KABA MAH GULAI SEHAT SAJO
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 40.42 -6.01 -6.10 -9.64 -12.71 -12.93 -5.28 -4.65 -9.67 -4.68 -4.79 -7.09 -0.75 >
INFO: cmn_live.c(105): Update to < 42.44 -5.80 -7.75 -11.39 -13.39 -14.76 -5.95 -4.28 -10.29 -4.09 -4.85 -7.56 -1.96 >

INFO: cmn_live.c(120): Update from < 42.44 -5.80 -7.75 -11.39 -13.39 -14.76 -5.95 -4.28 -10.29 -4.09 -4.85 -7.56 -1.96 >
INFO: cmn_live.c(138): Update to < 41.84 -6.57 -6.62 -10.54 -13.49 -14.16 -5.83 -3.73 -10.45 -4.83 -5.22 -8.21 -0.84 >

INFO: ngram_search_fwdtree.c(1550): 1384 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 18489 senones evaluated (54/fr)
INFO: ngram_search_fwdtree.c(1556): 50271 channels searched (146/fr), 20841 1st, 2631 last
INFO: ngram_search_fwdtree.c(1559): 2631 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1751 candidate words for entering last phone (5/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.005 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 8.81 wall 2.575 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 66 words
INFO: ngram_search_fwdflat.c(948): 1287 words recognized (4/fr)
INFO: ngram_search_fwdflat.c(950): 10974 senones evaluated (32/fr)
INFO: ngram_search_fwdflat.c(952): 16528 channels searched (48/fr)
INFO: ngram_search_fwdflat.c(954): 7291 words searched (21/fr)
INFO: ngram_search_fwdflat.c(957): 3754 word transitions (10/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.310
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 127 nodes, 221 links
INFO: ps_lattice.c(1376): Bestpath score: -10995
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>:310:340) = -810640
INFO: ps_lattice.c(1437): Joint P(0,S) = -851242 P(S|0) = -40602
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAKA KABA MAH LAI SEHAT SAJO
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : ba'a aba amak, lai sehat sajo (6 kata)
- Output model : baka kaba mah lai sehat sajo (6 kata)
- Analisis:

Substitusi (S): “ba’a” digantikan oleh “baka” (S = 1)
 “amak” digantikan oleh “mah” (S = 1)
 Total substitusi = 2

Deletion (D): Tidak ada kata yang dihapus (D = 0)
 Insertion (I): Tidak ada kata yang ditambahkan (I = 0)

Jumlah kata referensi (N): Ada 6 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 0 + 0}{6} \times 100\% = 33.3\%$$

Jadi, score WER untuk contoh ini sebesar 33.3%.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Perhitungan Output Model Eksperimen 4b

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang.ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 129 nodes, 149 links
INFO: ps_lattice.c(1376): Bestpath score: -7614
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>;240:273) = -531180
INFO: ps_lattice.c(1437): Joint P(O,S) = -582071 P(S|O) = -50891
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
PARUIK TAMUI TAK BAHANG
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 42.43 -4.99 -8.40 -9.07 -14.11 -14.18 -3.97 -9.50 -11.63 -4.27 -5.09 -6.02 0.04 >
INFO: cmn_live.c(105): Update to < 43.07 -5.23 -8.76 -9.42 -14.29 -13.97 -4.34 -9.03 -10.75 -3.60 -4.86 -5.68 0.09 >
INFO: cmn_live.c(88): Update from < 43.07 -5.23 -8.76 -9.42 -14.29 -13.97 -4.34 -9.03 -10.75 -3.60 -4.86 -5.68 0.09 >
INFO: cmn_live.c(105): Update to < 43.42 -6.07 -5.10 -6.74 -16.69 -12.23 -3.82 -8.37 -14.52 -5.22 -5.20 -6.49 0.96 >
INFO: cmn_live.c(120): Update from < 43.42 -6.07 -5.10 -6.74 -16.69 -12.23 -3.82 -8.37 -14.52 -5.22 -5.20 -6.49 0.96 >
INFO: cmn_live.c(138): Update to < 42.03 -6.47 -4.58 -5.80 -16.14 -12.00 -3.91 -7.76 -13.82 -5.13 -4.86 -5.96 1.09 >
INFO: ngram_search_fwdtree.c(1550): 1368 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 17256 senones evaluated (45/fr)
INFO: ngram_search_fwdtree.c(1556): 38694 channels searched (101/fr), 17573 1st, 2220 last
INFO: ngram_search_fwdtree.c(1559): 2220 words for which last channels evaluated (5/fr)
INFO: ngram_search_fwdtree.c(1561): 988 candidate words for entering last phone (2/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.05 CPU 0.012 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 7.41 wall 1.946 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 55 words
INFO: ngram_search_fwdflat.c(948): 1313 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 10584 senones evaluated (28/fr)
INFO: ngram_search_fwdflat.c(952): 13065 channels searched (34/fr)
INFO: ngram_search_fwdflat.c(954): 6456 words searched (16/fr)
INFO: ngram_search_fwdflat.c(957): 3348 word transitions (8/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.343
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 107 nodes, 48 links
INFO: ps_lattice.c(1376): Bestpath score: -11674
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>;343:379) = -801413
INFO: ps_lattice.c(1437): Joint P(O,S) = -839306 P(S|O) = -37893
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
TARASO TAKUIK BADAN AMBO KINI
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Taraso auih badan ambo kini (5 kata)
- Output model : Taraso takuik badan ambo kini (4 kata)
- Analisis:
 - Substitusi (S): : “auih” digantikan oleh “takuik” (S = 1)
 - Deletion (D): Tidak ada kata yang dihapus (D = 0).
 - Insertion (I): Tidak ada kata yang ditambahkan (I = 0).
 - Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal
- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{1 + 0 + 0}{5} \times 100\% = 20\%$$

Jadi, score WER untuk contoh ini sebesar 20%

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Percobaan 1

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kambang\output\kambang.ci_conf" -dict "C:\CMUSP...
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <S>.0 end node </S>.239
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 136 nodes, 205 links
INFO: ps_lattice.c(1376): Bestpath score: -8436
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha</S>.239:264) = -612981
INFO: ps_lattice.c(1437): Joint P(O,S) = -625130 P(S|O) = -12149
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAK PARAHU TINGGA BAKAMUDI
INFO: continuous.c(275): Ready....
INFO: cmn_live.c(88): Update from < 37.01 -6.74 -3.42 -0.75 -12.20 -15.83 0.51 -7.70 -9.38 0.19 -3.60 -0.40 4.52 >
INFO: cmn_live.c(105): Update to < 37.16 -6.77 -3.46 -0.68 -12.23 -16.13 0.56 -7.85 -9.46 0.11 -3.62 -0.34 4.52 >
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(120): Update from < 37.16 -6.77 -3.46 -0.68 -12.23 -16.13 0.56 -7.85 -9.46 0.11 -3.62 -0.34 4.52 >
INFO: cmn_live.c(138): Update to < 37.72 -8.02 -2.64 -0.86 -9.91 -17.01 -1.00 -7.21 -9.44 -1.22 -4.68 -0.04 5.11 >
INFO: ngram_search_fwdtree.c(1550): 1476 words recognized (6/fr)
INFO: ngram_search_fwdtree.c(1552): 14157 senones evaluated (62/fr)
INFO: ngram_search_fwdtree.c(1556): 49375 channels searched (214/fr), 17481 1st, 2709 last
INFO: ngram_search_fwdtree.c(1559): 2709 words for which last channels evaluated (11/fr)
INFO: ngram_search_fwdtree.c(1561): 1892 candidate words for entering last phone (8/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.007 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 2.43 wall 1.055 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 69 words
INFO: ngram_search_fwdflat.c(948): 1319 words recognized (6/fr)
INFO: ngram_search_fwdflat.c(950): 9229 senones evaluated (40/fr)
INFO: ngram_search_fwdflat.c(952): 15599 channels searched (67/fr)
INFO: ngram_search_fwdflat.c(954): 6400 words searched (27/fr)
INFO: ngram_search_fwdflat.c(957): 3563 word transitions (15/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <S>.0 end node </S>.220
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 142 nodes, 423 links
INFO: ps_lattice.c(1376): Bestpath score: -10363
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha</S>.220:228) = -699562
INFO: ps_lattice.c(1437): Joint P(O,S) = -720772 P(S|O) = -21210
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.004 xRT
JASO SASEK DALAM BARAGIAH
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
    
```

- Kalimat awal : Bak parahu indak bakamudi, biaso sasek dalam balayia. (9 kata)
- Output model : Bak parahu tingga bakamudi jaso sasek dalam baragiah (9 kata)
- Analisis:

Substitusi (S): “indak” digantikan oleh “tingga” (S = 1)
 “biaso” digantikan oleh “jaso” (S = 1)
 “balayia” digantikan oleh “baragiah” (S = 1)
 Total substitusi = 3

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 9 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{3 + 0 + 0}{9} \times 100\% = 33.3\%$$

Jadi, score WER untuk contoh ini sebesar 33.3%.

Percobaan 2

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang.ci_cont" -dict "C:\CMUSP...
INFO: ngram_search_fwdflat.c(957): 6146 word transitions (23/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.02 CPU 0.006 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.02 wall 0.006 xRT
INFO: ngram_search.c(1250): lattice start node <S>.0 end node </S>.227
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 172 nodes, 646 links
INFO: ps_lattice.c(1376): Bestpath score: -9766
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</S>.227:264) = -651468
INFO: ps_lattice.c(1437): Joint P(0,S) = -710685 P(S|0) = -59217
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.003 xRT
IMAN TAGAUAH BA PANDIRIAN
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(120): Update from < 29.97 -7.91 -7.30 -2.21 -10.71 -12.40 3.21 -5.93 -5.68 1.37 1.05 2.76 7.10 >
INFO: cmn_live.c(138): Update to < 28.55 -6.60 -6.15 -2.01 -9.09 -10.68 2.49 -5.27 -4.86 1.00 -0.08 1.71 5.95 >
INFO: ngram_search_fwdtree.c(1550): 222 words recognized (3/fr)
INFO: ngram_search_fwdtree.c(1552): 3045 senones evaluated (42/fr)
INFO: ngram_search_fwdtree.c(1556): 6589 channels searched (90/fr), 3413 1st, 235 last
INFO: ngram_search_fwdtree.c(1559): 235 words for which last channels evaluated (3/fr)
INFO: ngram_search_fwdtree.c(1561): 55 candidate words for entering last phone (0/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.021 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 10.97 wall 15.027 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 3 words
INFO: ngram_search_fwdflat.c(948): 239 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 618 senones evaluated (8/fr)
INFO: ngram_search_fwdflat.c(952): 345 channels searched (4/fr)
INFO: ngram_search_fwdflat.c(954): 279 words searched (3/fr)
INFO: ngram_search_fwdflat.c(957): 107 word transitions (1/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.003 xRT
INFO: ngram_search.c(1250): lattice start node <S>.0 end node </S>.36
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 6 nodes, 10 links
INFO: ps_lattice.c(1376): Bestpath score: -1147
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</S>.36:71) = -82417
INFO: ps_lattice.c(1437): Joint P(0,S) = -112392 P(S|0) = -29975
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.008 xRT
NO
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Imannyo taguah bapandirian (5 kata)
- Output model : Iman taguah ba pandirian (4 kata)
- Analisis:

Substitusi (S): Tidak ada kata yang digantikan (S = 0)

Deletion (D): “nyo” dihapus dalam transkripsi model (D = 1).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{0 + 1 + 0}{5} \times 100\% = 20\%$$

Jadi, score WER untuk contoh ini sebesar 20%.

Percobaan 3

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamus\output\kamus\cl_cont" -dict "C:\CMUSP...
INFO: ps_lattice.c(1437): Joint P(O,S) = -1662138 P(S|O) = -44361
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 37.95 -5.15 -10.05 -5.78 -17.03 -17.78 -0.19 -7.89 -5.79 2.55 0.85 -2.33 8.85 >
INFO: cmn_live.c(105): Update to < 38.26 -5.53 -9.08 -3.67 -16.65 -18.73 1.78 -9.05 -6.97 3.27 -0.24 1.15 8.34 >
INFO: cmn_live.c(120): Update from < 38.26 -5.53 -9.08 -3.67 -16.65 -18.73 1.78 -9.05 -6.97 3.27 -0.24 1.15 8.34 >
INFO: cmn_live.c(138): Update to < 36.30 -5.67 -7.58 -3.34 -17.59 -17.22 2.69 -10.37 -8.61 1.40 -1.02 0.31 4.51 >
INFO: ngram_search_fudtree.c(1550): 1504 words recognized (4/fr)
INFO: ngram_search_fudtree.c(1552): 21801 senones evaluated (51/fr)
INFO: ngram_search_fudtree.c(1556): 53861 channels searched (127/fr), 25112 1st, 2336 last
INFO: ngram_search_fudtree.c(1559): 2336 words for which last channels evaluated (5/fr)
INFO: ngram_search_fudtree.c(1561): 1063 candidate words for entering last phone (2/fr)
INFO: ngram_search_fudtree.c(1564): fudtree 0.08 CPU 0.018 xRT
INFO: ngram_search_fudtree.c(1567): fudtree 7.53 wall 1.776 xRT
INFO: ngram_search_fudflat.c(302): Utterance vocabulary contains 56 words
INFO: ngram_search_fudflat.c(948): 1417 words recognized (3/fr)
INFO: ngram_search_fudflat.c(950): 12686 senones evaluated (30/fr)
INFO: ngram_search_fudflat.c(952): 13866 channels searched (32/fr)
INFO: ngram_search_fudflat.c(954): 6439 words searched (15/fr)
INFO: ngram_search_fudflat.c(957): 2972 word transitions (7/fr)
INFO: ngram_search_fudflat.c(960): fudflat 0.02 CPU 0.004 xRT
INFO: ngram_search_fudflat.c(963): fudflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): Lattice start node <s>.0 end node </s>.380
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 102 nodes, 157 links
INFO: ps_lattice.c(1376): Bestpath score: -12893
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>:380:422) = -873391
INFO: ps_lattice.c(1437): Joint P(O,S) = -948010 P(S|O) = -74619
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
INFO: continuous.c(275): Ready....
INFO: cmn_live.c(88): Update from < 36.30 -5.67 -7.58 -3.34 -17.59 -17.22 2.69 -10.37 -8.61 1.40 -1.02 0.31 4.51 >
INFO: cmn_live.c(105): Update to < 36.22 -5.62 -7.65 -3.58 -17.39 -17.13 2.60 -10.44 -8.62 1.35 -0.86 0.45 4.35 >
INFO: cmn_live.c(120): Update from < 36.22 -5.62 -7.65 -3.58 -17.39 -17.13 2.60 -10.44 -8.62 1.35 -0.86 0.45 4.35 >
INFO: cmn_live.c(138): Update to < 37.01 -5.20 -6.79 -3.36 -19.12 -18.20 4.38 -11.93 -9.14 1.91 -1.53 2.02 3.62 >
INFO: cmn_live.c(105): Update to < 37.01 -5.20 -6.79 -3.36 -19.12 -18.20 4.38 -11.93 -9.14 1.91 -1.53 2.02 3.62 >
INFO: cmn_live.c(120): Update from < 37.01 -5.20 -6.79 -3.36 -19.12 -18.20 4.38 -11.93 -9.14 1.91 -1.53 2.02 3.62 >
INFO: cmn_live.c(138): Update to < 35.19 -6.40 -6.08 -1.19 -17.14 -16.66 4.55 -11.06 -10.55 1.13 -2.11 2.01 2.53 >
    
```

- Kalimat awal : Batanyo ka diri, untuak apo waang ko hiduik? (9 kata)
- Output model : Batanyo ka diri tutua wakatu iduik (8 kata)
- Analisis:

Substitusi (S): “untuak” digantikan oleh “tutua” (S = 1)

“apo” digantikan oleh “wakatu” (S = 1)

Total substitusi = 2

Deletion (D): “waang” dihapus dalam output model (D = 1).

“ko” dihapus dalam output model (D = 1).

Total substitusi = 2

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 9 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 2 + 0}{9} \times 100\% = 44.4\%$$

Jadi, score WER untuk contoh ini sebesar 44.4%.

Percobaan 4

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic 'yes' -hmm "C:\CMUSP\HMM\kambang\output\kambang_ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1381): Lattice has 86 nodes, 55 links
INFO: ps_lattice.c(1376): Bestpath score: -9336
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha</s>:263:287) = -672900
INFO: ps_lattice.c(1437): Joint P(0,S) = -682166 P(S|0) = -9266
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
KAMPARI DALAM BASO MINANG
INFO: continuous.c(275): Ready...
INFO: continuous.c(261): Listening...
INFO: cm_live.c(120): Update from < 46.40 -7.91 -2.85 -8.51 -8.81 -20.27 0.94 -13.70 -18.06 -5.54 -9.50 -4.06 0.14 >
INFO: cm_live.c(138): Update to < 48.05 -8.08 -4.19 -10.47 -9.92 -17.31 1.63 -13.49 -18.40 -6.82 -9.75 -3.86 1.19 >

INFO: ngram_search_fwdtree.c(1550): 964 words recognized (3/fr)
INFO: ngram_search_fwdtree.c(1552): 15099 senones evaluated (52/fr)
INFO: ngram_search_fwdtree.c(1556): 30905 channels searched (107/fr), 14352 1st, 1830 last
INFO: ngram_search_fwdtree.c(1559): 1830 words for which last channels evaluated (6/fr)
INFO: ngram_search_fwdtree.c(1561): 944 candidate words for entering last phone (3/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.05 CPU 0.016 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 6.13 wall 2.129 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 57 words
INFO: ngram_search_fwdflat.c(948): 691 words recognized (2/fr)
INFO: ngram_search_fwdflat.c(950): 9271 senones evaluated (32/fr)
INFO: ngram_search_fwdflat.c(952): 11528 channels searched (40/fr)
INFO: ngram_search_fwdflat.c(954): 5251 words searched (18/fr)
INFO: ngram_search_fwdflat.c(957): 2742 word transitions (9/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.264
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 87 nodes, 82 links
INFO: ps_lattice.c(1376): Bestpath score: -9250
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha</s>:264:286) = -681886
INFO: ps_lattice.c(1437): Joint P(0,S) = -724669 P(S|0) = -42783
INFO: ngram_search.c(872): bestpath 0.02 CPU 0.005 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
NAMO HARI DALAM BASO MINANG
INFO: continuous.c(275): Ready...
INFO: cm_live.c(88): Update from < 48.05 -8.08 -4.19 -10.47 -9.92 -17.31 1.63 -13.49 -18.40 -6.82 -9.75 -3.86 1.19 >
INFO: cm_live.c(105): Update to < 47.57 -8.14 -4.03 -10.27 -9.68 -17.18 1.54 -13.34 -18.28 -6.72 -9.60 -3.72 1.34 >

INFO: continuous.c(261): Listening...
    
```

- Kalimat awal : namo hari dalam bahaso minang (5 kata)
- Output model : namo hari dalam bahaso minang (5 kata)
- Analisis:

Substitusi (S): Tidak ada kata yang digantikan (S = 0)

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{0 + 0 + 0}{5} \times 100\% = 0\%$$

Jadi, score WER untuk contoh ini sebesar 0%.

Percobaan 5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang_ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1381): Lattice has 126 nodes, 130 links
INFO: ps_lattice.c(1376): Bestpath score: -8970
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha</s>:255:274) = -638831
INFO: ps_lattice.c(1437): Joint P(O,S) = -655901 P(S|O) = -17070
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
KALAH TANTU MINUM BAGAI
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 45.26 -6.36 -7.18 -7.14 -13.01 -16.33 -1.46 -11.77 -14.28 -6.52 -6.60 -4.88 0.92 >
INFO: cmn_live.c(105): Update to < 45.42 -6.16 -7.56 -8.02 -13.72 -16.46 -1.57 -10.92 -13.62 -6.45 -5.61 -4.87 1.52 >
INFO: cmn_live.c(120): Update from < 45.42 -6.16 -7.56 -8.02 -13.72 -16.46 -1.57 -10.92 -13.62 -6.45 -5.61 -4.87 1.52 >
INFO: cmn_live.c(138): Update to < 42.50 -6.28 -6.69 -8.42 -13.64 -16.61 -3.19 -10.47 -11.44 -3.81 -4.32 -5.62 0.12 >
INFO: ngram_search_fwdtree.c(1550): 1585 words recognized (5/fr)
INFO: ngram_search_fwdtree.c(1552): 19032 senones evaluated (56/fr)
INFO: ngram_search_fwdtree.c(1556): 53456 channels searched (156/fr), 23135 1st, 2686 last
INFO: ngram_search_fwdtree.c(1559): 2686 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1855 candidate words for entering last phone (5/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.005 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 6.04 wall 1.766 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 71 words
INFO: ngram_search_fwdflat.c(948): 1545 words recognized (5/fr)
INFO: ngram_search_fwdflat.c(950): 10926 senones evaluated (32/fr)
INFO: ngram_search_fwdflat.c(952): 15910 channels searched (46/fr)
INFO: ngram_search_fwdflat.c(954): 6953 words searched (20/fr)
INFO: ngram_search_fwdflat.c(957): 3894 word transitions (11/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.:298
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 119 nodes, 204 links
INFO: ps_lattice.c(1376): Bestpath score: -9267
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha</s>:298:340) = -646492
INFO: ps_lattice.c(1437): Joint P(O,S) = -670234 P(S|O) = -23742
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
MOLAH MAKAN AWAK LAI
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Molah makan awak lai. (4 kata)
- Output model : Molah makan awak lai. (4 kata)
- Analisis:

Substitusi (S): Tidak ada kata yang digantikan (S = 0)

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 4 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{0 + 0 + 0}{4} \times 100\% = 0\%$$

Jadi, score WER untuk contoh ini sebesar 0%.

Percobaan 6

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSP\HMM\kambang\output\kambang_ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 129 nodes, 149 links
INFO: ps_lattice.c(1376): Bestpath score: -7614
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>:240:273) = -531180
INFO: ps_lattice.c(1437): Joint P(O,S) = -582071 P(S|O) = -50891
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
PARUIK TAPUI TAK BANAWG
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cm_live.c(88): Update from < 42.43 -4.99 -8.40 -9.07 -14.11 -14.18 -3.97 -9.50 -11.63 -4.27 -5.09 -6.02 0.04 >
INFO: cm_live.c(105): Update to < 43.07 -5.23 -8.76 -9.42 -14.29 -13.97 -4.34 -9.03 -10.75 -3.60 -4.86 -5.68 0.09 >
INFO: cm_live.c(88): Update from < 43.07 -5.23 -8.76 -9.42 -14.29 -13.97 -4.34 -9.03 -10.75 -3.60 -4.86 -5.68 0.09 >
INFO: cm_live.c(105): Update to < 43.42 -6.07 -5.10 -6.74 -16.69 -12.23 -3.82 -8.37 -14.52 -5.22 -5.20 -6.49 0.96 >
INFO: cm_live.c(120): Update from < 43.42 -6.07 -5.10 -6.74 -16.69 -12.23 -3.82 -8.37 -14.52 -5.22 -5.20 -6.49 0.96 >
INFO: cm_live.c(138): Update to < 42.03 -6.47 -4.58 -5.80 -16.14 -12.00 -3.91 -7.76 -13.82 -5.13 -4.86 -5.96 1.09 >
INFO: ngram_search_fwdtree.c(1550): 1368 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 17256 senones evaluated (45/fr)
INFO: ngram_search_fwdtree.c(1556): 38694 channels searched (101/fr), 17573 1st, 2220 last
INFO: ngram_search_fwdtree.c(1559): 2220 words for which last channels evaluated (5/fr)
INFO: ngram_search_fwdtree.c(1561): 988 candidate words for entering last phone (2/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.05 CPU 0.012 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 7.41 wall 1.946 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 55 words
INFO: ngram_search_fwdflat.c(948): 1313 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 10584 senones evaluated (28/fr)
INFO: ngram_search_fwdflat.c(952): 13065 channels searched (34/fr)
INFO: ngram_search_fwdflat.c(954): 6456 words searched (16/fr)
INFO: ngram_search_fwdflat.c(957): 3348 word transitions (8/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.343
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 107 nodes, 48 links
INFO: ps_lattice.c(1376): Bestpath score: -11674
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>:343:379) = -801413
INFO: ps_lattice.c(1437): Joint P(O,S) = -839306 P(S|O) = -37893
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
TARASO TAKUIK BADAN AMBO KINI
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Taraso auih badan ambo kini (5 kata)
- Output model : Taraso takuik badan ambo kini (4 kata)
- Analisis:
 - Substitusi (S):) : “auih” digantikan oleh “takuik” (S = 1)
 - Deletion (D): Tidak ada kata yang dihapus (D = 0).
 - Insertion (I): Tidak ada kata yang ditambahkan (I = 0).
 - Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

• Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{1 + 0 + 0}{5} \times 100\% = 20\%$$

Jadi, score WER untuk contoh ini sebesar 20%.

Percobaan 7



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mempublikasikan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSP\HMM\kamus\output\kamus\cl_cont" -dict "C:\CMUSP...
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.002 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.270
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 73 nodes, 83 links
INFO: ps_lattice.c(1376): Bestpath score: -8144
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>.270:313) = -580435
INFO: ps_lattice.c(1437): Joint P(O,S) = -609001 P(S|O) = -28566
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
SIA NIMO USAHLAH TUN
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 38.23 -7.71 -4.34 -6.87 -13.72 -10.29 -4.07 -5.90 -9.38 -5.53 -5.05 -3.47 2.24 >
INFO: cmn_live.c(105): Update to < 39.18 -7.59 -4.89 -7.55 -14.07 -9.54 -4.18 -5.41 -10.17 -5.05 -4.62 -4.00 2.25 >
INFO: cmn_live.c(120): Update from < 39.18 -7.59 -4.89 -7.55 -14.07 -9.54 -4.18 -5.41 -10.17 -5.05 -4.62 -4.00 2.25 >
INFO: cmn_live.c(138): Update to < 39.07 -7.29 -6.33 -7.62 -13.07 -9.83 -3.38 -6.35 -9.24 -5.08 -4.27 -4.23 2.14 >
INFO: ngram_search_fwdtree.c(1550): 710 words recognized (3/fr)
INFO: ngram_search_fwdtree.c(1552): 11229 senones evaluated (51/fr)
INFO: ngram_search_fwdtree.c(1556): 28454 channels searched (128/fr), 14120 1st, 1133 last
INFO: ngram_search_fwdtree.c(1559): 1133 words for which last channels evaluated (5/fr)
INFO: ngram_search_fwdtree.c(1561): 699 candidate words for entering last phone (3/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.03 CPU 0.014 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 5.37 wall 2.430 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 35 words
INFO: ngram_search_fwdflat.c(948): 476 words recognized (2/fr)
INFO: ngram_search_fwdflat.c(950): 4995 senones evaluated (23/fr)
INFO: ngram_search_fwdflat.c(952): 5911 channels searched (26/fr)
INFO: ngram_search_fwdflat.c(954): 2933 words searched (13/fr)
INFO: ngram_search_fwdflat.c(957): 1600 word transitions (7/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.181
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 38 nodes, 43 links
INFO: ps_lattice.c(1376): Bestpath score: -5767
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>.181:219) = -404826
INFO: ps_lattice.c(1437): Joint P(O,S) = -445257 P(S|O) = -40431
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.002 xRT
DIMA ANGGU TINGGA
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Dima angku tingga? (3 kata)
- Output model : Dima angku tingga. (3 kata)
- Analisis:

Substitusi (S): Tidak ada kata yang digantikan (S = 0)

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 3 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{0 + 0 + 0}{3} \times 100\% = 0\%$$

Jadi, score WER untuk contoh ini sebesar 0%.

Percobaan 8



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang_ci_cont" -dict "C:\CMUSP...
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.264
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 113 nodes, 74 links
INFO: ps_lattice.c(1376): Bestpath score: -7354
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(<s>.264:274) = -604625
INFO: ps_lattice.c(1437): Joint P(O,S) = -618225 P(S|O) = -13600
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.003 xRT
LAT RANCAK BAJU KO MAH
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 42.28 -6.97 -7.35 -9.01 -15.06 -10.66 -2.96 -8.81 -9.91 -6.21 -4.77 -4.33 0.83 >
INFO: cmn_live.c(105): Update to < 43.65 -6.84 -7.36 -9.07 -15.11 -12.21 -3.65 -8.81 -10.08 -6.68 -4.59 -4.23 -0.15 >
INFO: cmn_live.c(120): Update from < 43.65 -6.84 -7.36 -9.07 -15.11 -12.21 -3.65 -8.81 -10.08 -6.68 -4.59 -4.23 -0.15 >
INFO: cmn_live.c(138): Update to < 43.02 -6.44 -8.17 -9.02 -14.26 -11.09 -3.05 -8.75 -9.21 -6.35 -5.39 -4.23 -0.97 >
INFO: ngram_search_fwdtree.c(1550): 678 words recognized (2/fr)
INFO: ngram_search_fwdtree.c(1552): 15735 senones evaluated (57/fr)
INFO: ngram_search_fwdtree.c(1556): 32939 channels searched (120/fr), 18634 1st, 1135 last
INFO: ngram_search_fwdtree.c(1559): 1135 words for which last channels evaluated (4/fr)
INFO: ngram_search_fwdtree.c(1561): 641 candidate words for entering last phone (2/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.05 CPU 0.017 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 6.16 wall 2.248 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 28 words
INFO: ngram_search_fwdflat.c(948): 706 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 5331 senones evaluated (19/fr)
INFO: ngram_search_fwdflat.c(952): 4465 channels searched (16/fr)
INFO: ngram_search_fwdflat.c(954): 2547 words searched (9/fr)
INFO: ngram_search_fwdflat.c(957): 1346 word transitions (4/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.267
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 93 nodes, 66 links
INFO: ps_lattice.c(1376): Bestpath score: -7383
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(<s>.267:272) = -608399
INFO: ps_lattice.c(1437): Joint P(O,S) = -626160 P(S|O) = -17761
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
LAT RANCAK BAJU KO MAH
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : Lai rancak baju ko mak (5 kata)
- Output model : Lai rancak baju ko mah (5 kata)
- Analisis:
 - Substitusi (S): "mak" digantikan oleh "mah" (S = 1)
 - Deletion (D): Tidak ada kata yang dihapus (D = 0).
 - Insertion (I): Tidak ada kata yang ditambahkan (I = 0).
 - Jumlah kata referensi (N): Ada 5 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{1 + 0 + 0}{5} \times 100\% = 20\%$$

Jadi, score WER untuk contoh ini sebesar 20%

Percobaan 9



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketsphinx_continuous -inmic "yes" -lmm "C:\CMUSP\HINX\kamang\output\kamang.ci_cont" -dict "C:\CMUSP...
INFO: ngram_search.c(1381): Lattice has 133 nodes, 249 links
INFO: ps_lattice.c(1376): Bestpath score: -10679
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>:310:340) = -783212
INFO: ps_lattice.c(1437): Joint P(O,S) = -831275 P(S|O) = -48063
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.002 xRT
BAKA KABA MAH LAI SEHAT SAJO
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(88): Update from < 42.44 -5.80 -7.75 -11.39 -13.39 -14.76 -5.95 -4.28 -10.29 -4.09 -4.85 -7.56 -1.96 >
INFO: cmn_live.c(105): Update to < 42.44 -5.80 -7.75 -11.39 -13.39 -14.76 -5.95 -4.28 -10.29 -4.09 -4.85 -7.56 -1.96 >
INFO: cmn_live.c(120): Update from < 42.44 -5.80 -7.75 -11.39 -13.39 -14.76 -5.95 -4.28 -10.29 -4.09 -4.85 -7.56 -1.96 >
INFO: cmn_live.c(138): Update to < 41.84 -6.57 -6.62 -10.54 -13.49 -14.16 -5.83 -3.73 -10.45 -4.83 -5.22 -8.21 -0.84 >
INFO: ngram_search_fwdtree.c(1550): 1384 words recognized (4/fr)
INFO: ngram_search_fwdtree.c(1552): 18489 senones evaluated (54/fr)
INFO: ngram_search_fwdtree.c(1556): 50271 channels searched (146/fr), 20841 1st, 2631 last
INFO: ngram_search_fwdtree.c(1559): 2631 words for which last channels evaluated (7/fr)
INFO: ngram_search_fwdtree.c(1561): 1751 candidate words for entering last phone (5/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.02 CPU 0.005 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 8.81 wall 2.575 xRT
INFO: ngram_search_fwdflat.c(948): Utterance vocabulary contains 66 words
INFO: ngram_search_fwdflat.c(948): 1287 words recognized (4/fr)
INFO: ngram_search_fwdflat.c(950): 10974 senones evaluated (32/fr)
INFO: ngram_search_fwdflat.c(952): 16528 channels searched (48/fr)
INFO: ngram_search_fwdflat.c(954): 7291 words searched (21/fr)
INFO: ngram_search_fwdflat.c(957): 3754 word transitions (10/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node < s>:0 end node </s>:310
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 127 nodes, 221 links
INFO: ps_lattice.c(1376): Bestpath score: -10995
INFO: ps_lattice.c(1380): Normalizer P(O) = alpha(</s>:310:340) = -810640
INFO: ps_lattice.c(1437): Joint P(O,S) = -851242 P(S|O) = -40602
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
BAKA KABA MAH LAI SEHAT SAJO
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : ba'a aba amak, lai sehat sajo (6 kata)
- Output model : baka kaba mah lai sehat sajo (6 kata)
- Analisis:

Substitusi (S): “ba’a” digantikan oleh “baka” (S = 1)
 “amak” digantikan oleh “mah” (S = 1)
 Total substitusi = 2

Deletion (D): Tidak ada kata yang dihapus (D = 0)

Insertion (I): Tidak ada kata yang ditambahkan (I = 0)

Jumlah kata referensi (N): Ada 6 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{2 + 0 + 0}{6} \times 100\% = 33.3\%$$

Jadi, score WER untuk contoh ini sebesar 44.4%.

Percobaan 10



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Command Prompt - pocketphinx_continuous -inmic "yes" -hmm "C:\CMUSPHINX\kamang\output\kamang.ci_cont" -dict "C:\CMUSP...
INFO: ngram_search_fwdflat.c(957): 2185 word transitions (5/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.01 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.372
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 114 nodes, 86 links
INFO: ps_lattice.c(1376): Bestpath score: -8029
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>.372:375) = -563085
INFO: ps_lattice.c(1437): Joint P(0,S) = -603471 P(S|0) = -40386
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.00 wall 0.001 xRT
KAU KAMEK RANCAK MANIH
INFO: continuous.c(275): Ready....
INFO: continuous.c(261): Listening...
INFO: cmn_live.c(120): Update from < 43.03 -7.98 -4.96 -6.48 -16.81 -10.77 -0.46 -7.62 -16.27 -7.70 -4.10 -4.96 1.83 >
INFO: cmn_live.c(138): Update to < 39.80 -7.29 -3.89 -5.54 -15.31 -10.75 -1.92 -8.41 -15.95 -7.87 -4.19 -4.57 1.92 >
INFO: ngram_search_fwdtree.c(1558): 182 words recognized (2/fr)
INFO: ngram_search_fwdtree.c(1552): 2817 senones evaluated (37/fr)
INFO: ngram_search_fwdtree.c(1556): 4316 channels searched (56/fr), 3332 1st, 219 last
INFO: ngram_search_fwdtree.c(1559): 219 words for which last channels evaluated (2/fr)
INFO: ngram_search_fwdtree.c(1561): 0 candidate words for entering last phone (0/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 0.08 CPU 0.103 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 32.91 wall 43.304 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 2 words
INFO: ngram_search_fwdflat.c(948): 213 words recognized (3/fr)
INFO: ngram_search_fwdflat.c(950): 225 senones evaluated (3/fr)
INFO: ngram_search_fwdflat.c(952): 219 channels searched (2/fr)
INFO: ngram_search_fwdflat.c(954): 219 words searched (2/fr)
INFO: ngram_search_fwdflat.c(957): 76 word transitions (1/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.00 CPU 0.000 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.00 wall 0.001 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.34
INFO: ngram_search.c(1276): Eliminated 0 nodes before end node
INFO: ngram_search.c(1381): Lattice has 6 nodes, 5 links
INFO: ps_lattice.c(1376): Bestpath score: -776
INFO: ps_lattice.c(1380): Normalizer P(0) = alpha(</s>.34:74) = -45854
INFO: ps_lattice.c(1437): Joint P(0,S) = -55073 P(S|0) = -9219
INFO: ngram_search.c(872): bestpath 0.00 CPU 0.000 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.009 xRT
INFO: continuous.c(275): Ready....
    
```

- Kalimat awal : kau kamek, rancak, manih (4 kata)
- Output model : kau kamek rancak manih (4 kata)
- Analisis:

Substitusi (S): Tidak ada kata yang digantikan (S = 0)

Deletion (D): Tidak ada kata yang dihapus (D = 0).

Insertion (I): Tidak ada kata yang ditambahkan (I = 0).

Jumlah kata referensi (N): Ada 4 kata dalam transkripsi Kalimat awal

- Perhitungan:

$$WER = \frac{S + D + I}{N} = \frac{0 + 0 + 0}{4} \times 100\% = 0\%$$

Jadi, score WER untuk contoh ini sebesar 0%.

LAMPIRAN E

DOKUMENTASI PENELITIAN



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Nama Lengkap : Syahrul Yusman
 Tempat / Tanggal Lahir : Pekanbaru/ 7 November 2000
 Nama Ayah : Nursal
 Nama Ibu : Yusniar
 Anak Ke - : 2 (Dua)
 Jumlah Saudara : 2 (Dua)
 Alamat : Jl. Dwi Utama Raya
 Perum. Dwi Utama Raya (Blok C.9)
 E-mail : 12050112675@students.uin-suska.ac.id

Pendidikan

- Tahun 2008-2014 : SDN 177 PEKANBARU
- Tahun 2014-2017 : SMPN 35 PEKANBARU
- Tahun 2017-2020 : SMAN 14 PEKANBARU
- Tahun 2020-2024 : Universitas Islam Negeri Sultan Syarif Kasim Riau
 Jurusan Teknik Informatika