

***WORD SENSE DISAMBIGUATION (WSD) PADA KALIMAT  
BAHASA INDONESIA MENGGUNAKAN ALGORITMA  
SIMPLIFIED LESK***

**Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Jurusan Teknik Informatika**



Oleh :

Nurul Akhni  
NIM : 09021181924024

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2023**

**LEMBAR PENGESAHAN SKRIPSI**

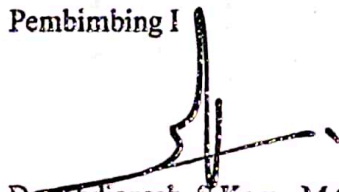
**WORD SENSE DISAMBIGUATION (WSD) PADA KALIMAT BAHASA  
INDONESIA MENGGUNAKAN ALGORITMA *SIMPLIFIED LESK***

Oleh :

Nurul Akhni  
NIM : 09021181924024

Inderalaya, 2 Agustus 2023

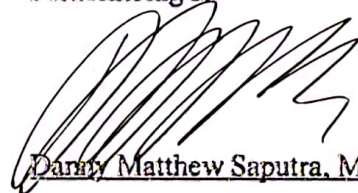
Pembimbing I



Dr. Abdiansah, S.Kom., M.CS

NIP. 198410012009121005

Pembimbing II



Denny Matthew Saputra, M.Sc.

NIP. 198505102015041002

Mengetahui,

Ketua Jurusan Teknik Informatika



Alvi Svahrini Utami, M.Kom.

NIP. 197812222006042003

## TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Senin tanggal 31 Juli 2023 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Nurul Akhni

NIM : 09021181924024

Judul : *Word Sense Disambiguation (WSD) pada Kalimat Bahasa Indonesia Menggunakan Algoritma Simplified Lesk*

dan dinyatakan LULUS.

1. Ketua Penguji

Yunita, M.Cs.

NIP. 198306062015042002

2. Penguji I

Novi Yusliani, S.Kom., M.T.

NIP. 198211082012122001

3. Pembimbing I

Dr. Abdiansah, S.Kom., M.Cs.

NIP. 198410012009121005

4. Pembimbing II

Danny Matthew Saputra, M.Sc.

NIP. 198505102015041002

Mengetahui,  
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.

NIP. 19781222200642003

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Nurul Akhni

NIM : 09021181924024

Program Studi : Teknik Informatika

Judul Skripsi : *Word Sense Disambiguation (WSD) pada Kalimat Bahasa*

*Indonesia Menggunakan Algoritma Simplified Lesk*

Hasil Pengecekan Software iThenticate/Turnitin : 8%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, 3 Agustus 2023



Nurul Akhni  
NIM. 09021181924024

## **MOTTO DAN PERSEMBAHAN**

*"Untuk hari yang terasa berat, cukuplah hanya membuatmu lelah. Jangan sampai  
membuatmu menyerah"*

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Orang Tua dan Keluarga
- Teman Seperjuangan
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

## **ABSTRACT**

*The Indonesian language is known to have several words that are commonly recognized for their ambiguous nature, meaning they possess more than one meaning. Word Sense Disambiguation (WSD), a branch of Natural Language Processing (NLP), deals with the challenge of resolving this ambiguity and identifying the precise meaning of a word based on its context. Among the algorithms used for WSD, the Simplified Lesk algorithm stands out as particularly popular. To assess its effectiveness, tests were conducted using the Kamus Besar Bahasa Indonesia (KBBI) as a reference for word definitions, and a dataset of 300 Indonesian sentences containing ambiguous words and their respective meanings as determined by human perception. The research reveals that the configuration of the preprocessing phase plays a crucial role in accurately identifying the intended meaning. After evaluation, the overall accuracy achieved was 58% for the dataset, incorporating preprocessing techniques such as stemming and stopword removal.*

*Keywords: Ambiguous, Word sense disambiguation (WSD), Simplified Lesk*

## ABSTRAK

Bahasa Indonesia memiliki beberapa kata yang biasa dikenal dengan kata ambigu, yaitu memiliki lebih dari satu makna. *Word Sense Disambiguation* (WSD) merupakan salah satu topik dari *Natural Language Processing* (NLP) yang berhubungan dengan penanganan ambiguitas untuk mengidentifikasi makna kata yang tepat berdasarkan konteksnya. Salah satu algoritma yang paling populer digunakan untuk WSD yaitu algoritma *Simplified Lesk*. Pengujian dilakukan menggunakan Kamus Besar Bahasa Indonesia (KBBI) sebagai acuan *glossary* kata dan 300 data Bahasa Indonesia berupa kalimat, kata ambigu serta maknanya yang ditentukan berdasarkan persepsi manusia. Berdasarkan penelitian yang dilakukan, konfigurasi proses pra-pengolahan berperan penting dalam mengidentifikasi makna yang tepat. Hasil evaluasi menunjukkan akurasi 58% untuk keseluruhan data, dengan melibatkan proses pra-pengolahan *stemming* dan *stopword removal*.

Kata Kunci: Ambigu, *Word Sense Disambiguation* (WSD), *Simplified Lesk*

## KATA PENGANTAR

Puji dan syukur atas kehadiran Allah Subhanahu Wa Ta'ala yang telah memberikan hidayah, rahmat, dan petunjuk sehingga penyusun dapat menyelesaikan Tugas Akhir ini dengan judul “*Word Sense Disambiguation (WSD) pada Kalimat Bahasa Indonesia Menggunakan Algoritma Simplified Lesk*” tepat pada waktunya. Dalam penulisan ini, penulis menyadari bahwa tanpa adanya bimbingan, bantuan, serta dukungan maupun petunjuk dari semua pihak, tidak mungkin Tugas Akhir ini dapat diselesaikan. Oleh karena itu, pada kesempatan ini penulis menyampaikan rasa terima kasih kepada:

1. Ayah, Ibu, dan Kakak Ulfa yang selalu mendoakan dan memberi dukungan tanpa henti.
2. Bapak Alm. Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Alvi Syahrini Utami, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Abdiansah, S.Kom., M.CS selaku pembimbing Tugas Akhir dan Bapak Danny Matthew Saputra, M.Sc. selaku pembimbing Tugas Akhir serta pembimbing akademik penulis yang telah berkenan membimbing, memberikan motivasi dan saran dalam menyelesaikan Tugas Akhir.
5. Kak Ricy Firnando selaku Admin Jurusan Teknik Informatika, yang telah membantu dalam kelancaran administrasi selama masa perkuliahan.



6. Teman-teman yang sampai akhir selalu mendukung dan menemani di masa-masa sulit (Fovv): Nilam Musdalifa, Tarisa Rafika, Zafira Galea, Rani Silvani Yousnaidi, Shabrina Putri Fadhillah, Reyhani Avissa, Aulia Mabbruca Putri, M. Raihan Almenata, Asyraf Syafiyurrahman, KMH Alvianyah, Fadel Muhammad, Bintang Dwitama.
7. Teman yang selalu hadir 24/7 dan melewati semua kondisi: Nilam Musdalifa.
8. Semua pihak yang telah membantu penulis dalam menyelesaikan Tugas Akhir ini yang tidak dapat disebutkan satu per satu.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan karena keterbatasan pengetahuan dan kemampuan, oleh karena itu kritik dan saran membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata, semoga Tugas Akhir ini dapat bermanfaat bagi kita semua.

Indralaya, 3 Agustus 2023

Penulis

Nurul Akhni

## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF.....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	I-1
1.1    Pendahuluan .....	I-1
1.2    Latar Belakang.....	I-1
1.3    Rumusan Masalah .....	I-4
1.4    Tujuan Penelitian.....	I-4
1.5    Manfaat Penelitian.....	I-4
1.6    Batasan Masalah.....	I-5
1.7    Sistematika Penulisan.....	I-5
1.8    Kesimpulan.....	I-7
BAB II KAJIAN LITERATUR .....	II-1
2.1    Pendahuluan .....	II-1
2.2    Landasan Teori .....	II-1
2.2.1 <i>Word Sense Disambiguation</i> .....	II-1
2.2.2 <i>Preprocessing</i> .....	II-4
2.2.3 <i>Algoritma Simplified Lesk</i> .....	II-5

2.2.4	Metrik Evaluasi .....	II-7
2.2.5	<i>Rational Unified Process</i> .....	II-7
2.3	Penelitian Lain yang Relevan.....	II-9
2.4	Kesimpulan.....	II-10
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan .....	III-1
3.2	Pengumpulan Data.....	III-1
3.2.1	Jenis dan Sumber Data .....	III-1
3.2.2	Metode Pengumpulan Data .....	III-1
3.3	Tahapan Penelitian .....	III-2
3.3.1	Pengumpulan dan Analisis Data .....	III-3
3.3.2	Membangun Sistem WSD Menggunakan Algoritma Simplified Lesk .....	III-4
3.3.3	Melakukan Pengujian Penelitian.....	III-6
3.3.4	Melakukan Analisis dan Menarik Kesimpulan .....	III-7
3.3.5	Membuat Laporan Penelitian .....	III-8
3.4	Kesimpulan.....	III-8
BAB IV PENGEMBANGAN PERANGKAT LUNAK.....		IV-1
4.1	Pendahuluan .....	IV-1
4.2	Fase Insepsi .....	IV-1
4.2.1	Pemodelan Bisnis .....	IV-1
4.2.2	Kebutuhan Sistem .....	IV-2
4.2.3	Analisis dan Desain.....	IV-3
4.3	Fase Elaborasi.....	IV-19
4.3.1	Pemodelan Bisnis .....	IV-20
4.3.2	Kebutuhan Sistem .....	IV-22
4.3.3	Analisis dan Perancangan .....	IV-22
4.4	Fase Konstruksi .....	IV-25
4.4.1	Kebutuhan Sistem .....	IV-25

4.4.2	Implementasi .....	IV-26
4.5	Fase Transisi .....	IV-29
4.5.1	Pemodelan Bisnis .....	IV-29
4.5.2	Rencana Pengujian .....	IV-29
4.5.3	Implementasi .....	IV-30
4.6	Kesimpulan .....	IV-31
BAB V HASIL DAN ANALISIS .....		V-1
5.1	Pendahuluan .....	V-1
5.2	Hasil Penelitian .....	V-1
5.2.1	Konfigurasi Percobaan .....	V-1
5.2.2	Hasil Konfigurasi 1 .....	V-2
5.2.3	Hasil Konfigurasi 2 .....	V-9
5.2.4	Hasil Konfigurasi 3 .....	V-14
5.2.5	Hasil Konfigurasi 4 .....	V-19
5.3	Analisis Hasil Penelitian .....	V-27
5.4	Kesimpulan .....	V-35
BAB VI KESIMPULAN DAN SARAN .....		VI-1
6.1	Pendahuluan .....	VI-1
6.2	Kesimpulan .....	VI-1
6.3	Saran .....	VI-1
DAFTAR PUSTAKA .....		xiv
LAMPIRAN .....		xvii

## DAFTAR TABEL

<b>Tabel III- 1.</b> Contoh Data Kalimat.....	III-4
<b>Tabel III- 2.</b> Tabel Analisis WSD.....	III-7
<b>Tabel III- 3.</b> Tabel Hasil Pengujian .....	III-7
<b>Tabel IV- 1.</b> Kebutuhan Fungsional.....	IV-2
<b>Tabel IV- 2.</b> Kebutuhan Non-Fungsional.....	IV-2
<b>Tabel IV- 3.</b> Data Sampel.....	IV-4
<b>Tabel IV- 4.</b> Hasil Tokenisasi Kalimat.....	IV-5
<b>Tabel IV- 5.</b> Hasil <i>Stopwords Removal</i> .....	IV-6
<b>Tabel IV- 6.</b> Hasil <i>Stemming</i> .....	IV-6
<b>Tabel IV- 7.</b> Hasil Pembentukan Kamus.....	IV-7
<b>Tabel IV- 8.</b> Hasil <i>Overlap</i> dan <i>Predicted Sense</i> .....	IV-11
<b>Tabel IV- 10.</b> Definisi <i>Actor</i> .....	IV-16
<b>Tabel IV- 11.</b> Definisi <i>Use Case</i> .....	IV-17
<b>Tabel IV- 12.</b> Skenario <i>Use Case</i> Disambiguasi Makna Kata .....	IV-17
<b>Tabel IV- 13.</b> Skenario <i>Use Case</i> Menampilkan Evaluasi.....	IV-18
<b>Tabel IV- 14.</b> Implementasi Kelas .....	IV-27
<b>Tabel IV- 15.</b> Rencana Pengujian Disambiguasi Makna Kata.....	IV-29
<b>Tabel IV- 16.</b> Rencana Pengujian Menampilkan Evaluasi .....	IV-29
<b>Tabel IV- 17.</b> Hasil Pengujian Disambiguasi Makna Kata.....	IV-30
<b>Tabel IV- 18.</b> Hasil Pengujian Menampilkan Evaluasi.....	IV-30
<b>Tabel V- 1.</b> Sampel Hasil Percobaan <i>Stemming</i> .....	V-3
<b>Tabel V- 2.</b> Sampel Hasil Analisis Proses WSD dengan <i>Stemming</i> .....	V-7
<b>Tabel V- 3.</b> Hasil Evaluasi Pengujian dengan <i>Stemming</i> .....	V-9
<b>Tabel V- 4.</b> Sampel Hasil Percobaan <i>Stopword Removal</i> .....	V-10
<b>Tabel V- 5.</b> Sampel Hasil Analisis Proses WSD dengan <i>Stopword Removal</i> . .	V-13
<b>Tabel V- 6.</b> Hasil Evaluasi Pengujian dengan <i>Stopword</i> .....	V-14
<b>Tabel V- 7.</b> Sampel Hasil Percobaan <i>Stemming</i> dan <i>Stopword Removal</i> .....	V-15

<b>Tabel V- 8.</b> Sampel Hasil Analisis Proses WSD dengan <i>Stemming</i> dan <i>Stopword Removal</i> .....	V-18
<b>Tabel V- 9.</b> Hasil Evaluasi Pengujian dengan <i>Stemming</i> dan <i>Stopword</i> .....	V-19
<b>Tabel V- 10.</b> Sampel Hasil Percobaan tanpa <i>Stemming</i> dan <i>Stopword Removal</i> .....	V-20
<b>Tabel V- 11.</b> Sampel Hasil Analisis Proses WSD tanpa <i>Stemming</i> dan <i>Stopword Removal</i> .....	V-26
<b>Tabel V- 12.</b> Hasil Pengujian Tanpa <i>Stemming</i> dan <i>Stopword</i> .....	V-27
<b>Tabel V- 13.</b> Sampel Proses Identifikasi WSD .....	V-29
<b>Tabel V- 14.</b> Sampel Hasil Pengujian WSD .....	V-32

## DAFTAR GAMBAR

<b>Gambar II- 1.</b> Arsitektur Umum WSD (Patwardhan et al., 2005) .....	II-2
<b>Gambar II- 2.</b> Alur <i>Simplified Lesk</i> (Aliwy & Abbas, 2015).....	II-6
<b>Gambar II- 3.</b> Arsitektur RUP .....	II-8
<b>Gambar III- 1.</b> Rincian Kegiatan Penelitian .....	III-3
<b>Gambar III- 2.</b> Arsitektur Sistem WSD .....	III-5
<b>Gambar IV- 1.</b> <i>Use Case Diagram</i> .....	IV-16
<b>Gambar IV- 2.</b> Rancangan Antarmuka Disambiguasi Makna Kata .....	IV-21
<b>Gambar IV- 3.</b> Rancangan Antarmuka Menampilkan Evaluasi.....	IV-21
<b>Gambar IV- 4.</b> Diagram <i>Activity</i> Disambiguasi Makna Kata.....	IV-23
<b>Gambar IV- 5.</b> Diagram <i>Activity</i> Menampilkan Evaluasi .....	IV-23
<b>Gambar IV- 6.</b> Diagram <i>Sequence</i> Disambiguasi Makna Kata.....	IV-24
<b>Gambar IV- 7.</b> Diagram <i>Sequence</i> Menampilkan Evaluasi .....	IV-25
<b>Gambar IV- 8.</b> Diagram Kelas .....	IV-26
<b>Gambar IV- 9.</b> Implementasi Antarmuka Disambiguasi Makna Kata .....	IV-28
<b>Gambar IV- 10.</b> Implementasi Antarmuka Menampilkan Evaluasi.....	IV-28
<b>Gambar V- 1.</b> Perbandingan Hasil Konfigurasi .....	V-28

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Kode Program.....	xvii
<b>Lampiran 2.</b> Hasil Evaluasi Sampel 100 Data.....	xiii



# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Bab ini akan memuat landasan dalam penelitian yang meliputi latar belakang, rumusan dan batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

### **1.2 Latar Belakang**

Bahasa Indonesia memiliki beberapa kata yang biasa dikenal dengan kata-kata yang ambigu, yaitu makna sebuah kalimat atau pernyataan membingungkan yang kurang dipahami atau bahkan tidak disampaikan. Hal ini berbeda dengan persepsi manusia yang memiliki kemampuan linguistik untuk menentukan makna ambiguitas atau lebih dari sekadar makna kata (Basuki et al., 2019). Kalimat yang bermakna terdiri dari kata-kata yang bermakna, setiap sistem diharapkan dapat memproses bahasa alami seperti yang dilakukan manusia, yaitu harus memiliki informasi tentang kata-kata dan makna (Hung, 2005).

*Word Sense Disambiguation* (WSD) merupakan bagian dari bidang Pengolahan Bahasa Alami (*Natural Language Processing/NLP*) yang berhubungan dengan penanganan ambiguitas. WSD merupakan proses komputasi linguistik yang bertujuan untuk mengidentifikasi makna kata yang tepat berdasarkan konteksnya (Basuki et al., 2019). Sebagai contoh, kata “buah” dapat memiliki makna berbeda seperti ‘bagian tanaman yang berasal dari bunga atau putik’ dan ‘beragam kategori untuk berbagai macam benda’. Kalimat apa pun

yang mengandung kata tersebut tidak dapat diterjemahkan ke bahasa lain tanpa mengetahui makna yang benar dalam kalimat, karena menemukan makna yang tepat untuk kata ambigu yang sesuai konteks bukanlah tugas yang mudah, oleh karena itu banyak algoritma dan teknik yang diusulkan untuk menyelesaikan permasalahan tersebut (Pal & Saha, 2015).

Adapun beberapa algoritma untuk permasalahan WSD, yaitu berdasarkan pendekatan *supervised* dan *unsupervised*, yang terdiri dari algoritma *Decision Tree*, *Naïve Bayes*, *Neural Network*, dan *Support Vector Machine* (SVM). Selain itu, terdapat pendekatan *knowledge-based* yang terdiri dari banyak metode salah satunya *Lesk*, di mana algoritma *Lesk* adalah algoritma yang paling banyak digunakan untuk memecahkan masalah WSD karena memiliki cara kerja yang sederhana, dan dapat secara efektif menangani kata-kata dengan banyak makna (Aliwy & Taher, 2019). Berdasarkan pernyataan tersebut, maka identifikasi WSD pada penelitian ini akan mengimplementasikan konsep dari algoritma *Lesk*.

Algoritma *Lesk* merupakan metode *knowledge-based* yang populer untuk WSD. Algoritma ini menggunakan definisi kamus (*gloss*) untuk membedakan kata polisemi dalam konteks kalimat. Algoritma *Lesk* memiliki 2 variasi yang terkenal yaitu *Original* dan *Simplified*. Keuntungan utama algoritma *Simplified Lesk* dibandingkan algoritma *Original Lesk* adalah bahwa algoritma ini jauh lebih cepat untuk dijalankan, karena memiliki kompleksitas waktu komputasi yang jauh lebih rendah dan akurasi yang lebih besar. Selain itu, algoritma tersebut juga lebih akurat dalam membedakan makna kata (Aliwy & Abbas, 2015), sehingga

algoritma *Simplified Lesk* akan diterapkan dalam meneliti *Word Sense Disambiguation* (WSD) pada kalimat Bahasa Indonesia.

Penetapan algoritma *Simplified Lesk* (variasi dari *Lesk*) yang akan diimplementasikan juga didukung oleh penelitian sebelumnya yang telah membuktikan keberhasilannya dalam menangani permasalahan WSD, seperti penelitian oleh (Basuki et al., 2019) yang menerapkan algoritma *Lesk* untuk implementasi WSD dalam menentukan makna kata homograf, dan menghasilkan akurasi sebesar 78,6%. Penelitian lainnya juga dilakukan oleh (Fazar & Widiastuti, 2017), dimana implementasi WSD menggunakan algoritma *Lesk* berperan dalam melengkapi metode utamanya yaitu *Lexical Chain* dalam memeriksa ambiguitas untuk mendapatkan akurasi dalam menentukan *Lexical Chain* yang sesuai dengan konteks kalimat. Hasil pengujian WSD menggunakan algoritma *Lesk* pada penelitian tersebut menghasilkan akurasi sebesar 77%.

Berdasarkan penjabaran sebelumnya, dapat disimpulkan bahwa penelitian mengenai *Word Sense Disambiguation* (WSD) dengan penerapan algoritma *Simplified Lesk* menunjukkan prospek sebagai perkembangan yang menjanjikan dari penelitian-penelitian sebelumnya yang menggunakan algoritma *Lesk*. Melalui pendekatan yang lebih efektif dan terarah, penelitian ini berpotensi untuk menghadirkan kemajuan dalam memahami makna kata-kata yang ambigu dalam berbagai konteks. Harapannya, hasil dari penelitian ini akan mengoptimalkan akurasi dan kesuksesan proses WSD, serta berkontribusi secara ilmiah dalam mengembangkan aplikasi NLP (*Natural Language Processing*) yang lebih canggih dan tepat sasaran.

### 1.3 Rumusan Masalah

Berikut beberapa masalah yang dirumuskan pada penelitian ini:

1. Bagaimana mengembangkan sistem *Word Sense Disambiguation* (WSD) menggunakan algoritma *Simplified Lesk* pada kalimat Bahasa Indonesia?
2. Bagaimana kinerja yang dihasilkan oleh sistem *Word Sense Disambiguation* (WSD) pada kalimat Bahasa Indonesia menggunakan algoritma *Simplified Lesk*?

### 1.4 Tujuan Penelitian

Beberapa tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Menghasilkan sistem yang dapat mengimplementasikan *Word Sense Disambiguation* (WSD) pada kalimat Bahasa Indonesia menggunakan algoritma *Simplified Lesk*
2. Mengetahui kinerja yang dihasilkan oleh sistem *Word Sense Disambiguation* (WSD) pada kalimat Bahasa Indonesia menggunakan algoritma *Simplified Lesk*.

### 1.5 Manfaat Penelitian

Berikut beberapa manfaat dari penelitian ini:

1. Mengatasi ambiguitas penafsiran pada kalimat Bahasa Indonesia yang mempunyai pemaknaan dan pemahaman yang luas

2. Hasil penelitian dapat dijadikan rujukan untuk penelitian terkait di masa yang akan datang

### **1.6 Batasan Masalah**

Batasan masalah dari penelitian ini adalah:

1. Kalimat Bahasa Indonesia yang menjadi fokus pada penelitian ini merupakan kalimat Bahasa Indonesia baku dan formal
2. Algoritma *Simplified Lesk* akan diimplementasikan bersama Kamus Besar Bahasa Indonesia (KBBI) sebagai acuan *glossary* kata

### **1.7 Sistematika Penulisan**

Tata cara penulisan yang akan diikuti adalah sesuai dengan standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya, yang dijelaskan sebagai berikut.

## **BAB I. PENDAHULUAN**

Bab ini akan memuat landasan dalam penelitian yang meliputi latar belakang, rumusan dan batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

## **BAB II. KAJIAN LITERATUR**

Topik yang akan dibahas pada bab ini yaitu kajian literatur yang menjadi dasar penelitian ini. Pembahasan akan memuat tentang *Word Sense*

*Disambiguation* (WSD), Algoritma *Simplified Lesk*, dan *Rational Unified Process* (RUP). Selain itu, terdapat penjelasan mengenai penelitian terkait.

### **BAB III. METODOLOGI PENELITIAN**

Bagian ini akan memuat pembahasan terkait pengumpulan data, langkah-langkah penelitian dan metode pengembangan sistem. Pada setiap tahap penelitian dan pengembangan perangkat lunak, bab ini akan menjadi panduan yang digunakan untuk mengatasi masalah yang dirumuskan dan mencapai tujuan penelitian.

### **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Bab ini memuat uraian terkait pengembangan perangkat lunak menggunakan metode *Rational Unified Process* (RUP) yang terdiri dari 4 fase yaitu Insepsi, Elaborasi, Konstruksi, dan Transisi.

### **BAB V. HASIL DAN ANALISIS PENELITIAN**

Bagian ini akan menampilkan evaluasi pengujian *Word Sense Disambiguation* (WSD) menggunakan algoritma *Simplified Lesk*, serta hasil analisis berdasarkan evaluasi.

### **BAB VI. KESIMPULAN DAN SARAN**

Bagian ini membahas mengenai kesimpulan dan saran berdasarkan analisis sebelumnya untk dijadikan rujukan penelitian di masa depan khususnya di bidang *Word Sense Disambiguation* (WSD).

## **1.8 Kesimpulan**

Bab ini telah menjelaskan tentang konteks dari penelitian dan referensi penting yang menjadi landasan, termasuk latar belakang, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan tata cara penulisan.

## **BAB II**

### **KAJIAN LITERATUR**

#### **2.1 Pendahuluan**

Topik yang akan dibahas pada bab ini yaitu kajian literatur yang menjadi dasar penelitian ini. Pembahasan akan memuat tentang *Word Sense Disambiguation* (WSD), Algoritma *Simplified Lesk*, Metrik Evaluasi, dan *Rational Unified Process* (RUP). Selain itu, terdapat penjelasan mengenai penelitian terkait.

#### **2.2 Landasan Teori**

##### **2.2.1 *Word Sense Disambiguation***

Ambiguitas adalah makna ganda dari kalimat yang diucapkan oleh pembicara sehingga mereka ragu atau tidak dipahami oleh pendengar sama sekali (bahasa lisan), maupun dalam bahasa tertulis. Ambiguitas dapat muncul dalam berbagai macam tulisan (Trismanto, 2018).

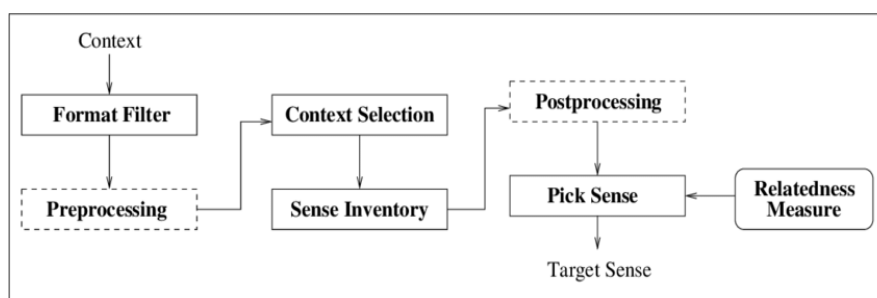
Ambiguitas leksikal adalah salah satu dari jenis ambiguitas yang terjadi ketika satu kata dikaitkan dengan banyak makna. Resolusi dari ambiguitas leksikal, yang biasa disebut "*Word Sense Disambiguation*" (WSD), diharapkan dapat meningkatkan kualitas dari beberapa penelitian seperti *Machine Translation*, *Parsing*, *Natural Language Understanding* (NLU) dan *lexicography* (Fujii & Tanaka, 1998).

*Word Sense Disambiguation* (WSD) adalah sub-bidang dalam pemrosesan bahasa alami, di mana sistem komputer dirancang untuk mengidentifikasi makna



yang benar dari sebuah kata dalam konteks tertentu (Mccarthy, 2009). Masalah WSD dapat dibagi menjadi dua bagian. Pertama, bagaimana mengungkapkan apa makna 1 dan makna 2 dari sebuah kata, dengan cara yang dapat ditafsirkan oleh algoritma WSD. Kedua, bagaimana mengetahui makna mana yang cocok dengan contoh kata yang akan disambiguasi (Kilgarriff, 1998).

Ada beberapa pendekatan yang telah diusulkan untuk memecahkan masalah WSD, yang paling banyak digunakan yaitu metode berbasis pengetahuan (*knowledge-based*), yaitu menggunakan sumber daya eksternal seperti kamus yang dapat dibaca mesin, kamus thesauri serta bilingual, dan metode berbasis korpus (*corpus-based*) yang sebagian besar menggunakan *supervised machine learning*, dimana setiap kemunculan kata ambigu diberi anotasi dengan label semantik dengan mempertimbangkan konteks kata tersebut (*hand-label training data*). Pada penelitian ini menerapkan pendekatan *knowledge-based* khususnya metode *dictionary-based*, dimana WSD dilakukan dengan cara menemukan tumpang tindih antara tanda kata-kata ambigu dengan tanda kata-kata sekitarnya dalam konteks (Martirosyan, 2005). Adapun arsitektur umum dari sistem WSD dapat dilihat pada Gambar II-1 berikut.



**Gambar II- 1.** Arsitektur Umum WSD (Patwardhan et al., 2005)

Gambar II-1 menunjukkan gambaran umum tentang arsitektur dari sistem, dan menampilkan berbagai sub-tugas yang perlu dilakukan untuk melaksanakan *Word Sense Disambiguation* (WSD). Sub-tugas yang berada dalam kotak putus-putus adalah opsional, artinya tidak semua sub-tugas harus dilakukan. Selain itu, setiap sub-tugas dapat dilaksanakan dengan berbagai cara yang berbeda, sehingga dapat disesuaikan dengan berbagai kebutuhan disambiguasi yang berbeda.

Penjelasan mengenai tahap-tahap utama pada Gambar II-1, yaitu mulai dari tahap *Format Filter* yang menetapkan tipe masukan user sesuai dengan kebutuhan dari proses algoritma yang akan dijalankan. Pada penelitian ini menggunakan tipe masukan teks. Kemudian dilakukan *Context Selection* untuk menghasilkan konteks yang akan diproses dalam menentukan relasi makna. Pada penelitian ini, tahap *Context Selection* akan menjadi satu dalam tahap *Preprocessing* dalam menghasilkan konteks. Pada tahap *Sense Inventory*, kandidat-kandidat makna yang akan dibandingkan dengan konteks akan diperoleh dari kamus data, dimana dalam penelitian ini menggunakan Kamus Besar Bahasa Indonesia (KBBI). Tahap *Postprocessing* merupakan tahap opsional dimana tahap ini melakukan *preprocessing* lanjutan seperti *filtering* terhadap makna-makna yang diperoleh dari kamus. Sebagai contoh, kita dapat mengecualikan makna kata kerja dari kata ambigu pada proses disambiguasi. Namun, penelitian ini akan menggunakan semua jenis makna sehingga tahap ini tidak diimplementasikan. Selanjutnya, tahap *Pick Sense* akan menggunakan informasi dari konteks dan kandidat makna yang diperoleh dari kamus, untuk memilih makna sebagai hasil

makna yang tepat. Tahap akhir yang juga akan diimplementasikan pada penelitian ini berdasarkan proses dalam arsitektur WSD diatas yaitu *Relatedness Measure*, atau yang dapat disebut juga sebagai evaluasi ketepatan dari makna yang dihasilkan.

### 2.2.2 *Preprocessing*

*Preprocessing* bertanggung jawab untuk mengambil kueri pengguna (instans / contoh yang berisi kata target) dan mengonversi kueri ini menjadi teks terstruktur. Proses *preprocessing* dimulai dengan memecah teks target menjadi kumpulan kalimat. Selanjutnya, setelah tokenisasi, langkah-langkah berikutnya adalah menghilangkan kata-kata yang tidak relevan (*stopwords removal*) dan menormalkan kata-kata ke bentuk dasarnya (*stemming*) (Abdo, 2023).

Tokenisasi adalah proses memecah aliran teks menjadi kata-kata, frasa, simbol, atau elemen bermakna lainnya yang disebut token. Tujuan dari tokenisasi adalah eksplorasi kata-kata dalam sebuah kalimat (Kannan et al., 2014).

*Stopword Removal* adalah proses menghapus kata-kata umum yang sering digunakan dalam banyak kalimat, sedangkan *stemming* adalah teknik umum yang digunakan dalam pencarian informasi untuk menghilangkan variasi morfologi untuk mendapatkan bentuk dasar dari sebuah kata. Penerapan *stemming* dan *stopword removal* dapat mempengaruhi kinerja sistem WSD. *Stemming* digunakan agar kata-kata dengan bentuk morfologi yang berbeda dapat dihitung sebagai kata-kata konten yang sama. Selain itu, *stopwords removal* digunakan untuk mencegah matriks yang mewakili kata-kata konten menjadi terlalu jarang,

serta untuk menghapus kata-kata yang tidak penting dari kata-kata konten (Mahendra et al., 2018).

### 2.2.3 Algoritma *Simplified Lesk*

Salah satu algoritma terdahulu dan paling populer yang menggunakan kamus yang dapat dibaca mesin adalah algoritma *Lesk*. Diterbitkan pada tahun 1986, ide di balik algoritma ini adalah untuk mengukur tumpang tindih antara definisi makna kata-kata dalam suatu konteks (Craggs, 2011).

Algoritma *Lesk* menggunakan definisi kamus (*gloss* / contoh) untuk mendisambiguasikan kata polisemi dalam konteks kalimat. Algoritma *Lesk* asli mengukur tumpang tindih antara definisi makna untuk semua kata dalam teks dan mengidentifikasi secara bersamaan makna yang benar untuk semua kata dalam teks (Aliwy & Abbas, 2015). Versi lain dari algoritma *Lesk* adalah algoritma *Simplified Lesk* yang akan diterapkan pada penelitian ini, dan merupakan variasi yang disederhanakan yang menjalankan proses WSD terpisah untuk setiap kata ambigu dalam teks input. Dalam algoritma *simplified* ini, makna yang benar dari setiap kata dalam sebuah teks ditentukan secara individual dengan menemukan pengertian yang mengarah pada tumpang tindih antara definisi kamus dan konteks saat ini (Mihalcea, 2007).

Berdasarkan cara kerja secara umum dari algoritma *Simplified Lesk* pada uraian sebelumnya, dapat dinyatakan bahwa variasi *Simplified Lesk* lebih sederhana dengan algoritma *Lesk* yang asli. Kita dapat melihat kompleksitas algoritma *Lesk* asli dari contoh berikut: jika kita memiliki sembilan kelas kata

dengan jumlah makna sebagai berikut: 26, 11, 4, 8, 5, 4, 10, 8, 3 maka kombinasi makna adalah 43.929.600, yang merupakan angka yang sangat besar dan sulit untuk diimplementasikan secara praktis. Itulah alasan utama mengapa digunakan versi lain dari algoritma Lesk, yaitu *Simplified Lesk* (Aliwy & Abbas, 2015). Adapun gambaran umum dari alur algoritma *Simplified Lesk* dapat dilihat pada Gambar II-2.

```

function Simplified_Lesk(word, sentence)
  best-sense ← most frequent sense for word
  max-overlap ← 0
  context ← set of words in sentence
  for each sense in senses of word do
    signature ← set of words in the gloss_examples of sense
    overlap ← Compute_overlap(signature, context)
    if overlap > max-overlap then
      max-overlap ← overlap
      best-sense ← sense
  end
  return(best-sense) { returns best sense of word }

```

**Gambar II- 2.** Alur *Simplified Lesk* (Aliwy & Abbas, 2015)

Sebelum masuk ke dalam proses algoritma, kalimat diambil satu per satu dari teks yang kemudian dilakukan *pre-processing* untuk menghasilkan konteks. Selanjutnya, berdasarkan Gambar II-2, masukan kata ambigu akan digunakan dalam mencari definisi kata pada kamus. Definisi kamus (*gloss*) dari semua kata-kata yang bermakna ini ditinjau dan dilakukan operasi irisan (*intersection*) antara setiap makna dari kata yang akan disambiguasi, lalu menentukan jumlah kata yang sama yang muncul dalam konteks dan juga muncul dalam *gloss* dari makna tersebut (Pal & Saha, 2014). Makna yang memiliki jumlah irisan kata yang paling

banyak dengan konteks dianggap sebagai makna yang paling sesuai (Pal et al., 2013).

#### 2.2.4 Metrik Evaluasi

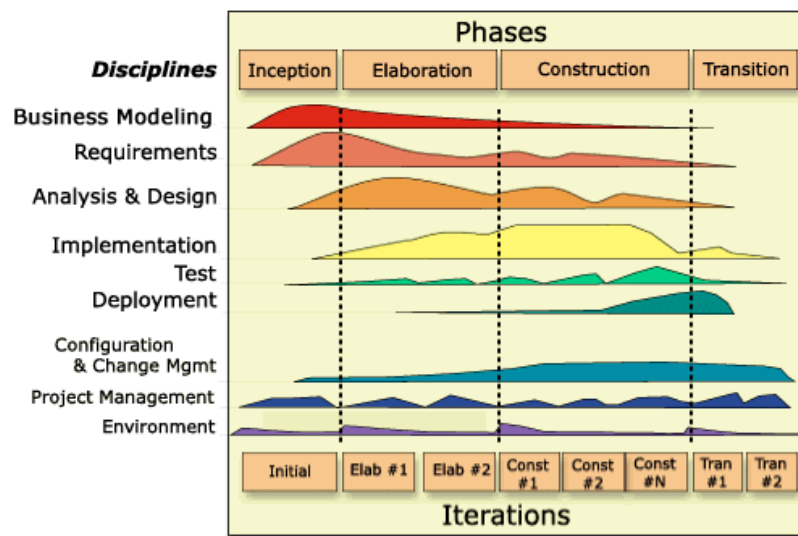
Saat ini, standar untuk evaluasi algoritma WSD adalah kriteria "pencocokan yang tepat" (Resnik & Yarowsky, 1997) dengan menghitung akurasi keseluruhan, di mana jumlah instans prediksi makna yang benar dibagi dengan jumlah total instans (Torres et al., 2009). Prediksi makna benar yaitu prediksi makna yang sesuai dengan makna sebenarnya. Rumus perhitungan akurasi didefinisikan sebagai berikut.

$$Akurasi = \frac{Jumlah\ Prediksi\ Makna\ Benar}{Jumlah\ Data} \times 100\% \quad (II-1)$$

#### 2.2.5 Rational Unified Process

*Rational Unified Process* (RUP) adalah pendekatan perangkat lunak yang berulang, arsitektur-sentris, dan lebih didorong oleh *use case* (Mubarok & Hadijah, 2015). Tujuannya yaitu untuk menjamin bahwa pengembangan perangkat lunak mencapai standar kualitas yang tinggi dan memenuhi persyaratan dari para pengguna. Pendekatan RUP terdiri dari empat tahap utama, yakni *Inception*, *Elaboration*, *Construction*, dan *Transition*. Tiap tahap melibatkan satu atau lebih siklus berulang dan dilengkapi dengan tonggak penting. Pada setiap tahap ini, kemajuan proyek dievaluasi, dan keputusan diambil berdasarkan

perkembangan tersebut. Tujuan dari setiap iterasi dalam setiap tahap adalah menghasilkan hasil teknis yang sesuai dengan tujuan tahap tersebut (Anwar, 2014).



**Gambar II- 3.** Arsitektur RUP

Berikut ini penjelasan dari setiap fase berdasarkan Gambar II-2.

1. Fase Insepsi (*Inception*)

Fase ini merupakan langkah pertama dalam pengembangan perangkat lunak. Proses pada tahap awal ini mencakup mengumpulkan data, menetapkan lingkup analisis, dan merancang awal perangkat lunak.

## 2. Fase Elaborasi (*Elaboration*)

Pada fase ini perancangan perangkat lunak akan dilakukan dengan menentukan spesifikasi, melakukan analisis, serta membuat desain teknis perangkat lunak.

## 3. Fase Konstruksi (*Construction*)

Tahap ini terdiri dari beberapa langkah, yakni menciptakan aplikasi atau sistem, serta menguji sistem yang telah dibangun.

## 4. Fase Transisi (*Transition*)

Fase ini merupakan tahap akhir pada RUP yang meliputi perbaikan perangkat lunak, serta menarik kesimpulan dari pengembangan perangkat lunak.

### **2.3 Penelitian Lain yang Relevan**

Dalam menyusun penelitian ini, banyak didukung dengan penelitian terkait yang relevan dengan latar belakang masalah serupa.

Penelitian yang telah dilakukan oleh (Basuki et al., 2019) yang menerapkan algoritma *Lesk* pada sistem pendefinisian makna kata yang sesuai dengan kata homograf, menghasilkan akurasi sebesar 78,6% dengan melakukan pengujian terhadap 140 kalimat Bahasa Indonesia yang diperoleh dari KBBI dan *end-user*.

Penelitian lainnya telah dilakukan oleh (Fazar & Widiastuti, 2017), dimana pada penelitian tersebut implementasi WSD menggunakan algoritma *Lesk*



berperan dalam melengkapi metode utama yaitu *Lexical Chain* dalam memeriksa ambiguitas untuk mencapai tingkat ketepatan yang tinggi dalam mengidentifikasi *Lexical Chain* yang sesuai dengan konteks kalimat. Hasil pengujian WSD menggunakan algoritma *Lesk* yaitu terdapat 46 kata yang dapat ditafsirkan dengan lebih dari satu arti (ambigu) dari total 60 kata yang ambigu, dan hasilnya menunjukkan akurasi sebesar 77%.

Penelitian lainnya dilakukan oleh (Ridwansyah, 2021) yang mengimplementasikan WSD dengan algoritma *Lesk* untuk digabungkan dengan metode utamanya yaitu *Non-Negative Matrix Factorization* untuk mengatasi kata-kata yang serupa namun sebenarnya memiliki makna yang berbeda untuk menentukan makna pada kalimat tersebut agar lebih dimengerti oleh pembaca. Berdasarkan hasil pengujian pada hasil ringkasan sistem dengan menerapkan WSD dan algoritma *Lesk*, mendapatkan akurasi yang cukup baik dengan nilai tertinggi precision dengan menggunakan metode sebesar 75,8%, recall sebesar 78,9%, dan f-measure sebesar 75,8%.

## 2.4 Kesimpulan

Bab ini telah memuat pembahasan mengenai kajian literatur yang menjadi dasar dalam penelitian ini. Pengertian atau teori yang dibahas mulai dari ambiguitas, *Word Sense Disambiguation (WSD)*, Algoritma *Simplified Lesk*, dan Metrik Evaluasi sebagai alat ukur pengujian dari algoritma yang diterapkan. Selain itu, disertai juga pembahasan mengenai penelitian terkait yang relevan sebagai validasi dari penelitian yang akan dilakukan.

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Pendahuluan**

Bagian ini akan memuat pembahasan terkait pengumpulan data, langkah-langkah penelitian dan metode pengembangan sistem. Pada setiap tahap penelitian dan pengembangan perangkat lunak, bab ini akan menjadi panduan yang digunakan untuk mengatasi masalah yang dirumuskan dan mencapai tujuan penelitian.

#### **3.2 Pengumpulan Data**

##### **3.2.1 Jenis dan Sumber Data**

Pada penelitian ini, jenis data yang digunakan merupakan 300 data sekunder berupa kalimat Bahasa Indonesia pada publikasi ilmiah dan *website* artikel resmi.

##### **3.2.2 Metode Pengumpulan Data**

Data 140 kalimat dan kata ambigu Bahasa Indonesia diperoleh melalui pengumpulan kalimat Bahasa Indonesia yang dijadikan data uji pada sebuah publikasi ilmiah, yang diunduh secara manual pada *UMM Institutional Repository*<sup>1</sup> dan dipublikasikan secara resmi pada *IEEE Xplore*<sup>2</sup>. Kemudian

---

<sup>1</sup> <https://eprints.umm.ac.id/58164/>

<sup>2</sup> <https://ieeexplore.ieee.org/abstract/document/8850957>

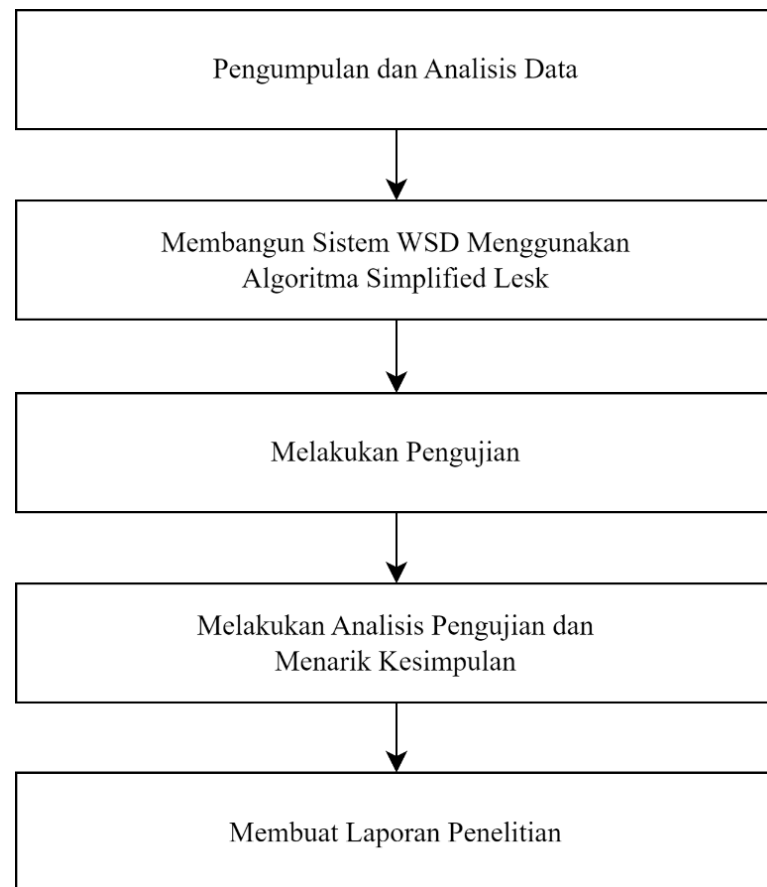
ditambahkan 160 data kalimat yang diperoleh dari beberapa artikel resmi pada menu Artikel di Website Badan Pengembangan dan Pembinaan Bahasa Kemendikbud<sup>3</sup> dengan menentukan kata ambigunya secara mandiri yang disesuaikan dengan kriteria yang dibutuhkan. Penambahan 160 data bertujuan untuk memperluas jumlah sampel yang digunakan dalam penelitian ini. Data yang lebih luas diharapkan dapat memberikan gambaran yang lebih akurat tentang kemampuan algoritma *Simplified Lesk* dalam menentukan makna kata-kata yang ambigu serta memberikan wawasan yang lebih mendalam tentang keandalan dan efektivitas WSD menggunakan algoritma *Simplified Lesk* dalam menghadapi dataset yang lebih besar dan beragam.

### **3.3 Tahapan Penelitian**

Tahapan penelitian merupakan langkah-langkah atau prosedur yang digunakan sebagai landasan dalam memecahkan masalah yang akan diteliti. Adapun rincian kegiatan yang akan dilakukan dapat dilihat pada Gambar III-1.

---

<sup>3</sup> <https://badanbahasa.kemdikbud.go.id/artikel>



**Gambar III- 1.** Rincian Kegiatan Penelitian

### 3.3.1 Pengumpulan dan Analisis Data

Sebelum menentukan dan mengumpulkan data yang akan digunakan sebagai bahan uji pada penelitian ini, dilakukan seleksi terlebih dahulu terhadap *dataset* berdasarkan beberapa kriteria berikut.

1. Kalimat pada dataset merupakan kalimat Bahasa Indonesia baku
2. Kalimat pada dataset terdiri dari kata-kata yang memiliki banyak definisi dan biasanya digunakan dalam berbagai konteks yang berbeda.

Setelah menemukan dataset yang telah memenuhi kriteria-kriteria tersebut, selanjutnya dilakukan pengumpulan 300 data kalimat dan kata ambigu Bahasa

Indonesia dengan menyusun dan mengambil data yang telah disesuaikan dengan aturan Kamus Besar Bahasa Indonesia (KBBI), kemudian disatukan kedalam satu *file excel*. Adapun contoh data Kalimat dan Kata Ambigu yang digunakan dapat dilihat pada Tabel III-1.

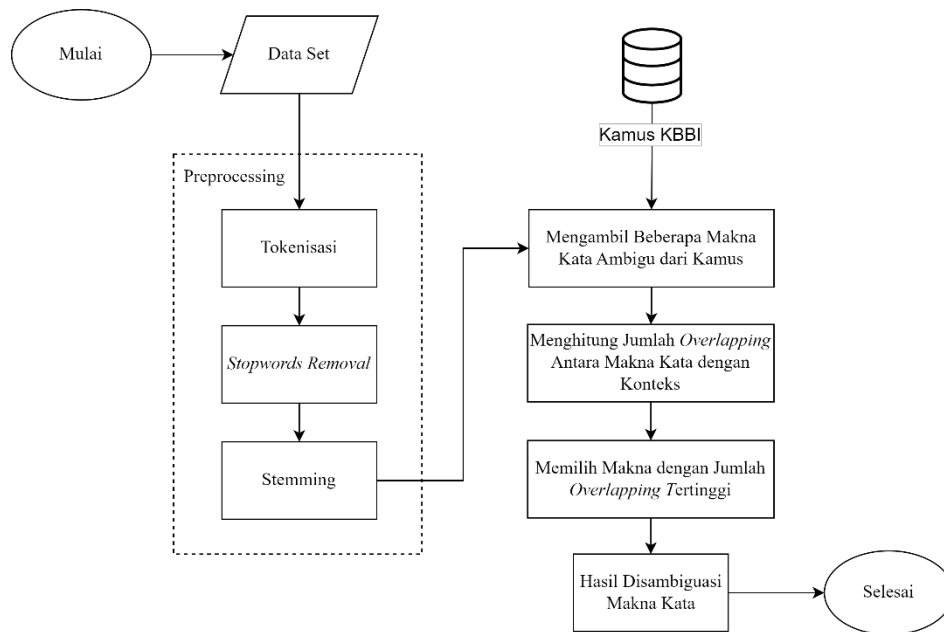
**Tabel III- 1.** Contoh Data Kalimat

Kalimat Bahasa Indonesia	Kata Ambigu
aku membaca <i>buku</i> tentang sejarah di perpustakaan	buku
tebu itu mempunyai banyak <i>buku</i>	buku
bukan hanya pembangunan fisik yang diperhatikan, melainkan juga pembangunan <i>mental</i>	mental
akibat tabrakan yang terjadi tubuh itu <i>mental</i> sekitar 2 meter	mental

### 3.3.2 Membangun Sistem WSD Menggunakan Algoritma Simplified Lesk

Metode yang diterapkan dalam membangun sistem WSD menggunakan algoritma *Simplified Lesk* yaitu metode *Rational Unified Process* (RUP). Dengan menggabungkan metode RUP dan algoritma *Simplified Lesk*, sistem WSD yang dibangun pada penelitian ini dapat disusun, direncanakan, dan dilaksanakan secara sistematis dan terorganisir.

Arsitektur umum sistem WSD yang dibangun pada penelitian ini dapat dilihat pada Gambar III-2.



**Gambar III- 2.** Arsitektur Sistem WSD

Proses yang dilakukan dalam membangun sistem WSD dengan menggunakan algoritma *Simplified Lesk* berdasarkan gambar diatas, akan melalui beberapa tahapan. Tahap pertama yaitu *preprocessing* yang bertujuan untuk mempersiapkan data sebelum diproses lebih lanjut. Pada tahap ini data akan melalui 3 proses yaitu *Tokenizing* (Tokenisasi), *Stopwords Removal*, dan *Stemming*.

Pada tahap selanjutnya kata ambigu yang ditemukan dalam kalimat pada data akan dijadikan target kata. Beberapa makna dari kata tersebut akan diambil untuk dilakukan penentuan makna sesuai dengan konteks kalimat. Selanjutnya setiap makna kata akan diproses dengan menghitung jumlah tumpang tindih (*overlapping*) antara makna dengan konteks. Makna dengan jumlah *overlapping* tertinggi akan dipilih sistem sebagai prediksi makna yang sesuai dari kata ambigu.

Konsep yang dijadikan kriteria pengujian pada penelitian ini yaitu pengujian klasifikasi dengan menghitung rata-rata dari jumlah data pada dataset yang digunakan pada nilai akurasi dari prediksi makna pada kata ambigu dalam kalimat Bahasa Indonesia yang telah dihasilkan sistem. Pengujian dilakukan dengan membandingkan prediksi makna yang telah dihasilkan dengan makna sebenarnya. Konfigurasi pada pengujian yaitu dengan menentukan makna dari kata yang memiliki jumlah tumpang tindih (*overlapping*) terbanyak dengan konteks.

Alat bantu yang digunakan pada penelitian WSD pada kalimat Bahasa Indonesia menggunakan Algoritma *Simplified Lesk* ini adalah sebagai berikut.

#### 1. Perangkat Keras

Prosesor : Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz

RAM : 8 GB

SSD : 512 GB

#### 2. Perangkat Lunak

Bahasa Pemrograman : Python

Framework : Flask

Pustaka Eksternal : nltk, numpy, kbabi, sklearn, sastrawi, pandas

### 3.3.3 Melakukan Pengujian Penelitian

Sistem yang telah dibuat nantinya akan diproses menggunakan data kalimat Bahasa Indonesia yang akan menghasilkan makna (*sense*) yang sesuai

dari kata ambigu pada kalimat. Pengujian dilakukan dengan membandingkan prediksi makna (*predicted sense*) yang telah dihasilkan dengan makna sebenarnya (*actual sense*). Konfigurasi pada pengujian yaitu dengan menentukan proses yang akan digunakan dalam menghasilkan konteks yang berperan penting dalam menghasilkan makna yang tepat. Hasil kinerja pengujian akan dibandingkan untuk mendapatkan kinerja terbaik.

### 3.3.4 Melakukan Analisis dan Menarik Kesimpulan

Kinerja *Word Sense Disambiguation* (WSD) menggunakan algoritma *Simplified Lesk* akan melalui perhitungan jumlah data yang digunakan, serta jumlah prediksi benar dan salah yang dihasilkan (akan dijelaskan secara rinci pada Bab V). Selanjutnya akan dicatat kedalam tabel hasil pengujian yang dapat dilihat pada Tabel III-3 untuk memperoleh nilai akurasi yang digunakan untuk mengukur kinerja dari WSD menggunakan algoritma *Simplified Lesk*. Kemudian kesimpulan penelitian akan dibuat berdasarkan hasil analisis pengujian.

**Tabel III- 2.** Tabel Analisis WSD

No	Kalimat	Kata Ambigu	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan

**Tabel III- 3.** Tabel Hasil Pengujian

Jumlah Data	Prediksi Benar	Prediksi Salah



### **3.3.5 Membuat Laporan Penelitian**

Setelah menyelesaikan seluruh proses pada tahapan penelitian, penelitian akan didokumentasikan berupa laporan penelitian yang disusun secara sistematis dari tahap awal hingga tahap akhir.

## **3.4 Kesimpulan**

Pada bab ini telah dibahas mengenai tahapan penelitian seperti mengumpulkan data, membangun sistem WSD menggunakan algoritma *Simplified Lesk*, melakukan pengujian, melakukan analisis hasil pengujian, dan membuat laporan penelitian.

## **BAB IV**

### **PENGEMBANGAN PERANGKAT LUNAK**

#### **4.1 Pendahuluan**

Bab ini memuat uraian terkait pengembangan perangkat lunak menggunakan metode *Rational Unified Process* (RUP) yang terdiri dari 4 fase yaitu Insepsi, Elaborasi, Konstruksi, dan Transisi.

#### **4.2 Fase Insepsi**

Hal pertama yang dilakukan pada fase ini, dimulai dengan menentukan ide untuk mendefinisikan kebutuhan sistem seperti ruang lingkup pengembangan yang dilakukan, dan melakukan analisis mengenai faktor-faktor yang berkaitan dengan pengembangan perangkat lunak.

##### **4.2.1 Pemodelan Bisnis**

Disambiguasi makna kata merupakan sebuah proses pada sistem yang diharapkan dapat memprediksi makna kata ambigu agar dapat mengetahui konteks kata pada kalimat seperti perspektif manusia untuk mengatasi permasalahan ambiguitas yang sering terjadi dalam bidang Pemrosesan Bahasa Alami.

Perangkat lunak yang dikembangkan pada penelitian ini berbasis *Website* dan menggunakan bahasa pemrograman Python. Hasil dari perangkat lunak bertujuan untuk menemukan makna yang tepat dari kata ambigu dalam kalimat Bahasa Indonesia, dan melihat kinerja algoritma yang digunakan yaitu *Simplified Lesk*.

#### 4.2.2 Kebutuhan Sistem

Kebutuhan sistem mencakup dua aspek utama, yaitu kebutuhan fungsional dan non-fungsional. Kebutuhan fungsional menggambarkan tindakan dan proses yang diharapkan dilakukan oleh sistem. Sementara itu, kebutuhan non-fungsional berfokus pada kinerja dan perilaku sistem yang harus diverifikasi. Kebutuhan fungsional dan non-fungsional pada sistem yang dibuat untuk penelitian ini dapat dilihat pada Tabel IV-1 dan Tabel IV-2.

**Tabel IV- 1.** Kebutuhan Fungsional

No	Kebutuhan Fungsional
1	Sistem dapat menerima masukan berupa kalimat dan kata ambigu
2	Sistem dapat menampilkan hasil disambiguasi makna kata menggunakan algoritma <i>Simplified Lesk</i>
3	Sistem memberikan akses ke pengguna untuk dapat memilih konfigurasi yang diinginkan terhadap evaluasi
4	Sistem dapat menampilkan hasil evaluasi <i>dataset</i> berdasarkan konfigurasi yang dipilih pengguna

**Tabel IV- 2.** Kebutuhan Non-Fungsional

No	Kebutuhan Non-Fungsional
1	Sistem memiliki antarmuka yang mudah dipahami dan digunakan oleh pengguna
2	Proses yang dibutuhkan sistem untuk menampilkan hasil disambiguasi makna kata relatif cepat, kurang dari 10 detik

### 4.2.3 Analisis dan Desain

Pada tahap ini, dilakukan serangkaian proses seperti analisis kebutuhan perangkat lunak, analisis preprocessing, analisis word sense disambiguation, analisis hasil word sense disambiguation, dan perancangan perangkat lunak.

#### 4.2.3.1 Analisis Kebutuhan Perangkat Lunak

Berdasarkan penjelasan pada pemodelan bisnis, perangkat lunak harus memenuhi dua kebutuhan yaitu melakukan disambiguasi makna kata, dan evaluasi. Kebutuhan tersebut akan dikelompokkan kedalam tiga mekanisme seperti proses *pre-processing*, proses *Word Sense Disambiguation* (WSD) dan proses pengujian WSD.

Proses *pre-processing* berfungsi untuk mempersiapkan data utuh menjadi data yang siap untuk dilakukan pemrosesan lebih lanjut. Proses ini melibatkan beberapa proses kembali seperti tokenisasi, *stopword removal*, dan *stemming*. Proses *Word Sense Disambiguation* (WSD) menggunakan algoritma Simplified Lesk yang meliputi proses pembentukan data kamus berdasarkan kata ambigu, proses perhitungan jumlah tumpang tindih (*overlap*) kata yang beririsan antara makna dengan kalimat atau konteks, untuk penentuan makna yang tepat berdasarkan jumlah *overlap* tertinggi. Proses terakhir adalah pengujian WSD, yaitu dengan melabeli data terlebih dahulu berdasarkan kesamaan antara hasil makna yang telah diprediksi oleh sistem dan makna sebenarnya, lalu terakhir akan menghitung akurasi dari hasil keseluruhan data sampel. Hasil - hasil yang didapat kemudian akan disatukan dan dijadikan sebagai hasil evaluasi kinerja dari proses WSD.

#### 4.2.3.2 Analisis Data

Data yang akan dimanfaatkan dalam pengembangan perangkat lunak ini adalah 300 data berupa Kalimat, Kata Ambigu, dan *Actual Sense* (makna sebenarnya dari kata ambigu pada kamus) yang disatukan kedalam satu *file excel*. Beberapa bagian data diperoleh dari sumber *paper* sebuah penelitian dengan topik yang sama yang berisi data Kalimat dan Kata Ambigu, dan bagian lainnya diambil secara manual dari beberapa artikel Bahasa Indonesia dan dijadikan sebagai data Kalimat. Selanjutnya dilakukan identifikasi kata dalam kalimat yang memiliki definisi lebih dari satu pada kamus, untuk dijadikan sebagai data Kata Ambigu. Data *Actual Sense* diperoleh dengan menentukan secara mandiri makna kata yang paling tepat dari beberapa makna pada kamus, yang sesuai dengan konteks kalimat berdasarkan perspektif manusia.

#### 4.2.3.3 Analisis *Pre-Processing*

*Pre-processing* bertujuan untuk menyiapkan data sehingga menjadi lebih terstruktur, dan siap untuk pemrosesan selanjutnya. Berikut merupakan 10 contoh data sampel kalimat yang diambil dari dataset sebelum dilakukan proses *pre-processing* yang dapat dilihat pada Tabel IV-3.

**Tabel IV- 3.** Data Sampel

No	Kalimat
S1	buka buku pada halaman 50
S2	Halaman yang indah menjadi idaman setiap orang yang punya rumah
S3	Saat kau ingin rumahmu terang alami, pakailah genting kaca di bagian tertentu atap rumahmu

S4	Keadaan bertambah genting mendengar kabar ani diculik
S5	Umumnya semua istri tidak akan pernah mau dimadu
S6	Madu mempunyai banyak manfaat untuk kesehatan tubuh
S7	pencuri itu kabur melarikan diri setelah dipergoki pemilik rumah
S8	mata nenek sudah kabur karena dimakan usia
S9	Sebidang tanah seluas 2 hektar tersebut milik pemerintah daerah setempat
S10	Perusahaan kami sedang mencari tenaga kerja baru di bidang pemasaran

**Keterangan : S : Sampel**

### 1. Tokenisasi

Tokenisasi merupakan proses memisahkan sepotong teks menjadi unit yang lebih kecil yang disebut token. Pada proses ini, kalimat akan dipisahkan menjadi potongan kata. Sebelum ditokenisasi, kalimat terlebih dahulu diubah menjadi huruf kecil dan dihilangkan karakter-karakter spesial yang tidak digunakan dalam perangkat lunak. Hasil dari tokenisasi kalimat dapat dilihat pada Tabel IV-4.

**Tabel IV- 4.** Hasil Tokenisasi Kalimat

No	Kalimat
S1	'buka', 'buku', 'pada', 'halaman'
S2	'halaman', 'yang', 'indah', 'menjadi', 'idaman', 'setiap', 'orang', 'punya', 'rumah'
S3	'saat', 'kau', 'ingin', 'rumahmu', 'terang', 'alami', 'pakailah', 'genting', 'kaca', 'di', 'bagian', 'tertentu', 'atap'
S4	'keadaan', 'bertambah', 'genting', 'mendengar', 'kabar', 'ani', 'diculik'
S5	'umumnya', 'semua', 'istri', 'tidak', 'akan', 'pernah', 'mau', 'dimadu'
S6	'madu', 'mempunyai', 'banyak', 'manfaat', 'untuk', 'kesehatan', 'tubuh'
S7	'pencuri', 'itu', 'kabur', 'melarikan', 'diri', 'setelah', 'dipergoki', 'pemilik', 'rumah'
S8	'mata', 'nenek', 'sudah', 'kabur', 'karena', 'dimakan', 'usia'
S9	'sebidang', 'tanah', 'seluas', 'hektar', 'tersebut', 'milik', 'pemerintah',

	'daerah', 'setempat'
S10	'perusahaan', 'kami', 'sedang', 'mencari', 'tenaga', 'kerja', 'baru', 'di', 'bidang', 'pemasaran'

### 2. *Stopwords Removal*

*Stopwords Removal* adalah langkah awal dalam tahap preprocessing yang bertujuan untuk menghapus kata-kata yang dianggap kurang penting dan tidak relevan. Hasil *stopwords removal* dapat dilihat pada Tabel IV-5.

**Tabel IV- 5.** Hasil *Stopwords Removal*

No	Kalimat
S1	'buku', 'halaman'
S2	'halaman', 'indah', 'idaman', 'rumah'
S3	'kau', 'rumahmu', 'terang', 'alami', 'pakailah', 'genting', 'kaca', 'atap'
S4	'bertambah', 'genting', 'mendengar', 'kabar', 'ani', 'diculik'
S5	'istri', 'dimadu'
S6	'madu', 'manfaat', 'kesehatan', 'tubuh'
S7	'pencuri', 'kabur', 'melarikan', 'dipergoki', 'pemilik', 'rumah'
S8	'matanya', 'kabur', 'dimakan', 'umur'
S9	'sebidang', 'tanah', 'seluas', 'hektar', 'milik', 'pemerintah', 'daerah'
S10	'perusahaan', 'mencari', 'tenaga', 'bidang', 'pemasaran'

### 3. *Stemming*

Pada tahap stemming akan dilakukan proses mengubah bentuk suatu kata menjadi bentuk dasarnya. Hasil dari tahap ini dapat dilihat pada Tabel IV-6.

**Tabel IV- 6.** Hasil *Stemming*

No	Kalimat
S1	'buku', 'halaman'

S2	'halaman', 'indah', 'idam', 'rumah'
S3	'kau', 'rumah', 'terang', 'alami', 'pakai', 'genting', 'kaca', 'atap'
S4	'tambah', 'genting', 'dengar', 'kabar', 'ani', 'culik'
S5	'istri', 'madu'
S6	'madu', 'manfaat', 'sehat', 'tubuh'
S7	'curi', 'kabur', 'lari', 'pergok', 'milik', 'rumah'
S8	'mata', 'kabur', 'makan', 'umur'
S9	'bidang', 'tanah', 'luas', 'hektar', 'milik', 'perintah', 'daerah'
S10	'usaha', 'cari', 'tenaga', 'bidang', 'pasar'

#### 4.2.3.4 Analisis Algoritma *Simplified Lesk*

##### 1. Pembentukan Makna Kamus

Proses ini merupakan proses awal dari WSD menggunakan algoritma *Simplified Lesk*, dimana data kamus yang didefinisikan terdiri dari definisi dan contoh yang diperoleh dari pencarian kata ambigu pada Kamus Besar Bahasa Indonesia (KBBI) yang dapat dilihat pada Tabel IV-7. Arti dan contoh tersebut akan digabungkan sebagai satu kesatuan yang disebut makna (*senses*), dan akan dijadikan acuan dalam menentukan makna yang tepat dari kata ambigu pada kalimat.

**Tabel IV- 7.** Hasil Pembentukan Kamus

No	Kata Ambigu	Makna	
		Definisi	Contoh
S1, S2	halaman	'pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)'	'halaman rumahnya ditanami cemara'
		'muka dari lembaran buku (majalah, surat kabar, dan sebagainya); pagina'	'lihat gambar pada halaman 127'
S3, S4	genting	'kecil (tipis, sempit) pada bagian tengah'	'pinggangnya genting'



		‘hampir putus (tentang tali dan sebagainya)’	‘tali ini genting’
		‘tegang; berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)’	‘setelah perundingan menemui jalan buntu, keadaan bertambah genting’
		‘tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya’	
S5, S6	madu	‘cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)’	‘sarang lebah ini jika diperas keluar madunya’
		‘manis sekali’	‘senyumnya manis bagai madu’
		‘istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya’	‘semua orang heran bahwa ia dapat bersikap baik terhadap madunya’
		‘orang yang menjadi saingan dalam percintaan; pesaing dalam percintaan’	
S7, S8	kabur	‘tidak dapat melihat sesuatu dengan jelas (tentang mata)’	‘matanya telah kabur dimakan umur’
		‘kurang terang (tentang pemandangan); kurang nyata (tentang lukisan); kurang jernih (tentang kaca); kurang jelas (tentang pertanyaan dan sebagainya)’	
		‘keadaan atmosfer yang tampak berkabut akibat adanya partikel yang sangat kecil dan kering yang cukup banyak terdapat di dalamnya’	
		‘berlari cepat-cepat; melarikan diri’	‘pengebut kabur dikejar Polantas’

		‘meninggalkan tugas (pekerjaan, keluarga, dan sebagainya) tanpa pamit; menghilang’	
		‘kumbang tanduk; kumbang kelapa; bangbung2’	
S9, S10	bidang	‘permukaan (yang) rata dan tentu batasnya’	‘kubus itu mempunyai enam bidang’
		‘ukuran panjang (5 hasta) untuk mengukur panjang (tikar, layar, kulit, dan sebagainya)’	‘berapa bidang tikar ini?’
		‘segi pandangan; aspek’	‘masalah itu harus ditinjau dari bidang ilmu ketatanegaraan’
		‘kata penggolong bagi barang-barang yang luas seperti tanah, sawah, ladang’	‘dua bidang tanahnya ditanami sayur-sayuran’
		‘lebar’	‘dadanya bidang’
		‘kolom yang terdapat pada kartu berlubang untuk menuliskan informasi khusus’	
		‘bagian tertentu dalam rekaman komputer, misalnya dalam rekaman bibliografi, pengarang, atau tanggal publikasi dokumen’	

## 2. Menghitung Jumlah *Overlapping*

Proses ini akan melibatkan konteks (kalimat) dan makna yang didapat dari proses sebelumnya (dan telah melewati *pre-processing*), untuk pencarian kata yang beririsan antara makna dengan kalimat. Kata-kata yang beririsan akan dihitung sebagai jumlah tumpang tindih (*overlap*) dengan mengecualikan kata ambigu. Prediksi makna (*predicted sense*) merupakan makna yang dipilih berdasarkan jumlah *overlap* tertinggi. Jika tidak terdapat kata yang beririsan (jumlah *overlap* 0), maka makna yang paling pertama yang akan dipilih sebagai

prediksi makna yang tepat oleh sistem. Selain itu, jika terdapat makna dengan kata dan jumlah irisan yang sama, maka makna pertama dari makna dengan irisan yang sama yang akan dipilih sebagai prediksi makna yang tepat oleh sistem. Tabel IV-8 akan menunjukkan hasil dari perhitungan overlap dan prediksi makna oleh sistem.

Tabel IV- 8. Hasil *Overlap* dan *Predicted Sense*

No	Kata	Konteks	Kandidat Makna		<i>Overlap</i>	<i>Predicted Sense</i>
			Indeks	Makna		
S1	halaman	'buku', ' <del>halaman</del> '	0	'tanah', 'pekarang', 'cemara', 'tanam', 'rumah', ' <del>halaman</del> ', 'sekolah'	0	[1] muka dari lembaran buku (majalah, surat kabar, dan sebagainya); pagina
			1	'pagina', ' <b>buku</b> ', 'majalah', 'kabar', 'gambar', 'surat', 'lembar', ' <del>halaman</del> ', 'muka'	1	
S2	halaman	' <del>halaman</del> ', 'indah', 'idam', 'rumah'	0	'tanah', 'pekarang', 'cemara', 'tanam', ' <b>rumah</b> ', ' <del>halaman</del> ', 'sekolah'	1	[0] pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)
			1	'pagina', 'buku', 'majalah', 'kabar', 'gambar', 'surat', 'lembar', ' <del>halaman</del> ', 'muka'	0	
S3	genting	'kau', 'rumah', 'terang', 'alami', 'pakai', ' <del>genting</del> ', 'kaca', 'atap'	0	' <del>genting</del> ', 'tipis', 'sempit', 'pinggang'	0	[3] tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya
			1	' <del>genting</del> ', 'putus', 'tali'	0	
			2	'jalan', 'runding', 'bahaya', 'tegang', 'bencana', 'buntu', 'perang', 'timbul', ' <del>genting</del> ', 'temu', 'tambah'	0	
			3	'jalan', ' <b>atap</b> ', 'cetak', 'buat', 'tanah', 'liat', 'bentuk', 'runding', 'tutup', 'buntu', ' <b>rumah</b> ', ' <del>genting</del> ', 'bakar', 'temu', 'bermacam-macam', 'tambah'	2	
S4	genting	'tambah', ' <del>genting</del> ',	0	' <del>genting</del> ', 'tipis', 'sempit', 'pinggang'	0	[2] tegang;

		'dengar', 'kabar', 'ani', 'culik'	1	' <del>genting</del> ', 'putus', 'tali'	0	berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)
			2	'jalan', 'runding', 'bahaya', 'tegang', 'bencana', 'buntu', 'perang', 'timbul', ' <del>genting</del> ', 'temu', ' <b>tambah</b> '	1	
			3	'jalan', 'atap', 'cetak', 'buat', 'tanah', 'liat', 'bentuk', 'runding', 'tutup', 'buntu', 'rumah', ' <del>genting</del> ', 'bakar', 'temu', 'bermacammacam', ' <b>tambah</b> '	1	
S5	madu	'istri', ' <del>madu</del> '	0	'cair', 'gula', 'bunga', 'sarang', 'kandung', ' <del>madu</del> ', 'lebah', 'manis', 'zat', 'peras'	0	[2] istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya
			1	' <del>madu</del> ', 'senyum', 'manis'	0	
			2	'dasar', 'suami', 'sikap', ' <b>istri</b> ', ' <del>madu</del> ', 'pertama', 'pandang', 'heran'	1	
			3	' <del>madu</del> ', 'cinta', 'sikap', 'saing', 'heran'	0	
S6	madu	' <del>madu</del> ', 'manfaat', 'sehat', 'tubuh'	0	'cair', 'gula', 'bunga', 'sarang', 'kandung', ' <del>madu</del> ', 'lebah', 'manis', 'zat', 'peras'	0	[0] cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)
			1	' <del>madu</del> ', 'senyum', 'manis'	0	
			2	'dasar', 'suami', 'sikap', 'istri', ' <del>madu</del> ', 'pertama', 'pandang', 'heran'	0	
			3	' <del>madu</del> ', 'cinta', 'sikap', 'saing', 'heran'	0	
S7	kabur	'curi', ' <del>kabur</del> ', 'lari', 'pergok', 'milik', 'rumah'	0	'makan', ' <del>kabur</del> ', 'mata', 'umur'	0	[3] berlari cepat-cepat; melarikan diri
			1	' <del>kabur</del> ', 'makan', 'lukis', 'terang', 'umur', 'jernih', 'mata', 'pandang', 'kaca'	0	
			2	'kering', ' <del>kabur</del> ', 'kabut', 'atmosfer', 'akibat',	0	

				'dalam', 'umur', 'mata', 'partikel', 'makan'		
			3	'kebut', <del>'kabur'</del> , 'cepatcepat', 'polantas', 'kejar', 'lari'	1	
			4	'kebut', <del>'kabur'</del> , 'tinggal', 'kejar', 'hilang', 'polantas', 'keluarga', 'pamit', 'kerja', 'tugas'	0	
			5	'kebut', <del>'kabur'</del> , 'bangbung', 'kejar', 'kelapa', 'tanduk', 'polantas', 'kumbang'	0	
S8	kabur	'nenek', <del>'kabur'</del> , 'makan', 'usia'	0	'makan', <del>'kabur'</del> , 'mata', 'umur'	1	[0] tidak dapat melihat sesuatu dengan jelas (tentang mata)
			1	<del>'kabur'</del> , 'makan', 'lukis', 'terang', 'umur', 'jernih', 'mata', 'pandang', 'kaca'	1	
			2	'kering', <del>'kabur'</del> , 'kabut', 'atmosfer', 'akibat', 'dalam', 'umur', 'mata', 'partikel', 'makan'	1	
			3	'kebut', <del>'kabur'</del> , 'cepatcepat', 'polantas', 'kejar', 'lari'	0	
			4	'kebut', <del>'kabur'</del> , 'tinggal', 'kejar', 'hilang', 'polantas', 'keluarga', 'pamit', 'kerja', 'tugas'	0	
			5	'kebut', <del>'kabur'</del> , 'bangbung', 'kejar', 'kelapa', 'tanduk', 'polantas', 'kumbang'	0	
S9	bidang	<del>'bidang'</del> , 'tanah', 'luas', 'hektar', 'milik', 'perintah', 'daerah'	0	'kubus', <del>'bidang'</del> , 'batas', 'muka'	0	kata penggolong bagi barang-barang yang luas seperti tanah, sawah, ladang
			1	'hasta', 'layar', 'ukur', 'tikar', 'bidang', 'kulit'	0	
			2	'lingkung', 'buruh', 'tahu', 'kerja', <del>'bidang'</del> , 'lapang'	0	
			3	'ilmu', 'segi', 'aspek', 'ketatanegaraan', 'pandang', <del>'bidang'</del> , 'tinjau'	0	
			4	'tanah', 'sawah', 'sayursayuran', <del>'bidang'</del> , 'luas',	2	

				'ladang', 'tanam', 'barangbarang', 'golong'		
			5	'lebar', 'bidang', 'dada'	0	
			6	'lubang', 'kartu', 'tulisi', 'bidang', 'dada', 'kolom', 'informasi'	0	
			7	'komputer', 'bibliografi', 'tanggal', 'bidang', 'arang', 'dada', 'publikasi', 'dokumen', 'rekam'	0	
S10	bidang	'usaha', 'cari', 'tenaga', 'bidang', 'pasar'	0	'kubus', 'bidang', 'batas', 'muka'	0	permukaan (yang) rata dan tentu batasnya
			1	'hasta', 'layar', 'ukur', 'tikar', 'bidang', 'kulit'	0	
			2	'lingkung', 'buruh', 'tahu', 'kerja', 'bidang', 'lapang'	0	
			3	'ilmu', 'segi', 'aspek', 'ketatanegaraan', 'pandang', 'bidang', 'tinjau'	0	
			4	'tanah', 'sawah', 'sayursayuran', 'bidang', 'luas', 'ladang', 'tanam', 'barangbarang', 'golong'	0	
			5	'lebar', 'bidang', 'dada'	0	
			6	'lubang', 'kartu', 'tulisi', 'bidang', 'dada', 'kolom', 'informasi'	0	
			7	'komputer', 'bibliografi', 'tanggal', 'bidang', 'arang', 'dada', 'publikasi', 'dokumen', 'rekam'	0	

#### 4.2.3.5 Analisis Pengujian *Word Sense Disambiguation* (WSD)

Sistem WSD menggunakan algoritma *Simplified Lesk* yang dibuat akan dievaluasi menggunakan pengujian akurasi, yaitu menghitung data dengan prediksi makna yang benar dari jumlah keseluruhan data. Data tersebut telah dilabeli terlebih dahulu dalam format *binary* yaitu 1 dan 0 yang merepresentasikan benar dan salah untuk melihat kesesuaian antara hasil prediksi makna oleh sistem dengan makna sebenarnya. Evaluasi yang ditambahkan yaitu dengan menampilkan tabel yang berisi setiap detail data beserta label data yang telah didefinisikan menjadi ‘Sesuai’(1) dan ‘Tidak Sesuai’(0). Hasil dari pengujian WSD menggunakan *Simplified Lesk* serta analisis kinerjanya akan dijelaskan lebih lanjut pada Bab V.

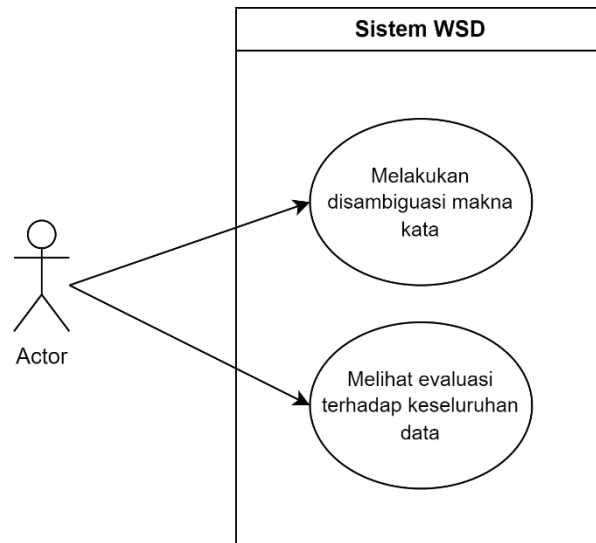
#### 4.2.3.6 Desain Perangkat Lunak

Desain perangkat lunak yang dibuat akan didefinisikan dengan diagram *use case*, tabel definisi *use case*, dan skenario *use case*.

1. Diagram *Use Case*

Interaksi antara *actor* dan perangkat lunak yang dibuat akan didefinisikan menggunakan diagram *use case* seperti pada Gambar IV-1.





**Gambar IV- 1.** *Use Case Diagram*

## 2. Tabel Definisi *Actor*

Seorang aktor (*Actor*) adalah individu yang akan terlibat secara langsung dengan perangkat lunak WSD menggunakan algoritma *Simplified Lesk* ini. Definisi actor pada perangkat lunak yang dibuat akan didefinisikan pada Tabel IV-10.

**Tabel IV- 9.** Definisi *Actor*

No	<i>Actor</i>	Definisi
1	<i>User</i>	Seseorang yang dapat berinteraksi secara langsung dengan perangkat lunak dan menyusuri semua fitur yang tersedia

## 3. Tabel Definisi *Use Case*

Definisi dari *use case* yang telah dimodelkan sebelumnya dapat dilihat pada Tabel IV-11.

**Tabel IV- 10.** Definisi *Use Case*

No	<i>Use Case</i>	Definisi
1	Melakukan disambiguasi makna kata	Melakukan proses disambiguasi untuk memprediksi makna kata yang tepat berdasarkan masukan Kalimat, dan Kata Ambigu
2	Melihat evaluasi terhadap keseluruhan data	Melihat hasil evaluasi yang ditampilkan dari <i>dataset</i> berdasarkan konfigurasi yang dipilih

#### 4. Skenario *Use Case*

Setelah pemodelan dan definisi *use case*, selanjutnya skenario *use case* akan memuat secara detail setiap perilaku maupun hubungan antara *actor* dan perangkat lunak. Tabel IV-12 merupakan scenario *use case* dari perangkat lunak yang dibuat.

**Tabel IV- 11.** Skenario *Use Case* Disambiguasi Makna Kata

<b>Identifikasi</b>	
<b>Nomor</b>	01
<b>Nama</b>	Melakukan disambiguasi makna kata
<b>Tujuan</b>	Menampilkan makna dari kata ambigu pada kalimat
<b>Deskripsi</b>	Melakukan proses disambiguasi makna kata berdasarkan masukan kalimat dan kata ambigu

<b>Aktor</b>	Pengguna
<b>Kondisi Awal</b>	Sistem menampilkan halaman awal berupa form masukan dan hasil yang akan ditampilkan
<b>Skenario Utama</b>	
<b>Aksi Aktor</b>	<b>Reaksi Sistem</b>
1. Memasukkan kalimat pada <i>field</i> Kalimat 1 dan Kalimat 2 2. Memasukkan kata ambigu pada <i>field</i> Kata Ambigu 3. Menekan tombol 'Disambiguasi'	
	4. Menampilkan hasil disambiguasi makna kata dari kata ambigu pada kalimat yang telah diinputkan
<b>Kondisi Akhir</b>	Sistem menampilkan hasil prediksi makna kata yang paling beririsan dengan kalimat

**Tabel IV- 12.** Skenario *Use Case* Menampilkan Evaluasi

<b>Identifikasi</b>	
<b>Nomor</b>	02
<b>Nama</b>	Melihat evaluasi terhadap keseluruhan data
<b>Tujuan</b>	Melihat dan mengetahui hasil evaluasi disambiguasi makna kata menggunakan algoritma <i>Simplified Lesk</i>
<b>Deskripsi</b>	Menampilkan data jumlah prediksi benar-salah, akurasi, dan detail data disambiguasi berdasarkan konfigurasi yang dipilih

<b>Aktor</b>	Pengguna
<b>Kondisi Awal</b>	Tampilan halaman awal (demo)
<b>Skenario Utama</b>	
<b>Aksi Aktor</b>	<b>Reaksi Sistem</b>
1. Menekan menu 'Evaluasi'	
	2. Menampilkan halaman evaluasi dengan tampilan awal berupa kolom untuk memilih jumlah data
3. Memilih opsi jumlah data dari evaluasi dataset yang ingin ditampilkan 4. Menekan tombol 'Proses'	
	5. Menampilkan data hasil evaluasi berupa data statistik prediksi benar-salah dan akurasi, serta hasil disambiguasi setiap data yang disajikan dalam bentuk tabel dari dataset berdasarkan jumlah data yang dipilih
<b>Kondisi Akhir</b>	Sistem menampilkan hasil evaluasi WSD dari dataset berdasarkan jumlah data yang dipilih

### 4.3 Fase Elaborasi

Tahap ini akan mengidentifikasi kebutuhan seperti merancang model bisnis, data, dan antarmuka (*interface*), serta membuat diagram *activity* dan diagram *sequence*.

### **4.3.1 Pemodelan Bisnis**

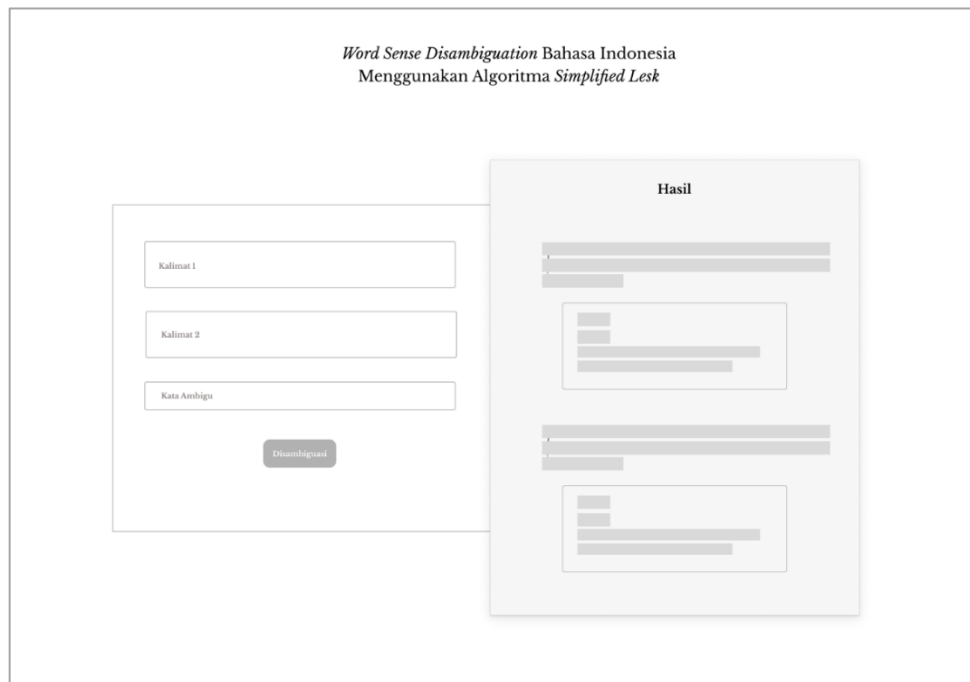
Tahap ini akan menerangkan tentang bagaimana merancang perangkat lunak, termasuk perancangan data dan tampilan antarmuka.

#### **4.3.1.1 Perancangan Data**

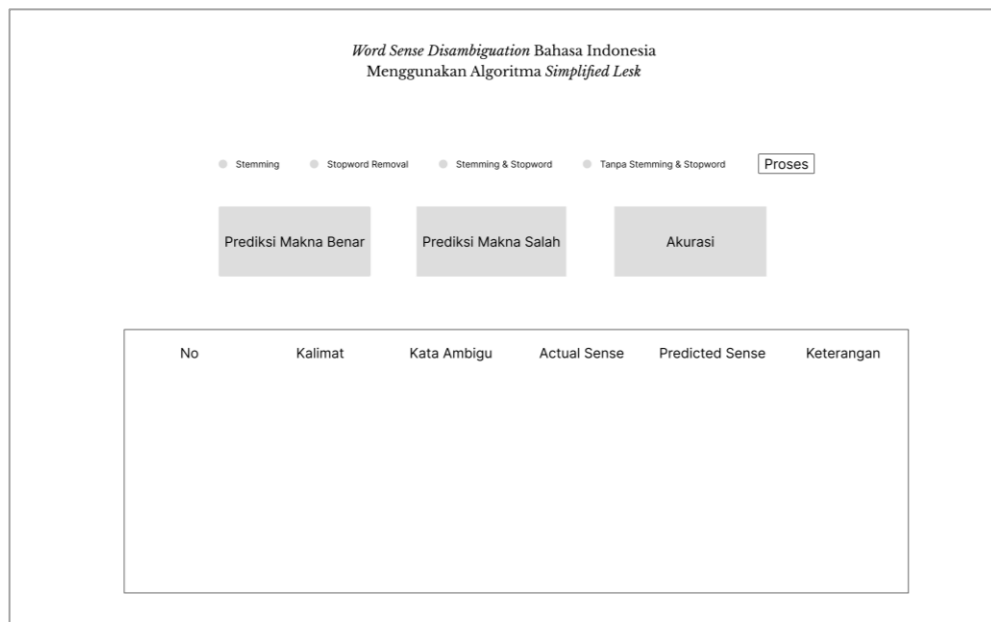
Data yang dipergunakan yaitu 300 data berupa kalimat, kata ambigu, dan *actual sense* yang disatukan kedalam *file excel*. Selain itu, terdapat data *stopwords* yang dikumpulkan kedalam 1 *file* berekstensi *.txt* dan berperan dalam proses WSD menggunakan *Simplified Lesk*.

#### **4.3.1.2 Perancangan Antarmuka**

Antarmuka merupakan fitur-fitur yang memfasilitasi interaksi pengguna dengan perangkat lunak atau aplikasi. Perancangan antarmuka bertujuan untuk memenuhi kebutuhan pengguna dan mendukung fungsionalitas perangkat lunak yang efektif. Rancangan antarmuka yang dibangun dapat dilihat pada Gambar IV-2 dan Gambar IV-3.



**Gambar IV- 2.** Rancangan Antarmuka Disambiguasi Makna Kata



**Gambar IV- 3.** Rancangan Antarmuka Menampilkan Evaluasi

### **4.3.2 Kebutuhan Sistem**

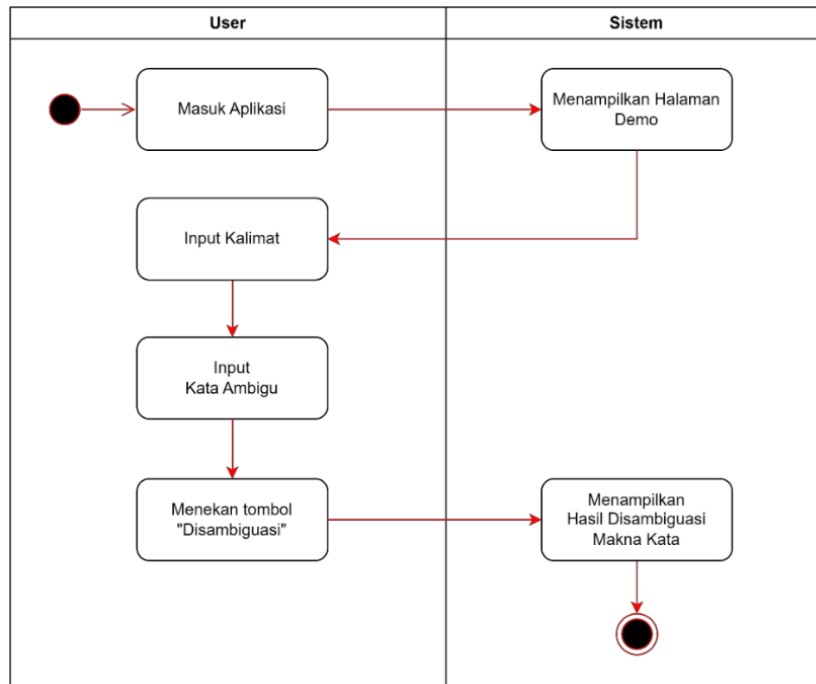
Pengembangan sistem atau perangkat lunak secara umum diperlukan tiga elemen, yaitu perangkat keras, perangkat lunak, serta sebuah bahasa pemrograman, dengan Python sebagai bahasa pemrograman yang digunakan.

### **4.3.3 Analisis dan Perancangan**

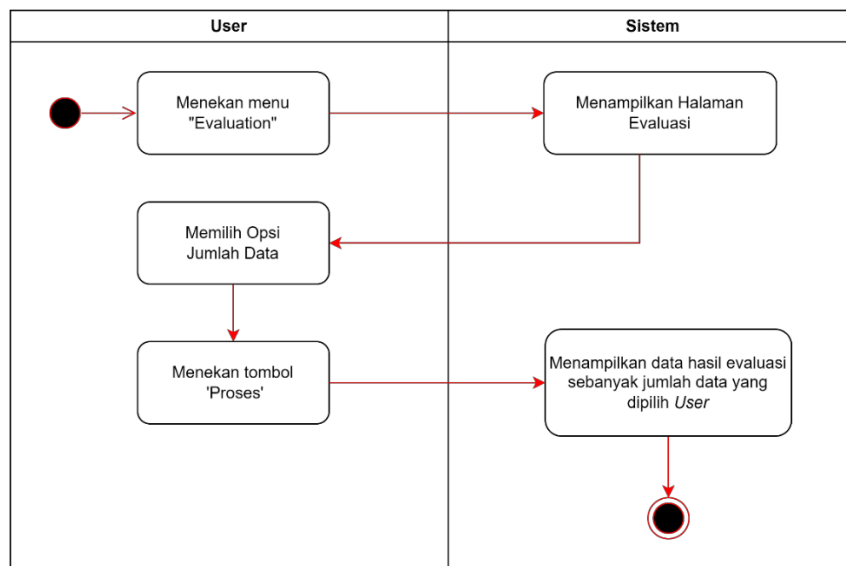
Bagian ini akan memuat penjelasan tentang perancangan yang dilakukan dalam mengembangkan perangkat lunak, yaitu Diagram *Activity* dan Diagram *Sequence*.

#### **4.3.3.1 Diagram *Activity***

Diagram activity menggambarkan alur dan perpindahan setiap kegiatan pada perangkat lunak. Diagram activity pada penelitian ini terdiri dari 2 aktivitas yaitu Disambiguasi Makna Kata dan Menampilkan Evaluasi yang dapat dilihat pada Gambar IV-4 dan Gambar IV-5.



**Gambar IV- 4.** Diagram *Activity* Disambiguasi Makna Kata



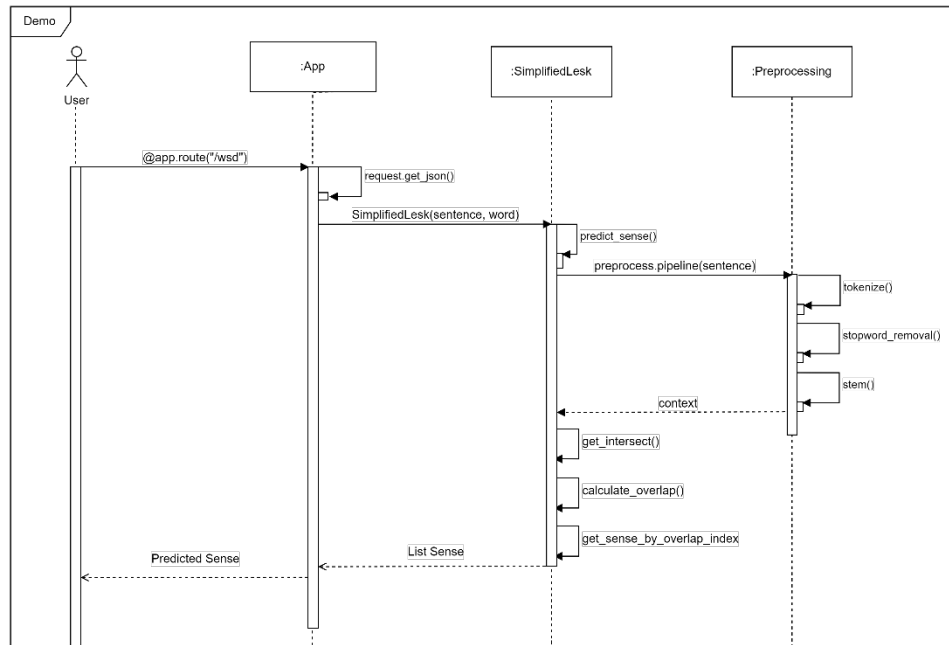
**Gambar IV- 5.** Diagram *Activity* Menampilkan Evaluasi

#### 4.3.3.2 Diagram *Sequence*

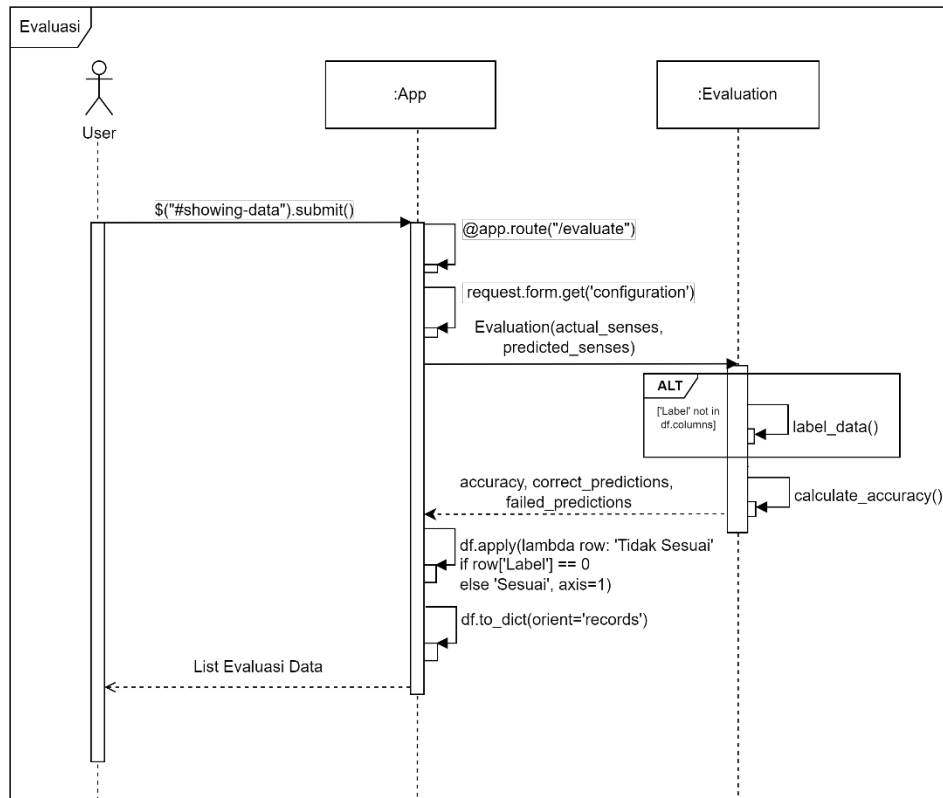
Diagram ini adalah diagram interaksi yang mendefinisikan setiap proses dan obyek yang terlibat yang diperlukan untuk melaksanakan fungsionalitas.



Gambar IV-6 dan Gambar IV-7 merupakan diagram *sequence* Disambiguasi Makna Kata dan Menampilkan Evaluasi yang didefinisikan berdasarkan *use case*.



**Gambar IV- 6.** Diagram *Sequence* Disambiguasi Makna Kata



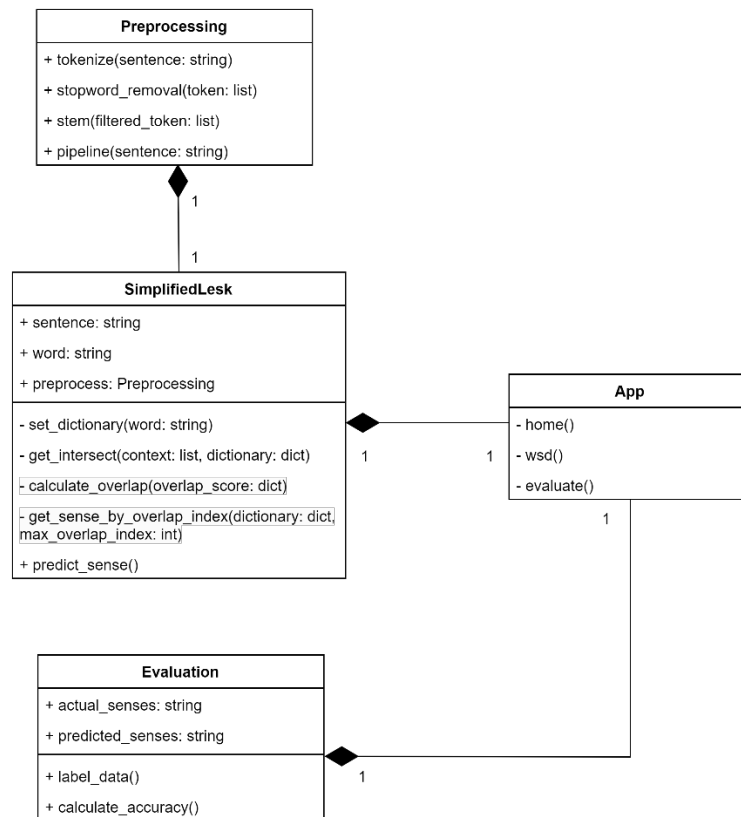
**Gambar IV- 7.** Diagram *Sequence* Menampilkan Evaluasi

#### 4.4 Fase Konstruksi

Fase Konstruksi merupakan fase ketiga pengembangan yang berfokus pada implementasi dan menghasilkan sistem sebagai pendukung penelitian ini.

##### 4.4.1 Kebutuhan Sistem

Topik pada bagian ini yaitu mengilustrasikan pemodelan perangkat lunak dalam bentuk diagram kelas yang akan memuat penjelasan terkait struktur sistem dengan menunjukkan kelas, atribut, dan relasi antar objek. Diagram kelas pada penelitian ini dapat dilihat pada Gambar IV-8.



**Gambar IV- 8.** Diagram Kelas

#### 4.4.2 Implementasi

Tahap ini akan mengimplementasikan rancangan yang dibuat menjadi sebuah perangkat lunak. Perangkat lunak juga didukung dengan Bahasa pemrograman Python dan *library* Flask. Berikut ini akan menjelaskan mengenai implementasi kelas dan *interface* dari diagram kelas dan rancangan antarmuka yang telah dibuat.

#### 4.4.2.1 Implementasi Kelas

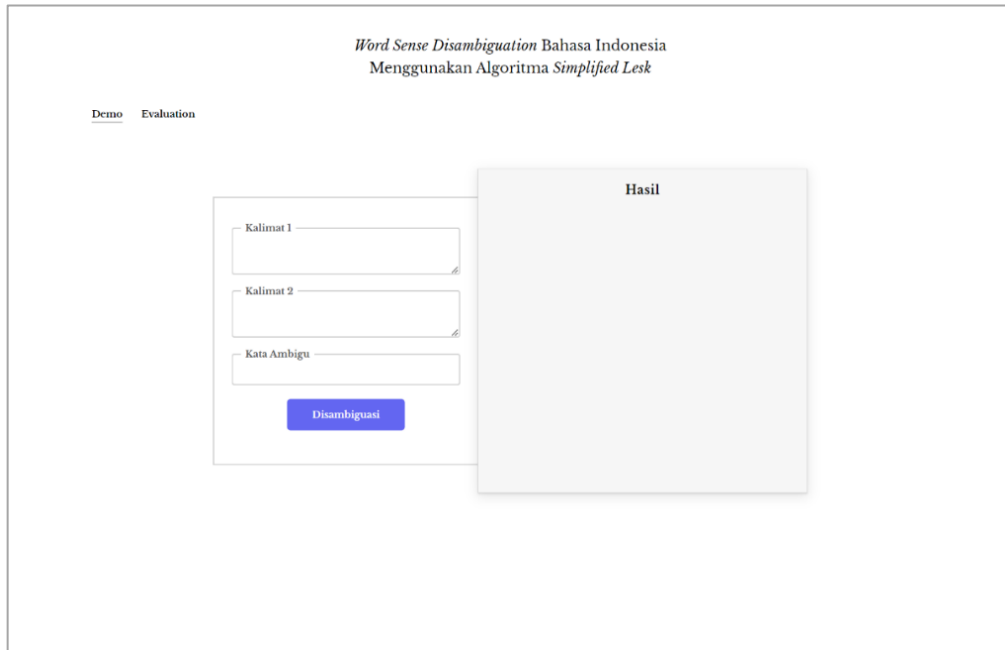
Rancangan kelas yang dibuat sebelumnya akan diterapkan menggunakan Bahasa pemrograman Python. Hasil implementasi kelas dapat dilihat pada Tabel IV-13

**Tabel IV- 13.** Implementasi Kelas

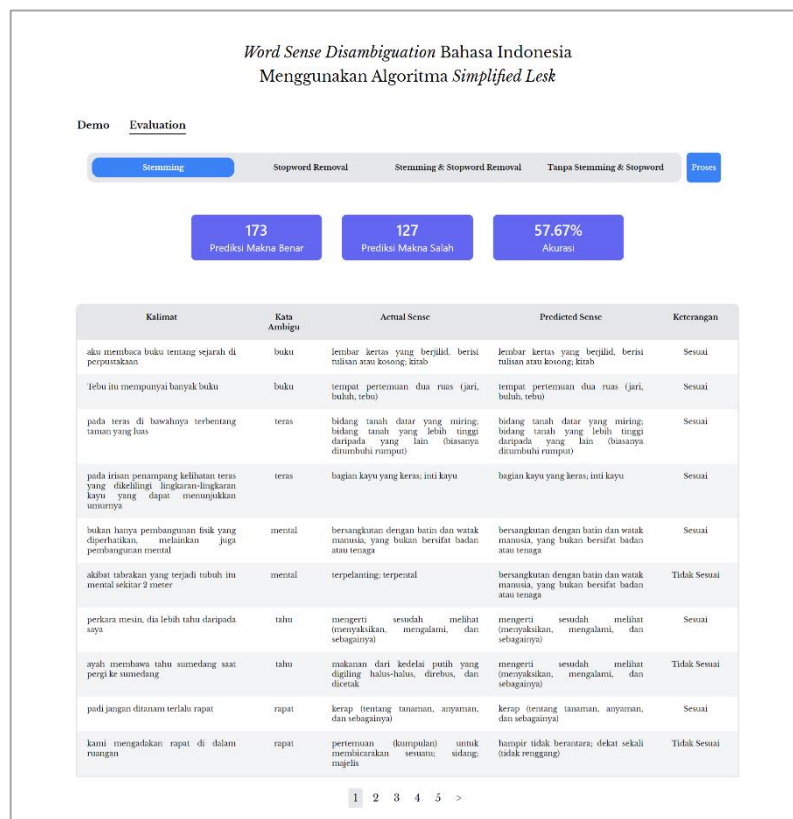
No	Nama Kelas	Nama File	Keterangan
1	Preprocessing	WSD.py	Kelas ini melakukan proses pra-pengolahan untuk menyiapkan data sebelum diproses lebih lanjut
2	SimplifiedLesk	WSD.py	Kelas ini akan melakukan seluruh proses utama sistem WSD dengan algoritma <i>Simplified Lesk</i>
3	Evaluation	evaluation.py	Kelas ini melakukan evaluasi dataset yang digunakan dari hasil pemrosesan algoritma WSD
4	App	app.py	Kelas ini menghubungkan antara tampilan dengan proses sistem

#### 4.4.2.2 Implementasi *Interface*

Gambar IV-9 dan Gambar IV-10 merupakan hasil implementasi dari rancangan antarmuka yang telah dibuat.



Gambar IV- 9. Implementasi Antarmuka Disambiguasi Makna Kata



Gambar IV- 10. Implementasi Antarmuka Menampilkan Evaluasi

## 4.5 Fase Transisi

Fase ini merupakan fase terakhir pengembangan yang berfokus pada pengujian perangkat lunak yang telah dibuat.

### 4.5.1 Pemodelan Bisnis

Perangkat lunak yang telah dikembangkan pada penelitian ini akan diuji menggunakan metode *Black Box Testing*, dengan menyusun rencana atau strategi pengujian sebagai langkah awal dari pengujian.

### 4.5.2 Rencana Pengujian

Rencana atau strategi pengujian pada perangkat lunak yang telah dikembangkan pada penelitian ini mengacu pada *use case* yang dibuat sebelumnya. Tabel IV-14 merupakan rencana pengujian *use case* Disambiguasi Makna Kata, dan Tabel IV-15 merupakan rencana pengujian Menampilkan Evaluasi

**Tabel IV- 14.** Rencana Pengujian Disambiguasi Makna Kata

No	ID	Pengujian	Tingkat Pengujian
1	UC-1-1	Disambiguasi Makna Kata	<i>Unit Testing</i>

**Tabel IV- 15.** Rencana Pengujian Menampilkan Evaluasi

No	ID	Pengujian	Tingkat Pengujian
1	UC-2-1	Menampilkan Evaluasi	<i>Unit Testing</i>

### 4.5.3 Implementasi

Bagian ini menjelaskan hasil implementasi dari rencana pengujian perangkat lunak yang telah dibuat. Tabel IV-16 dan Tabel IV-17 merupakan hasil pengujian Disambiguasi Makna Kata dan pengujian Menampilkan Evaluasi.

**Tabel IV- 16.** Hasil Pengujian Disambiguasi Makna Kata

<b>ID</b>	<b>Prosedur Pengujian</b>	<b>Input</b>	<b>Output yang diharapkan</b>	<b>Output yang didapatkan</b>	<b>Kesimpulan</b>
UC-1-1	Menekan tombol ‘Disambiguasi’	Kalimat, Kata Ambigu	Menampilkan hasil evaluasi WSD dari dataset berdasarkan jumlah data yang dipilih	Menampilkan hasil evaluasi WSD dari dataset berdasarkan jumlah data yang dipilih	Terpenuhi

**Tabel IV- 17.** Hasil Pengujian Menampilkan Evaluasi

<b>ID</b>	<b>Prosedur Pengujian</b>	<b>Input</b>	<b>Output yang diharapkan</b>	<b>Output yang didapatkan</b>	<b>Kesimpulan</b>
UC-2-1	Menekan tombol ‘Proses’	Jumlah data	Menampilkan hasil evaluasi WSD dari dataset berdasarkan	Menampilkan hasil evaluasi WSD dari dataset berdasarkan	Terpenuhi

			konfigurasi yang dipilih	konfigurasi yang dipilih	
--	--	--	-----------------------------	-----------------------------	--

#### 4.6 Kesimpulan

Bab ini telah menjelaskan keseluruhan proses pengembangan sistem atau perangkat lunak dengan menerapkan RUP sebagai metode pendukung. Seluruh fase telah dipaparkan dengan detail sehingga menghasilkan perangkat lunak yang sesuai dengan kebutuhan penelitian.



## **BAB V**

### **HASIL DAN ANALISIS**

#### **5.1 Pendahuluan**

Bagian ini akan menampilkan evaluasi pengujian *Word Sense Disambiguation* (WSD) menggunakan algoritma *Simplified Lesk*, serta hasil analisis berdasarkan evaluasi.

#### **5.2 Hasil Penelitian**

##### **5.2.1 Konfigurasi Percobaan**

Pengujian akan melibatkan 300 data berupa Kalimat Bahasa Indonesia, Kata ambigu, dan *Actual Sense*. Kalimat dan Kata Ambigu digunakan untuk melakukan proses disambiguasi makna untuk memperoleh makna yang tepat (*Predicted Sense*) sesuai dengan konteks kalimat. *Actual Sense* digunakan untuk mengukur kinerja WSD menggunakan algoritma *Simplified Lesk* pada kalimat Bahasa Indonesia.

Pengujian akan menggunakan 4 skenario berbeda untuk menentukan proses yang harus dilakukan dalam menghasilkan konteks untuk menentukan makna yang sesuai. Pengujian akan mencoba proses WSD dengan pra-pengolahan *stemming*, proses WSD dengan pra-pengolahan *stopword removal*, proses WSD dengan pra-pengolahan *stemming* dan *stopword removal*, dan proses WSD tanpa pra-pengolahan *stemming* dan *stopword removal*. Selanjutnya dari hasil konfigurasi percobaan, dilakukan analisis untuk mendapatkan nilai akurasi yang

diperoleh dengan menghitung jumlah data prediksi benar dari percobaan yang diuji.

### **5.2.2 Hasil Konfigurasi 1**

Hasil dari percobaan menggunakan pra-pengolahan stemming dalam menghasilkan konteks dan proses dalam menentukan makna yang tepat akan ditunjukkan dengan menggunakan 2 sampel data dari *dataset* secara rinci pada Tabel V-1. Selanjutnya hasil prediksi makna akan dilabeli dengan keterangan ‘Sesuai’ atau ‘Tidak Sesuai’ berdasarkan kesesuaian makna yang dihasilkan sistem dengan makna sebenarnya, yang dapat dilihat pada Tabel V-2.

Tabel V- 1. Sampel Hasil Percobaan *Stemming*

No	Kalimat	Kata	Konteks	Kandidat Makna		Overlap	Predicted Sense
				Indeks	Makna		
1	Surat tersebut dikirim melalui kantor pos	pos	['surat', 'sebut', 'kirim', 'lalu', 'kantor', 'pos']	0	{'standar', 'prosedur', 'operasional'}	0	[1] jawatan yang menyelenggarakan kirim-mengirim barang, surat, uang, dan sebagainya
				1	{'pos', 'barang', 'dan', 'jawat', 'uang', 'kirim', 'bagai', 'lalu', 'surat', 'selenggara', 'pegawai', 'yang'}	3	
				2	{'pos', 'ia', 'dan', 'beli', 'prangko', 'uang', 'untuk', 'kirim', 'bagai', 'tempat', 'surat', 'pergi', 'kantor', 'ke'}	3	
				3	{'antar', 'pos', 'tukang', 'dan', 'antara', 'dengan', 'kirim', 'bagai', 'surat', 'kartu', 'warkat', 'pegawai', 'yang'}	2	
				4	{'pos', 'angkut', 'semarang', 'tumpang', 'dan', 'kereta', 'bagai', 'pada', 'surat', 'dahulu', 'dari', 'kami', 'zaman', 'kuda'}	1	
				5	{'segar', 'dahulu', 'kuda', 'tiba', 'di', 'ganti', 'kusir', 'pos',	0	

				'zaman', 'dengan', 'yang', 'masih', 'tempat', 'kereta', 'henti', 'pada'}		
			6	{'rumah', 'mereka', 'di', 'pos', 'belakang', 'lama', 'dan', 'kereta', 'henti', 'bagai', 'trem', 'api'}	0	
			7	{'tulis', 'sosial', 'lalu', 'rumah', 'sebar', 'mereka', 'atau', 'di', 'pos', 'yang', 'belakang', 'lama', 'unggah', 'media'}	1	
			8	{'polisi', 'amat', 'di', 'rap', 'tentara', 'dua', 'tempat', 'jaga', 'dan', 'gunung', 'regu', 'pos', 'bagai', 'depan'}	0	
			9	{'tugas', 'selalu', 'harus', 'di', 'laku', 'yang', 'siap', 'masing', 'dokter', 'tempat', 'orang', 'dan', 'duduk', 'pos', 'bagai'}	0	
			10	{'kumpul', 'selalu', 'harus', 'di', 'siap', 'dokter', 'tempat', 'anggota', 'kelompok', 'pos', 'masing'}	0	
			11	{'jawab', 'selalu', 'untuk',	0	

					'harus', 'jumlah', 'di', 'perlu', 'yang', 'siap', 'masing', 'dokter', 'dan', 'bagai', 'pos', 'suatu', 'uang', 'sedia'}		
				12	{'pos', 'dokter', 'di', 'selalu', 'dari', 'harus', 'siap', 'belanja', 'bagi', 'anggar', 'mata', 'masing'}	0	
2	Pos tentara militer tepat berada di depan	pos	['pos', 'tentara', 'militer', 'tepat', 'ada', 'di', 'depan']	0	{'standar', 'prosedur', 'operasional'}	0	[8] tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)
				1	{'pos', 'barang', 'dan', 'jawab', 'uang', ' kirim', 'bagai', 'lalu', 'surat', 'selenggara', 'pegawai', 'yang'}	0	
				2	{'pos', 'ia', 'dan', 'beli', 'prangko', 'uang', 'untuk', ' kirim', 'bagai', 'tempat', 'surat', 'pergi', 'kantor', 'ke'}	0	
				3	{'antar', 'pos', 'tukang', 'dan', 'antara', 'dengan', ' kirim', 'bagai', 'surat', 'kartu', 'warkat', 'pegawai', 'yang'}	0	
				4	{'pos', 'angkut', 'semarang', 'tumpang', 'dan', 'kereta',	0	

				'bagai', 'pada', 'surat', 'dahulu', 'dari', 'kami', 'zaman', 'kuda'}	
			5	{'pos', 'masih', 'di', 'dengan', 'kereta', 'kusir', 'kuda', 'segar', 'pada', 'tempat', 'dahulu', 'henti', 'tiba', 'ganti', 'zaman', 'yang'}	1
			6	{'pos', 'di', 'dan', 'belakang', 'trem', 'kereta', 'bagai', 'mereka', 'rumah', 'lama', 'henti', 'api'}	1
			7	{'pos', 'media', 'di', 'belakang', 'mereka', 'atau', 'lalu', 'unggah', 'sebar', 'lama', 'sosial', 'rumah', 'tulis', 'yang'}	1
			8	{'depan', 'pos', 'rap', 'di', 'dan', 'dua', 'polisi', 'jaga', 'tentara', 'regu', 'bagai', 'gunung', 'amat', 'tempat'}	3
			9	{'pos', 'dokter', 'di', 'dan', 'duduk', 'orang', 'tugas', 'bagai', 'selalu', 'tempat', 'harus', 'siap', 'laku', 'masing', 'yang'}	1
			10	{'pos', 'dokter', 'di', 'kelompok',	1

					'anggota', 'selalu', 'kumpul', 'tempat', 'harus', 'siap', 'masing'}		
				11	{'pos', 'jumlah', 'dokter', 'di', 'dan', 'jawab', 'uang', 'perlu', 'untuk', 'bagai', 'selalu', 'harus', 'siap', 'yang', 'suatu', 'masing', 'sedia'}	1	
				12	{'pos', 'dokter', 'di', 'selalu', 'dari', 'harus', 'siap', 'belanja', 'bagi', 'anggar', 'mata', 'masing'}	1	

**Tabel V- 2.** Sampel Hasil Analisis Proses WSD dengan *Stemming*

No	Kata	Kalimat	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan
1	pos	Surat tersebut dikirim melalui kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	jawatan yang menyelenggarakan kirim-mengirim barang, surat, uang, dan sebagainya	Tidak Sesuai

2	pos	Pos tentara militer tepat berada di depan	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	Sesuai
---	-----	---	--	--	--------



Tabel V-3 akan menampilkan hasil analisis jumlah data, prediksi benar, dan prediksi salah yang diperoleh dari hasil percobaan menggunakan pra-pengolahan *stemming* berdasarkan kedua tabel diatas.

**Tabel V- 3.** Hasil Evaluasi Pengujian dengan *Stemming*

Jumlah Data	Prediksi Benar	Prediksi Salah
300	173	127

Berdasarkan hasil analisis diatas, akan dilakukan perhitungan akurasi berikut sebagai evaluasi dari pengujian WSD menggunakan proses pra-pengolahan *stemming*.

$$\text{Akurasi} = \frac{173}{300} \times 100\% = 57.67\%$$

### 5.2.3 Hasil Konfigurasi 2

Hasil dari percobaan menggunakan pra-pengolahan *stopword removal* dalam menghasilkan konteks dan proses dalam menentukan makna yang tepat akan ditunjukkan dengan menggunakan 2 sampel data dari *dataset* secara rinci pada Tabel V-4. Selanjutnya hasil prediksi makna akan dilabeli dengan keterangan ‘Sesuai’ atau ‘Tidak Sesuai’ berdasarkan kesesuaian makna yang dihasilkan sistem dengan makna sebenarnya, yang dapat dilihat pada Tabel V-5.

Tabel V- 4. Sampel Hasil Percobaan *Stopword Removal*

No	Kalimat	Kata	Konteks	Kandidat Makna		Overlap	Predicted Sense
				Indeks	Makna		
1	Surat tersebut dikirim melalui kantor pos	pos	[ 'surat', 'dikirim', 'kantor', 'pos' ]	0	{ 'standar', 'prosedur', 'operasional' }	0	
				1	{ 'mengirim', 'dikirim', 'jawatan', 'pegawai', 'menyelenggarakan', 'pos', 'surat', 'uang', 'barang', 'kirim' }	2	
				2	{ 'prangko', 'kirim', 'mengirim', 'pergi', 'membeli', 'pos', 'kantor', 'uang', 'surat' }	2	
				3	{ 'mengantarkan', 'kartu', 'warkat', 'tukang', 'pegawai', 'pos', 'pengantar', 'perantaraan', 'dikirim', 'surat' }	2	
				4	{ 'menumpang', 'pengangkut', 'semarang', 'pos', 'kuda', 'kereta', 'zaman', 'surat' }	1	
				5	{ 'kudanya', 'mengganti', 'perhentian', 'pos', 'segar', 'kereta', 'kusir', 'zaman' }	0	
				6	{ 'perhentian', 'pos', 'rumah',	0	

					'kereta', 'trem'}		
				7	{'sosial', 'unggahan', 'media', 'tulisan', 'pos', 'rumah', 'disebarkan'}	0	
				8	{'berapi', 'tentara', 'penjagaan', 'terdepan', 'regu', 'ditempatkan', 'pos', 'pengamat', 'gunung', 'polisi'}	0	
				9	{'dokter', 'tugas', 'pos', 'kedudukan'}	0	
				10	{'pos', 'dokter', 'berkumpul', 'anggota', 'sekelompok'}	0	
				11	{'disediakan', 'pos', 'dokter', 'uang', 'keperluan', 'jawatan'}	0	
				12	{'belanja', 'pos', 'anggaran', 'dokter'}	0	
2	Pos tentara militer tepat berada di depan	pos	['pos', 'tentara', 'militer']	0	{'standar', 'prosedur', 'operasional'}	0	
				1	{'mengirim', 'dikirim', 'jawatan', 'pegawai', 'menyelenggarakan', 'pos', 'surat', 'uang', 'barang', 'kirim'}	0	
				2	{'prangko', 'kirim', 'mengirim', 'pergi', 'membeli', 'pos', 'kantor',	0	

				'uang', 'surat'}	
			3	{'mengantarkan', 'kartu', 'warkat', 'tukang', 'pegawai', 'pos', 'pengantar', 'perantaraan', 'dikirim', 'surat'}	0
			4	{'menumpang', 'pengangkut', 'semarang', 'pos', 'kuda', 'kereta', 'zaman', 'surat'}	0
			5	{'kudanya', 'mengganti', 'perhentian', 'pos', 'segar', 'kereta', 'kusir', 'zaman'}	0
			6	{'perhentian', 'pos', 'rumah', 'kereta', 'trem'}	0
			7	{'sosial', 'unggahan', 'media', 'tulisan', 'pos', 'rumah', 'disebarkan'}	0
			8	{'berapi', 'tentara', 'penjagaan', 'terdepan', 'regu', 'ditempatkan', 'pos', 'pengamat', 'gunung', 'polisi'}	1
			9	{'dokter', 'tugas', 'pos', 'kedudukan'}	0
			10	{'pos', 'dokter', 'berkumpul', 'anggota', 'sekelompok'}	0

				11	{'disediakan', 'pos', 'dokter', 'uang', 'keperluan', 'jawatan'}	0	
				12	{'belanja', 'pos', 'anggaran', 'dokter'}	0	

**Tabel V- 5.** Sampel Hasil Analisis Proses WSD dengan *Stopword Removal*

No	Kata	Kalimat	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan
1	pos	Surat tersebut dikirim melalui kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	jawatan yang menyelenggarakan kirim-mengirim barang, surat, uang, dan sebagainya	Tidak Sesuai
2	pos	Pos tentara militer tepat berada di depan	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	Sesuai

Tabel V-6 akan menampilkan hasil analisis jumlah data, prediksi benar, dan prediksi salah yang diperoleh dari hasil percobaan menggunakan pra-pengolahan *stopword removal* berdasarkan kedua tabel diatas.

**Tabel V- 6.** Hasil Evaluasi Pengujian dengan *Stopword*

Jumlah Data	Prediksi Benar	Prediksi Salah
300	164	36

Berdasarkan hasil analisis diatas, akan dilakukan perhitungan akurasi berikut sebagai evaluasi dari pengujian WSD menggunakan proses pra-pengolahan *stopword removal*.

$$\text{Akurasi} = \frac{164}{300} \times 100\% = \mathbf{54.67\%}$$

#### 5.2.4 Hasil Konfigurasi 3

Hasil dari percobaan menggunakan pra-pengolahan *stemming* dan *stopword removal* dalam menghasilkan konteks dan proses dalam menentukan makna yang tepat akan ditunjukkan dengan menggunakan 2 sampel data dari *dataset* secara rinci pada Tabel V-7. Selanjutnya hasil prediksi makna akan dilabeli dengan keterangan ‘Sesuai’ atau ‘Tidak Sesuai’ berdasarkan kesesuaian makna yang dihasilkan sistem dengan makna sebenarnya, yang dapat dilihat pada Tabel V-8.

**Tabel V- 7.** Sampel Hasil Percobaan *Stemming* dan *Stopword Removal*

No	Kalimat	Kata	Konteks	Kandidat Makna		Overlap	Predicted Sense
				Indeks	Makna		
1	Surat tersebut dikirim melalui kantor pos	pos	['surat', 'kirim', 'kantor', 'pos']	0	{'standar', 'prosedur', 'operasional'}	0	[2] kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos
				1	{'kirim', 'pegawai', 'uang', 'jawat', 'selenggara', 'surat', 'barang', 'pos'}	2	
				2	{'kirim', 'uang', 'beli', 'kantor', 'prangko', 'pergi', 'surat', 'pos'}	3	
				3	{'kirim', 'tukang', 'pegawai', 'antar', 'kartu', 'warkat', 'antara', 'surat', 'pos'}	2	
				4	{'zaman', 'semarang', 'kuda', 'angkut', 'kereta', 'tumpang', 'surat', 'pos'}	1	
				5	{'zaman', 'ganti', 'henti', 'kusir', 'segar', 'kuda', 'kereta', 'pos'}	0	
				6	{'trem', 'rumah', 'kereta', 'henti', 'pos'}	0	

				7	{'rumah', 'sosial', 'media', 'tulis', 'sebar', 'unggah', 'pos'}	0	
				8	{'jaga', 'gunung', 'depan', 'polisi', 'amat', 'tentara', 'regu', 'tempat', 'rap', 'pos'}	0	
				9	{'dokter', 'duduk', 'pos', 'tugas'}	0	
				10	{'anggota', 'dokter', 'kumpul', 'kelompok', 'pos'}	0	
				11	{'jawab', 'sedia', 'dokter', 'perlu', 'uang', 'pos'}	0	
				12	{'anggar', 'dokter', 'belanja', 'pos'}	0	
2	Pos tentara militer tepat berada di depan	pos	['pos', 'tentara', 'militer']	0	{'standar', 'prosedur', 'operasional'}	0	[8] tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)
				1	{' kirim', 'pegawai', 'uang', 'jawab', 'selenggara', 'surat', 'barang', 'pos'}	0	
				2	{' kirim', 'uang', 'beli', 'kantor', 'prangko', 'pergi', 'surat', 'pos'}		



				3	{'kirim', 'tukang', 'pegawai', 'antar', 'kartu', 'warkat', 'antara', 'surat', 'pos'}	0	
				4	{'zaman', 'semarang', 'kuda', 'angkut', 'kereta', 'tumpang', 'surat', 'pos'}	0	
				5	{'zaman', 'ganti', 'henti', 'kusir', 'segar', 'kuda', 'kereta', 'pos'}	0	
				6	{'trem', 'rumah', 'kereta', 'henti', 'pos'}	0	
				7	{'rumah', 'sosial', 'media', 'tulisi', 'sebar', 'unggah', 'pos'}	0	
				8	{'jaga', 'gunung', 'depan', 'polisi', 'amat', 'tentara', 'regu', 'tempat', 'rap', 'pos'}	1	
				9	{'dokter', 'duduk', 'pos', 'tugas'}	0	
				10	{'anggota', 'dokter', 'kumpul', 'kelompok', 'pos'}	0	

				11	{'jawat', 'sedia', 'dokter', 'perlu', 'uang', 'pos'}	0	
				12	{'anggar', 'dokter', 'belanja', 'pos'}	0	

**Tabel V- 8.** Sampel Hasil Analisis Proses WSD dengan *Stemming* dan *Stopword Removal*

No	Kata	Kalimat	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan
1	pos	Surat tersebut dikirim melalui kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	Sesuai
2	pos	Pos tentara militer tepat berada di depan	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	Sesuai

Tabel V-9 akan menampilkan hasil analisis jumlah data, prediksi benar, dan prediksi salah yang diperoleh dari hasil percobaan menggunakan pra-pengolahan *stemming* berdasarkan kedua tabel diatas.

**Tabel V- 9.** Hasil Evaluasi Pengujian dengan *Stemming* dan *Stopword*

Jumlah Data	Prediksi Benar	Prediksi Salah
300	174	126

Berdasarkan hasil analisis diatas, akan dilakukan perhitungan akurasi berikut sebagai evaluasi dari pengujian WSD menggunakan proses pra-pengolahan *stemming* dan *stopword removal*.

$$\text{Akurasi} = \frac{174}{300} \times 100\% = \mathbf{58\%}$$

#### 5.2.5 Hasil Konfigurasi 4

Hasil dari percobaan menggunakan pra-pengolahan *stemming* dalam menghasilkan konteks dan proses dalam menentukan makna yang tepat akan ditunjukkan dengan menggunakan 2 sampel data dari *dataset* secara rinci pada Tabel V-10. Selanjutnya hasil prediksi makna akan dilabeli dengan keterangan ‘Sesuai’ atau ‘Tidak Sesuai’ berdasarkan kesesuaian makna yang dihasilkan sistem dengan makna sebenarnya, yang dapat dilihat pada Tabel V-11.

**Tabel V- 10.** Sampel Hasil Percobaan tanpa *Stemming* dan *Stopword Removal*

No	Kalimat	Kata	Konteks	Kandidat Makna		Overlap	Predicted Sense
				Indeks	Makna		
1	Surat tersebut dikirim melalui kantor pos	pos	[surat, 'tersebut', 'dikirim', 'melalui', 'kantor', 'pos']	0	{'operasional', 'prosedur', 'standar'}	0	[1] jawatan yang menyelenggarakan kirim-mengirim barang, surat, uang, dan sebagainya
				1	{'menyelenggarakan', 'pos', 'yang', 'kirim', 'dikirim', 'melalui', 'surat', 'sebagainya', 'mengirim', 'uang', 'dan', 'barang', 'jawatan', 'pegawai'}	3	
				2	{'pergi', 'pos', 'ke', 'kirim', 'kantor', 'surat', 'mengirim', 'uang', 'prangko', 'dan', 'ia', 'membeli', 'untuk', 'sebagainya', 'tempat'}	2	
				3	{'perantaraan', 'dengan', 'pos', 'warkat', 'yang', 'tukang', 'pengantar', 'dikirim', 'mengantarkan', 'surat', 'dan', 'kartu', 'sebagainya', 'pegawai'}	2	
				4	{'pengangkut', 'dari',	1	

					'semarang', 'pos', 'menumpang', 'zaman', 'dahulu', 'pada', 'kuda', 'kereta', 'surat', 'dan', 'sebagainya', 'kami'}		
				5	{'mengganti', 'pos', 'dengan', 'yang', 'zaman', 'dahulu', 'pada', 'segar', 'masih', 'kereta', 'kudanya', 'tiba', 'tempat', 'di', 'kusir', 'perhentian'}	0	
				6	{'pos', 'trem', 'lama', 'kereta', 'mereka', 'belakang', 'dan', 'rumah', 'api', 'sebagainya', 'di', 'perhentian'}	0	
				7	{'atau', 'disebarkan', 'pos', 'yang', 'sosial', 'media', 'melalui', 'lama', 'mereka', 'belakang', 'tulisan', 'unggahannya', 'di', 'rumah'}	1	
				8	{'pos', 'regu', 'dua', 'ditempatkan', 'terdepan', 'penjagaan', 'polisi',	0	

					'tentara', 'berapi', 'dan', 'sebagainya', 'tempat', 'di', 'gunung', 'pengamat'}		
				9	{'kedudukan', 'pos', 'harus', 'yang', 'siap', 'selalu', 'melakukan', 'sebagainya', 'dokter', 'dan', 'tugas', 'tempat', 'di', 'orang', 'masing'}	0	
				10	{'anggota', 'pos', 'harus', 'siap', 'selalu', 'berkumpul', 'dokter', 'sekelompok', 'tempat', 'di', 'masing'}	0	
				11	{'pos', 'harus', 'yang', 'siap', 'keperluan', 'selalu', 'sebagainya', 'masing', 'uang', 'dokter', 'dan', 'suatu', 'untuk', 'jawatan', 'disediakan', 'di', 'jumlah'}	0	
				12	{'dari', 'pos', 'harus', 'siap', 'selalu', 'anggaran', 'belanja', 'dokter', 'mata', 'bagian', 'di', 'masing'}	0	

2	Pos tentara militer tepat berada di depan	pos	['pos', 'tentara', 'militer', 'tepat', 'berada', 'di', 'depan']	0	{'operasional', 'prosedur', 'standar'}	0
				1	{'menyelenggarakan', 'pos', 'yang', ' kirim', 'dikirim', 'melalui', 'surat', 'sebagainya', 'mengirim', 'uang', 'dan', 'barang', 'jawatan', 'pegawai'}	0
				2	{'pergi', 'pos', 'ke', ' kirim', 'kantor', 'surat', 'mengirim', 'uang', 'prangko', 'dan', 'ia', 'membeli', 'untuk', 'sebagainya', 'tempat'}	0
				3	{'perantaraan', 'dengan', 'pos', 'warkat', 'yang', 'tukang', 'pengantar', 'dikirim', 'mengantarkan', 'surat', 'dan', 'kartu', 'sebagainya', 'pegawai'}	0
				4	{'pengangkut', 'dari', 'semarang', 'pos', 'menumpang', 'zaman', 'dahulu', 'pada', 'kuda',	0

					'kereta', 'surat', 'dan', 'sebagainya', 'kami'}		
				5	{'mengganti', 'pos', 'dengan', 'yang', 'zaman', 'dahulu', 'pada', 'segar', 'masih', 'kereta', 'kudanya', 'tiba', 'tempat', 'di', 'kusir', 'perhentian'}	1	
				6	{'pos', 'trem', 'lama', 'kereta', 'mereka', 'belakang', 'dan', 'rumah', 'api', 'sebagainya', 'di', 'perhentian'}	1	
				7	{'atau', 'disebarkan', 'pos', 'yang', 'sosial', 'media', 'melalui', 'lama', 'mereka', 'belakang', 'tulisan', 'unggah', 'di', 'rumah'}	1	
				8	{'pos', 'regu', 'dua', 'ditempatkan', 'terdepan', 'penjagaan', 'polisi', 'tentara', 'berapi', 'dan', 'sebagainya', 'tempat', 'di', 'gunung', 'pengamat'}	2	



				9	{'kedudukan', 'pos', 'harus', 'yang', 'siap', 'selalu', 'melakukan', 'sebagainya', 'dokter', 'dan', 'tugas', 'tempat', 'di', 'orang', 'masing'}	1	
				10	{'anggota', 'pos', 'harus', 'siap', 'selalu', 'berkumpul', 'dokter', 'sekelompok', 'tempat', 'di', 'masing'}	1	
				11	{'pos', 'harus', 'yang', 'siap', 'keperluan', 'selalu', 'sebagainya', 'masing', 'uang', 'dokter', 'dan', 'suatu', 'untuk', 'jawatan', 'disediakan', 'di', 'jumlah'}	1	
				12	{'dari', 'pos', 'harus', 'siap', 'selalu', 'anggaran', 'belanja', 'dokter', 'mata', 'bagian', 'di', 'masing'}	1	

**Tabel V- 11.** Sampel Hasil Analisis Proses WSD tanpa *Stemming* dan *Stopword Removal*

No	Kata	Kalimat	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan
1	pos	Surat tersebut dikirim melalui kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	jawatan yang menyelenggarakan kirim-mengirim barang, surat, uang, dan sebagainya	Tidak Sesuai
2	pos	Pos tentara militer tepat berada di depan	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	Sesuai

Tabel V-12 akan menampilkan hasil analisis jumlah data, prediksi benar, dan prediksi salah yang diperoleh dari hasil percobaan menggunakan pra-pengolahan *stemming*.

**Tabel V- 12.** Hasil Pengujian Tanpa *Stemming* dan *Stopword*

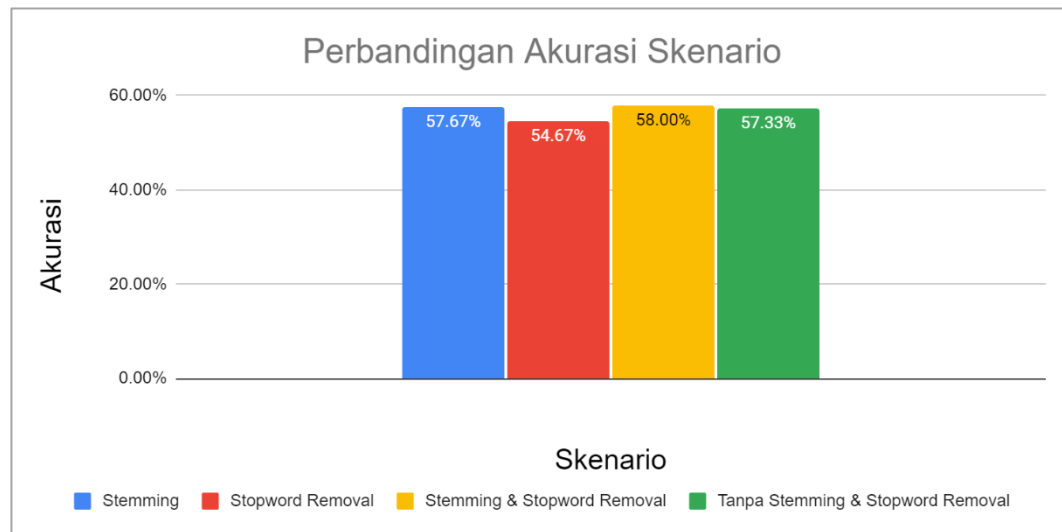
Jumlah Data	Prediksi Benar	Prediksi Salah
300	172	128

Berdasarkan hasil analisis diatas, akan dilakukan perhitungan akurasi berikut sebagai evaluasi dari pengujian WSD tanpa menggunakan proses pra-pengolahan *stemming* dan *stopword removal*.

$$\text{Akurasi} = \frac{172}{300} \times 100\% = \mathbf{57.33\%}$$

### 5.3 Analisis Hasil Penelitian

Gambar V-1 akan menunjukkan perbandingan hasil perhitungan akurasi yang diperoleh dari analisis setiap skenario sebagai evaluasi dalam memilih proses terbaik yang akan digunakan dalam WSD menggunakan algoritma *Simplified Lesk* pada kalimat Bahasa Indonesia.



**Gambar V- 1.** Perbandingan Hasil Konfigurasi

Berdasarkan perbandingan hasil akurasi pada gambar diatas, maka proses yang akan ditetapkan untuk menghasilkan konteks dan relasi dengan makna sehingga dapat mengidentifikasi makna yang tepat yaitu proses pra-pengolahan dengan *stemming* dan *stopword removal* dengan perolehan akurasi tertinggi yaitu 58% dari keseluruhan data yang diuji, sehinggann konfigurasi percobaan dengan proses pra-pengolahan *stemming* dan *stopword removal* merupakan konfigurasi terbaik yang akan diimplementasikan pada WSD menggunakan algoritma *Simplified Lesk*.

Tabel V-13 akan berisi 5 data sampel untuk menampilkan proses WSD dalam mengidentifikasi makna yang tepat dan telah mengimplementasikan konfigurasi terbaik, dan Tabel V-14 menunjukkan hasil WSD menggunakan algoritma *Simplified Lesk* yang diperoleh serta keterangan berdasarkan kesesuaiannya dengan makna sebenarnya. Selanjutnya akan dilakukan analisis terhadap sampel yang prediksi maknanya tidak sesuai dengan makna sebenarnya.

**Tabel V- 13.** Sampel Proses Identifikasi WSD

No	Kalimat	Kata	Konteks	Kandidat Makna		Overlap	Predicted Sense
				Indeks	Makna		
1	Paman membawa banyak oleh-oleh makanan, yang paling terkenal adalah tahu sumedang	tahu	['paman', 'bawa', 'makan', 'kenal', 'sumedang']	0	{'erti', 'alami', 'mesin', 'saksi', 'tolong', 'perkara'}	0	[1] kenal (akan); mengenal
				1	{'kenal', 'sanak', 'saudara'}	1	
				2	{'peduli', 'anak', 'indah'}	0	
				3	{'erti'}	0	
				4	{'cakap', 'mesin', 'pandai'}	0	
				5	{'kurang', 'insaf', 'sadar'}	0	
				6	{'tinju', 'bolos', 'adik', 'menang'}	0	
7	{'kedelai', 'makan', 'menang', 'tinju', 'cetak', 'putih', 'rebus', 'adik', 'giling', 'bolos', 'halus'}	1					
2	Halaman yang indah menjadi idaman setiap orang yang punya rumah	halaman	['halaman', 'indah', 'idam', 'rumah']	0	{'tanah', 'pekarang', 'cemara', 'tanam', 'rumah', 'halaman', 'sekolah'}	1	[0] pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)
				1	{'pagina', 'buku', 'majalah', 'kabar', 'gambar', 'surat', 'lembar',	0	

					'halaman', 'muka'}		
3	Saat kau ingin rumahmu terang alami, pakailah genting kaca di bagian tertentu atap rumahmu	genting	['kau', 'rumah', 'terang', 'alami', 'pakai', 'genting', 'kaca', 'atap']	0	{'genting', 'tipis', 'sempit', 'pinggang'}	0	[3] tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya
				1	{'genting', 'putus', 'tali'}	0	
				2	{'jalan', 'runding', 'bahaya', 'tegang', 'bencana', 'buntu', 'perang', 'timbul', 'genting', 'temu', 'tambah'}	0	
				3	{'jalan', 'atap', 'cetak', 'buat', 'tanah', 'liat', 'bentuk', 'runding', 'tutup', 'buntu', 'rumah', 'genting', 'bakar', 'temu', 'bermacam-macam', 'tambah'}	2	
4	Setiap tahun, ani menerima bunga 5 persen dari bank	bunga	['ani', 'terima', 'bunga', 'persen', 'bank']	0	{'kembang', 'warna', 'harum', 'elok', 'tumbuh', 'mangga', 'bau', 'bunga', 'rambutan'}	0	[0] bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang
				1	{'jenis', 'melati', 'mawar', 'bunga'}	0	
				2	{'pamor', 'ukir', 'hias', 'melati', 'mawar', 'gambar', 'bunga', 'kain'}	0	

				3	{'bunga', 'tambah', 'cerita', 'indah'}	0	
				4	{'bunga', 'hasil', 'tuntut'}	0	
				5	{'kampung', 'cantik', 'gadis', 'bunga', 'elok', 'anggap'}	0	
				6	{'nama', 'pakai', 'bangkai', 'tongkeng', 'jenis', 'kampung', 'tumbuh', 'kerabat', 'gadis', 'bunga', 'manila'}	0	
				7	{'persentase', 'dasar', 'tentu', 'modal', 'guna', 'sepakat', 'nyata', 'pokok', 'kampung', 'imbal', 'bayar', 'gadis', 'jasa', 'bunga', 'uang'}	0	
				8	{'kampung', 'gadis', 'dapat', 'bunga', 'investasi', 'modal'}	0	
5	Risqi Aris babak belur dihajar oleh tetangganya	babak	['risqi', 'babak', 'bur', 'hajar', 'tetangga']	0	{'tunjuk', 'babak', 'lakon', 'adegan', 'drama'}	0	[0] bagian besar dalam suatu drama atau lakon (terdiri atas
				1	{'proses', 'runding', 'mula', 'tiga', 'sengketa', 'babak',	0	

					'jadi', 'peristiwa', 'ada', 'negara'}		beberapa adegan)
			2		{'babak', 'main', 'sebelas', 'ronde', 'kuat'}	0	
			3		{'babak', 'sebelas', 'main', 'kuat', 'lecet', 'kulit'}	0	
			4		{'babak', 'cari', 'sebelas', 'main', 'kuat'}	0	
			5		{'babak', 'cari', 'babakan', 'sebelas', 'main', 'kuat'}	0	

**Tabel V- 14.** Sampel Hasil Pengujian WSD

No	Kata	Kalimat	<i>Actual Sense</i>	<i>Predicted Sense</i>	Keterangan
1	tahu	Paman membawa banyak oleh-oleh makanan, yang paling terkenal adalah tahu sumedang	makanan dari kedelai putih yang digiling halus-halus, direbus, dan dicetak	kenal (akan); mengenal	Tidak Sesuai
2	halaman	Halaman yang indah menjadi idaman setiap orang yang punya rumah	pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan	pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan	Sesuai



			sebagainya)	sebagainya)	
3	genting	Saat kau ingin rumahmu terang alami, pakailah genting kaca di bagian tertentu atap rumahmu	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	Sesuai
4	bunga	Setiap tahun, ani menerima bunga 5 persen dari bank	imbalan jasa untuk penggunaan uang atau modal yang dibayar pada waktu tertentu berdasarkan ketentuan atau kesepakatan, umumnya dinyatakan sebagai persentase dari modal pokok	bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang	Tidak Sesuai
5	babak	Risqi Aris babak belur dihajar oleh tetangganya	lecet (tentang kulit)	bagian besar dalam suatu drama atau lakon (terdiri atas beberapa adegan)	Tidak Sesuai

Berdasarkan hasil pada kedua tabel, terdapat 2 faktor yang memungkinkan dari penyebab hasil prediksi yang tidak sesuai dengan makna sebenarnya.

**1. Tidak adanya makna yang beririsan dengan konteks**

Konteks yang dihasilkan serta kamus data untuk mengambil makna masih kurang lengkap, sehingga akan mempengaruhi jumlah irisan antara makna dengan konteks serta acuan kamus data.

**2. Makna yang dipilih diantara kandidat makna yang memiliki jumlah *overlapping* yang sama tidak selalu tepat**

Proses pemilihan makna diantara kandidat makna dengan jumlah *overlapping* yang sama masih kurang maksimal karena akan selalu memilih makna paling pertama dan masih minim dalam mendeteksi makna yang paling sesuai dengan konteks.

Poin-poin yang disebutkan diatas juga disimpulkan sebagai kelemahan dari cara kerja algoritma *Simplified Lesk* dalam mengidentifikasi makna kata berdasarkan relasi makna kata dengan kalimat, serta ketergantungannya terhadap definisi yang diperoleh dari kamus data. Semakin banyaknya konteks dan lengkapnya makna yang diperoleh dari kamus sehingga meningkatnya jumlah *overlapping*, maka akan semakin tepat algoritma dalam mengidentifikasi makna yang sesuai dari kata ambigu sesuai konteks kalimat.

#### 5.4 Kesimpulan

Pengujian serta analisis hasil dari WSD menggunakan algoritma Simplified Lesk telah dijelaskan dan ditampilkan secara detail, dan dapat disimpulkan bahwa algoritma *Simplified Lesk* memperoleh hasil yang kurang memuaskan karena terdapat beberapa kekurangan. Analisis juga menunjukkan bahwa data kalimat dan kamus yang dijadikan sebagai acuan sangat berpengaruh dalam mengidentifikasi makna yang tepat.

## **BAB VI**

### **KESIMPULAN DAN SARAN**

#### **6.1 Pendahuluan**

Bagian ini membahas mengenai kesimpulan dan saran berdasarkan analisis sebelumnya untuk dijadikan rujukan penelitian di masa depan khususnya di bidang *Word Sense Disambiguation* (WSD).

#### **6.2 Kesimpulan**

Setelah melakukan analisis penelitian yang telah diuraikan dalam bab sebelumnya, dapat ditarik kesimpulan bahwa:

1. Perangkat lunak atau sistem *Word Sense Disambiguation* (WSD) menggunakan algoritma *Simplified Lesk* berhasil diimplementasikan.
2. Program dapat melakukan disambiguasi makna kata dengan akurasi yang kurang memuaskan yaitu 58% dipengaruhi oleh data yang digunakan. Semakin sedikit konteks yang diperoleh serta makna yang kurang lengkap pada kamus, maka akurasi yang didapat akan semakin menurun dan tidak dapat mengidentifikasi makna yang sesuai.

#### **6.3 Saran**

Pada penelitian selanjutnya diharapkan:

1. Menggunakan *dataset* Bahasa Indonesia untuk menangani kasus ambiguitas yang lebih bervariasi.

2. Menambahkan kamus data berbeda sebagai acuan *glossary* kata yang lebih kompleks seperti *Wordnet* untuk Bahasa Indonesia atau kamus lainnya.
3. Menambahkan metode lain dalam melakukan disambiguasi makna kata yang tepat untuk memperoleh kinerja yang lebih baik.

## DAFTAR PUSTAKA

- Abdo, A. (2023). Enhanced Word Sense Disambiguation Algorithm for Afaan Oromoo. *International Journal of Information Engineering and Electronic Business*, 15(1), 41–50. <https://doi.org/10.5815/ijieeb.2023.01.04>
- Aliwy, A., & Abbas, A. (2015). *Improvement WSD Dictionary Using Annotated Corpus and Testing it with Simplified Lesk Algorithm*. 89–97. <https://doi.org/10.5121/csit.2015.50409>
- Aliwy, A. H., & Taher, H. A. (2019). Word Sense Disambiguation: Survey study. In *Journal of Computer Science* (Vol. 15, Issue 7, pp. 1004–1011). Science Publications. <https://doi.org/10.3844/jcssp.2019.1004.1011>
- Anahit Martirosyan. (2005). *Word Sense Disambiguation And Context*.
- Anwar, A. (2014). A Review of RUP (Rational Unified Process). In *Ashraf Anwar International Journal of Software Engineering (IJSE)* (Issue 5).
- Basuki, S., Sofyan Kholimi, A., Eko Minarno, A., Sumadi, F., & Effendy, M. (2019). *Word Sense Disambiguation (WSD) for Indonesian Homograph Word Meaning Determination by LESK Algorithm Application*.
- Craggs, D. J. (2011). *An analysis and comparison of predominant word sense disambiguation algorithms*. [https://ro.ecu.edu.au/theses\\_hons/4](https://ro.ecu.edu.au/theses_hons/4)
- Fazar, D., & Widiastuti, N. (2017). *Lexical Chain dan Word Sense Disambiguation Untuk Peringkasan Artikel Berbahasa Indonesia Lexical Chain and Word Sense Disambiguation to Summarize Article in Bahasa Indonesia*. 16(2), 195–207.
- Fujii, A., & Tanaka, H. (1998). *Corpus-Based Word Sense Disambiguation*.

- Hung, J. C. (2005). *Applying Word Sense Disambiguation to Question Answering System for E-Learning*.
- Kannan, S., Gurusamy, V., Vijayarani, S., Ilamathi, J., & Nithya, M. (2014). *Preprocessing Techniques for Text Mining*.
- Kilgarriff, A. (1998). Gold standard datasets for evaluating word sense disambiguation programs. In *Computer Speech and Language* (Vol. 12).
- Mahendra, R., Septiantri, H., Wibowo, H. A., Manurung, R., & Adriani, M. (2018). *Cross-Lingual and Supervised Learning Approach for Indonesian Word Sense Disambiguation Task*.
- Mccarthy, D. (2009). Word sense disambiguation: An overview. In *Linguistics and Language Compass* (Vol. 3, Issue 2, pp. 537–558). Blackwell Publishing Inc. <https://doi.org/10.1111/j.1749-818X.2009.00131.x>
- Mubarok, F., & Hadijah, I. (2015). *Perbandingan Antara Metode RUP dan Prototype Dalam Aplikasi Penerimaan Siswa Baru Berbasis Web*.
- Pal, A. R., Maiti, P. K., & Saha, D. (2013). An Approach to Automatic Text Summarization Using Simplified Lesk Algorithm and Wordnet. *International Journal of Control Theory and Computer Modeling*, 3(5), 15–23. <https://doi.org/10.5121/ijctcm.2013.3502>
- Pal, A., & Saha, D. (2014). *An Approach to Automatic Text Summarization using WordNet*.
- Pal, A., & Saha, D. (2015). Word Sense Disambiguation: A Survey. *International Journal of Control Theory and Computer Modeling*, 5(3), 1–16. <https://doi.org/10.5121/ijctcm.2015.5301>
- Patwardhan, S., Banerjee, S., & Pedersen, T. (2005). *SenseRelate::TargetWord-A Generalized Framework for Word Sense Disambiguation*. <http://search.cpan.org/dist/WordNet-QueryData>

- Rada Mihalcea. (2007). *5 Knowledge-Based Methods for WSD*.
- Resnik, P., & Yarowsky, D. (1997). *A Perspective on Word Sense Disambiguation Methods and Their Evaluation*.
- Ridwansyah, D. A. (2021). *Peringkasan Teks Otomatis Ekstraktif Menggunakan Metode Non-Negative Matrix Factorization Dan Word Sense Disambiguation Pada Berita Berbahasa Indonesia*.
- Torres, S., Gelbukh, A., Profesional Adolfo-López Mateos, U., Juan de Dios Bátiz, A., & Othón de Mendizábal, M. (2009). *Comparing Similarity Measures for Original WSD Lesk Algorithm*. [www.gelbukh.com](http://www.gelbukh.com)
- Trismanto. (2018). *Ambiguitas Dalam Bahasa Indonesia*.



## **LAMPIRAN**

### **Lampiran 1. Kode Program**

<https://gitlab.com/nurulakhni12/wsd-app.git>

**Lampiran 2.** Hasil Evaluasi Sampel 100 Data

No	Kalimat	Kata Ambigu	Actual Sense	Predicted Sense	Keterangan
1	aku membaca buku tentang sejarah di perpustakaan	buku	lembar kertas yang berjilid, berisi tulisan atau kosong; kitab	lembar kertas yang berjilid, berisi tulisan atau kosong; kitab	Sesuai
2	Tebu itu mempunyai banyak buku	buku	tempat pertemuan dua ruas (jari, buluh, tebu)	tempat pertemuan dua ruas (jari, buluh, tebu)	Sesuai
3	pada teras di bawahnya terbentang taman yang luas	teras	bidang tanah datar yang miring; bidang tanah yang lebih tinggi daripada yang lain (biasanya ditumbuhi rumput)	bidang tanah datar yang miring; bidang tanah yang lebih tinggi daripada yang lain (biasanya ditumbuhi rumput)	Sesuai
4	pada irisan penampang kelihatan teras yang dikelilingi lingkaran-lingkaran kayu yang dapat menunjukkan umurnya	teras	bagian kayu yang keras; inti kayu	bagian kayu yang keras; inti kayu	Sesuai
5	bukan hanya pembangunan	mental	bersangkutan dengan batin	bersangkutan dengan batin	Sesuai

	fisik yang diperhatikan, melainkan juga pembangunan mental		dan watak manusia, yang bukan bersifat badan atau tenaga	dan watak manusia, yang bukan bersifat badan atau tenaga	
6	akibat tabrakan yang terjadi tubuh itu mental sekitar 2 meter	mental	terpelanting; terpentak	bersangkutan dengan batin dan watak manusia, yang bukan bersifat badan atau tenaga	Tidak Sesuai
7	perkara mesin, dia lebih tahu daripada saya	tahu	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	Sesuai
8	ayah membawa tahu sumedang saat pergi ke sumedang	tahu	makanan dari kedelai putih yang digiling halus-halus, direbus, dan dicetak	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	Tidak Sesuai
9	padi jangan ditanam terlalu rapat	rapat	kerap (tentang tanaman, anyaman, dan sebagainya)	kerap (tentang tanaman, anyaman, dan sebagainya)	Sesuai
10	kami mengadakan rapat di dalam ruangan	rapat	pertemuan (kumpulan) untuk membicarakan sesuatu; sidang; majelis	hampir tidak berantara; dekat sekali (tidak renggang)	Tidak Sesuai

11	matahari terbenam menandakan malam telah tiba	malam	waktu setelah matahari terbenam hingga matahari terbit	waktu setelah matahari terbenam hingga matahari terbit	Sesuai
12	ia kehabisan malam saat membatik	malam	lilin (dipakai untuk membatik)	lilin (dipakai untuk membatik)	Sesuai
13	dia mendapatkan bunga mawar merah yang harum dari penggemarnya	bunga	jenis untuk berbagai-bagai bunga	bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang	Tidak Sesuai
14	dia mendapatkan bunga bank atas jasa investasi modalnya	bunga	pendapatan atas setiap investasi modal	imbalan jasa untuk penggunaan uang atau modal yang dibayar pada waktu tertentu berdasarkan ketentuan atau kesepakatan, umumnya dinyatakan sebagai persentase dari modal pokok	Tidak Sesuai
15	sepatu dengan hak tinggi sedang digemari oleh wanita karier	hak	telapak sepatu pada bagian tumit yang relatif tinggi	telapak sepatu pada bagian tumit yang relatif tinggi	Sesuai

16	ia juga punya hak tinggal di sini	hak	kekuasaan yang benar atas sesuatu atau untuk menuntut sesuatu	kekuasaan yang benar atas sesuatu atau untuk menuntut sesuatu	Sesuai
17	satu minggu ini, dia sudah empat kali datang ke rumahku	kali	kata untuk menyatakan kekerapan tindakan	kata untuk menyatakan kekerapan tindakan	Sesuai
18	ia melihat bebek berenang di kali	kali	sungai	kata untuk menyatakan kekerapan tindakan	Tidak Sesuai
19	ia mendapat gelar srikandi dari kawan-kawannya	gelar	julukan yang berhubungan dengan keadaan atau tabiat orang; sebutan	julukan yang berhubungan dengan keadaan atau tabiat orang; sebutan	Sesuai
20	tolong gelar tikar itu	gelar	hampar	sebutan kehormatan, kebangsawanan, atau kesarjanaan yang biasanya ditambahkan pada nama orang seperti raden, tengku, doktor, sarjana ekonomi	Tidak Sesuai
21	jarak antara Mekah dan Medinah kami tempuh	jarak	ruang sela (panjang atau jauh) antara dua benda atau tempat	ruang sela (panjang atau jauh) antara dua benda atau tempat	Sesuai

	dengan bus dalam 5 jam				
22	ia sedang mencari biji jarak	jarak	biji jarak	perdu, tinggi 2 m, batangnya mudah patah, daunnya besar berwarna hijau kebiru-biruan, buahnya berbiji polong, bijinya sebesar kacang tanah, ketika tua berwarna hitam, dapat digunakan sebagai bahan minyak pelumas; dulang	Tidak Sesuai
23	ia mendapat salam dari ayahnya	salam	pernyataan hormat; tabik	damai	Tidak Sesuai
24	ia lupa memasukkan salam kemasakannya	salam	tanaman termasuk suku Myrtaceae, pohonnya bertajuk lebat, tingginya mencapai 25 m, daunnya berbentuk bulat telur atau lonjong, berbau harum, biasanya digunakan sebagai	damai	Tidak Sesuai

			bumbu masak		
25	hal ini sudah jamak terjadi	jamak	lazim; tidak aneh; lumrah; wajar	lazim; tidak aneh; lumrah; wajar	Sesuai
26	ia sedang menjamak salat	jamak	penggabungan (tentang salat wajib); antara salat Zuhur dan salat Asar atau salat Magrib dan salat Isya	penggabungan (tentang salat wajib); antara salat Zuhur dan salat Asar atau salat Magrib dan salat Isya	Sesuai
27	roman wajahnya berubah menjadi sedih	roman	rupa muka; wajah	rupa muka; wajah	Sesuai
28	Film Roman Picisan adalah film drama percintaan Indonesia tahun 2018 yang disutradarai oleh Monty Tiwa	roman	(cerita) percintaan	(cerita) percintaan	Sesuai
29	buka buku pada halaman 50	halaman	muka dari lembaran buku (majalah, surat kabar, dan sebagainya); pagina	muka dari lembaran buku (majalah, surat kabar, dan sebagainya); pagina	Sesuai
30	halaman rumahnya	halaman	pekarangan rumah (sekolah	pekarangan rumah (sekolah	Sesuai

	ditanami cemara		dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)	dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)	
31	angin terjadi karena gerakan hawa	hawa	campuran berbagai-bagai gas yang meliputi bumi; udara	campuran berbagai-bagai gas yang meliputi bumi; udara	Sesuai
32	ia tidak dapat menahan hawa nafsunya	hawa	keinginan; nafsu	keinginan; nafsu	Sesuai
33	baik sanak keluarga maupun orang helat banyak yang hadir pada resepsi pernikahannya	helat	asing; lain; bukan keluarga	asing; lain; bukan keluarga	Sesuai
34	banyak tamu yang datang untuk meramaikan helat putri tunggalnya	helat	pesta perkawinan dan sebagainya	pesta perkawinan dan sebagainya	Sesuai
35	hemat pangkal kaya, rajin pangkal pandai	hemat	berhati-hati dalam membelanjakan uang, dan sebagainya; tidak boros; cermat	berhati-hati dalam membelanjakan uang, dan sebagainya; tidak boros; cermat	Sesuai



36	ia mendengarkan pelajaran dengan hemat dan cermat	hemat	penuh minat dan perhatian; (dengan) saksama; teliti	penuh minat dan perhatian; (dengan) saksama; teliti	Sesuai
37	majikan itu sangat kejam, tidak mau menaikkan upah buruhnya barang sedikit juga	kejam	sangat kikir	sangat kikir	Sesuai
38	sungguhpun matanya kejam, ia tidak tidur	kejam	tertutup rapat (tentang mata); pejam	tertutup rapat (tentang mata); pejam	Sesuai
39	perkataannya lemak manis	lemak	sedap; enak	sedap; enak	Sesuai
40	lemak penyelar daging	lemak	zat minyak yang melekat pada daging; gemuk	zat minyak yang melekat pada daging; gemuk	Sesuai
41	berbagai macam pengaruh negatif dapat disebabkan oleh sejarah politik	negatif	kurang baik; menyimpang dari ukuran umum	kurang baik; menyimpang dari ukuran umum	Sesuai
42	kabel positif jangan langsung kau hubungkan dengan kabel negatif	negatif	kabel yang aliran listriknya berpotensi lebih rendah	kabel yang aliran listriknya berpotensi lebih rendah	Sesuai
43	pupus sudah harapannya	pupus	habis sama sekali; hilang	habis sama sekali; hilang	Sesuai

	selama ini		lenyap; punah	lenyap; punah	
44	pupus daun pisang tersebut terjatuh	pupus	pucuk yang masih muda; daun muda	pucuk yang masih muda; daun muda	Sesuai
45	surat tersebut dikirim melalui kantor pos	pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	kantor tempat kirim-mengirim surat, uang, dan sebagainya; kantor pos	Sesuai
46	pos tentara militer tepat berada di depan	pos	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	tempat penjagaan (tentara, polisi, pengamat gunung berapi, dan sebagainya)	Sesuai
47	di negara yang rusuh itu sering timbul pemberontakan	rusuh	kacau; ribut; gaduh; huru-hara	kacau; ribut; gaduh; huru-hara	Sesuai
48	pada babak pertama, kedua kesebelasan masih bermain sama kuat	babak	bagian permainan yang waktunya tertentu; ronde	bagian permainan yang waktunya tertentu; ronde	Sesuai
49	roni babak belur dihajar oleh lawannya	babak	lecet (tentang kulit)	bagian besar dalam suatu drama atau lakon (terdiri atas beberapa adegan)	Tidak Sesuai

50	Ular kobra memiliki bisa yang sangat berbahaya dan mematikan	bisa	zat racun yang dapat menyebabkan luka, busuk, atau mati bagi sesuatu yang hidup (biasanya terdapat pada binatang)	zat racun yang dapat menyebabkan luka, busuk, atau mati bagi sesuatu yang hidup (biasanya terdapat pada binatang)	Sesuai
51	Dia bisa membaca tapi tidak dengan menulis	bisa	mampu (kuasa melakukan sesuatu); dapat	mampu (kuasa melakukan sesuatu); dapat	Sesuai
52	buah apel sangat kaya akan vitamin C	apel	buah apel	naik banding kepada pengadilan yang lebih tinggi; permohonan pemeriksaan ulang pada pengadilan tingkat kedua (pengadilan tinggi) terhadap keputusan pengadilan tingkat pertama (pengadilan negeri)	Tidak Sesuai
53	apabila tidak puas dengan putusan pengadilan negeri boleh minta apel kepada pengadilan tinggi	apel	naik banding kepada pengadilan yang lebih tinggi; permohonan pemeriksaan ulang pada pengadilan tingkat	naik banding kepada pengadilan yang lebih tinggi; permohonan pemeriksaan ulang pada pengadilan tingkat	Sesuai

			kedua (pengadilan tinggi) terhadap keputusan pengadilan tingkat pertama (pengadilan negeri)	kedua (pengadilan tinggi) terhadap keputusan pengadilan tingkat pertama (pengadilan negeri)	
54	amanat apel pagi dibacakan oleh pemimpin apel	apel	upacara	wajib hadir dalam suatu upacara resmi (bersifat kemiliteran) untuk diketahui hadir tidaknya atau untuk mendengar amanat	Tidak Sesuai
55	Genting rumahku rusak akibat angin puting beliung kemarin malam	genting	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	Sesuai
56	setelah perundingan menemui jalan buntu, keadaan bertambah genting	genting	tegang; berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)	tegang; berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)	Sesuai
57	dia merasakan sesak akibat asma yang dideritanya	asma	gangguan pernapasan yang sering bersifat alergis,	gangguan pernapasan yang sering bersifat alergis,	Sesuai

			ditandai dengan sulit bernapas dan rasa sesak dalam dada; penyakit sesak napas; bengek	ditandai dengan sulit bernapas dan rasa sesak dalam dada; penyakit sesak napas; bengek	
58	kami berbuat baik terhadap siapapun semata-mata untuk meluhurkan asma Tuhan	asma	nama (bagi Tuhan)	nama (bagi Tuhan)	Sesuai
59	semua orang heran bahwa istrinya dapat bersikap baik pada madunya	madu	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	Sesuai
60	sarang lebah ini jika diperas mengeluarkan madu	madu	cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)	cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)	Sesuai
61	tamu itu gondok karena diperlakukan tidak sewajarnya	gondok	mendongkol; marah yang tertahan	mendongkol; marah yang tertahan	Sesuai
62	gondok merupakan	gondok	penyakit bengkak pada leher	penyakit bengkak pada leher	Sesuai

	penyakit pembengkakan pada leher		depan karena kelenjar yang menjadi besar; beguk	depan karena kelenjar yang menjadi besar; beguk	
63	Para pihak yang merasa terancam karena tindakan pembajakan harus segera melaporkan hal tersebut kepada pihak berwajib	bajak	ambil alih secara paksa, disertai ancaman (tentang pesawat dan sebagainya)	ambil alih secara paksa, disertai ancaman (tentang pesawat dan sebagainya)	Sesuai
64	Petani harus membajak sawah sebelum menanam padi	bajak	perkakas pertanian yang terbuat dari kayu atau besi untuk menggemburkan dan membalikkan tanah; luku; tenggala	perkakas pertanian yang terbuat dari kayu atau besi untuk menggemburkan dan membalikkan tanah; luku; tenggala	Sesuai
65	Saya telah berbicara dengan ahli waris keluarga itu tentang penyelesaian warisan	ahli	keluarga atau kaum	keluarga atau kaum	Sesuai
66	dia seorang yang ahli menjalankan mesin itu	ahli	mahir benar; sangat mendalam dalam suatu ilmu	mahir benar; sangat mendalam dalam suatu ilmu	Sesuai
67	tiang itu rusak dimakan	bubuk	binatang kecil yang berkulit	binatang kecil yang berkulit	Sesuai

	bubuk		keras, berupa kumpang yang suka makan kayu, balok, beras, dan sebagainya	keras, berupa kumpang yang suka makan kayu, balok, beras, dan sebagainya	
68	bubuk kopi tersebut adalah yang paling hitam dan halus	bubuk	barang yang ditumbuk menjadi tepung halus	barang yang ditumbuk menjadi tepung halus	Sesuai
69	sabun yang baik banyak mengandung busa	busa	gelembung-gelembung putih kecil-kecil (seperti pada sabun dan sebagainya); buih	gelembung-gelembung putih kecil-kecil (seperti pada sabun dan sebagainya); buih	Sesuai
70	Pilihan kasur dengan lapisan busa memastikan tidur Anda lebih empuk dan nyaman	busa	karet gembus dan empuk, biasanya untuk isi jok atau kasur	karet gembus dan empuk, biasanya untuk isi jok atau kasur	Sesuai
71	Dasar bodoh, tidakkah kau mengerti betapa berbahayanya tindakanmu?	dasar	memang	pokok atau pangkal suatu pendapat (ajaran, aturan); asas	Tidak Sesuai
72	Buah apel batu sangat segar jika dijus	apel	buah apel	naik banding kepada pengadilan yang lebih tinggi; permohonan pemeriksaan	Tidak Sesuai

				ulang pada pengadilan tingkat kedua (pengadilan tinggi) terhadap keputusan pengadilan tingkat pertama (pengadilan negeri)	
73	Apel kebangsaan pagi di lapangan tersebut dihadiri seluruh pns kota malang	apel	upacara	wajib hadir dalam suatu upacara resmi (bersifat kemiliteran) untuk diketahui hadir tidaknya atau untuk mendengar amanat	Tidak Sesuai
74	Tahu asli kediri sangat gemar dijadikan oleh oleh khas	tahu	makanan dari kedelai putih yang digiling halus-halus, direbus, dan dicetak	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	Tidak Sesuai
75	Dia tahu sifat asli temannya yang dikira jahat tersebut	tahu	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	mengerti sesudah melihat (menyaksikan, mengalami, dan sebagainya)	Sesuai
76	toilet untuk kaum hawa pada tahap perbaikan	hawa	(ditulis dengan huruf kecil) perempuan	(ditulis dengan huruf kecil) perempuan	Sesuai
77	hawa pada pagi hari ini	hawa	keadaan udara (bukan suhu,	keadaan udara (bukan suhu,	Sesuai



	terasa sejuk sekali		cuaca, dan lain-lain)	cuaca, dan lain-lain)	
78	gunakan jangka sebagai alat bantu untuk membuat lingkaran	jangka	alat untuk membuat bulatan (lingkaran, mengukur jarak pada peta, dan sebagainya), berupa benda yang berkaki dua yang dapat dilebarkan dan disempitkan langkahnya atau ukurannya	alat untuk membuat bulatan (lingkaran, mengukur jarak pada peta, dan sebagainya), berupa benda yang berkaki dua yang dapat dilebarkan dan disempitkan langkahnya atau ukurannya	Sesuai
79	menuju teras rumah hanya butuh beberapa langkah dari pintu	teras	tanah atau lantai yang agak tinggi di depan rumah	tanah atau lantai yang agak tinggi di depan rumah	Sesuai
80	udara malam ini tidak begitu dingin	malam	waktu setelah matahari terbenam hingga matahari terbit	waktu setelah matahari terbenam hingga matahari terbit	Sesuai
81	bunga di pagi hari dibasahi oleh embun pagi	bunga	bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang	bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang	Sesuai
82	dia rela dimadu oleh	madu	istri sah yang lain dari	istri sah yang lain dari	Sesuai

	suaminya		seorang suami berdasarkan pandangan istri pertamanya	seorang suami berdasarkan pandangan istri pertamanya	
83	setiap balita punya hak untuk imunisasi	hak	kewenangan	benar	Tidak Sesuai
84	jangan menggunakan sepatu hak tinggi dalam prosesi wisuda	hak	telapak sepatu pada bagian tumit yang relatif tinggi	telapak sepatu pada bagian tumit yang relatif tinggi	Sesuai
85	martabak asal kota malang memiliki cita rasa yang nikmat	asal	keadaan (tempat, wujud, rupa, dan sebagainya) yang semula; pangkal permulaan	keadaan (tempat, wujud, rupa, dan sebagainya) yang semula; pangkal permulaan	Sesuai
86	babak pertama kontes take me out dimenangkan oleh effendy	babak	bagian dari suatu keseluruhan proses, kejadian, atau peristiwa	bagian besar dalam suatu drama atau lakon (terdiri atas beberapa adegan)	Tidak Sesuai
87	risqi aris babak belur dihajar oleh tetangganya	babak	lecet (tentang kulit)	bagian besar dalam suatu drama atau lakon (terdiri atas beberapa adegan)	Tidak Sesuai
88	Selai rasa apel memang paling banyak disukai	apel	buah apel	naik banding kepada pengadilan yang lebih tinggi;	Tidak Sesuai

	anak-anak			permohonan pemeriksaan ulang pada pengadilan tingkat kedua (pengadilan tinggi) terhadap keputusan pengadilan tingkat pertama (pengadilan negeri)	
89	Tahu makanan kesukaanku	tahu	makanan dari kedelai putih yang digiling halus-halus, direbus, dan dicetak	makanan dari kedelai putih yang digiling halus-halus, direbus, dan dicetak	Sesuai
90	Istri mana yang mau dimadu	madu	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	Sesuai
91	Setiap tahun, ani menerima bunga 5 persen dari bank	bunga	imbalan jasa untuk penggunaan uang atau modal yang dibayar pada waktu tertentu berdasarkan ketentuan atau kesepakatan, umumnya dinyatakan sebagai persentase dari modal pokok	bagian tumbuhan yang akan menjadi buah, biasanya elok warnanya dan harum baunya; kembang	Tidak Sesuai

92	Bunga mawar di taman tampak layu	bunga	jenis untuk berbagai-bagai bunga	jenis untuk berbagai-bagai bunga	Sesuai
93	Madu mempunyai banyak manfaat untuk kesehatan tubuh	madu	cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)	cairan yang banyak mengandung zat gula pada sarang lebah atau bunga (rasanya manis)	Sesuai
94	Umumnya semua istri tidak akan pernah mau dimadu	madu	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	istri sah yang lain dari seorang suami berdasarkan pandangan istri pertamanya	Sesuai
95	Saat kau ingin rumahmu terang alami, pakailah genting kaca di bagian tertentu atap rumahmu	genting	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	tutup atap rumah yang terbuat dari tanah liat yang dicetak dan dibakar, bermacam-macam bentuknya	Sesuai
96	Keadaan bertambah genting mendengar kabar ani diculik	genting	tegang; berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)	tegang; berbahaya (tentang keadaan yang mungkin segera menimbulkan bencana perang dan sebagainya)	Sesuai
97	Langkah yang diambil harus dapat	langkah	sikap; tindak-tanduk; perbuatan	sikap; tindak-tanduk; perbuatan	Sesuai

	dipertanggungjawabkan dan sesuai dengan aturan yang berlaku				
98	Seiring bertambahnya umur, langkah kaki seseorang dapat menjadi semakin lambat ketika berjalan	langkah	gerakan kaki (ke depan, ke belakang, ke kiri, ke kanan) waktu berjalan	gerakan kaki (ke depan, ke belakang, ke kiri, ke kanan) waktu berjalan	Sesuai
99	Halaman yang indah menjadi idaman setiap orang yang punya rumah	halaman	pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)	pekarangan rumah (sekolah dan sebagainya); tanah di sekitar rumah (sekolah dan sebagainya)	Sesuai
100	Bagian HRM mau mengadakan rapat nanti sore untuk membicarakan penghargaan atas kinerja karyawan	rapat	pertemuan (kumpulan) untuk membicarakan sesuatu; sidang; majelis	pertemuan (kumpulan) untuk membicarakan sesuatu; sidang; majelis	Sesuai