

**PENGEMBANGAN U-TAPIS UNTUK DETEKSI KESALAHAN
PENGUNAAN KATA KONJUNSI MENGGUNAKAN SEMANTIC
ANALYSIS DAN BERT**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

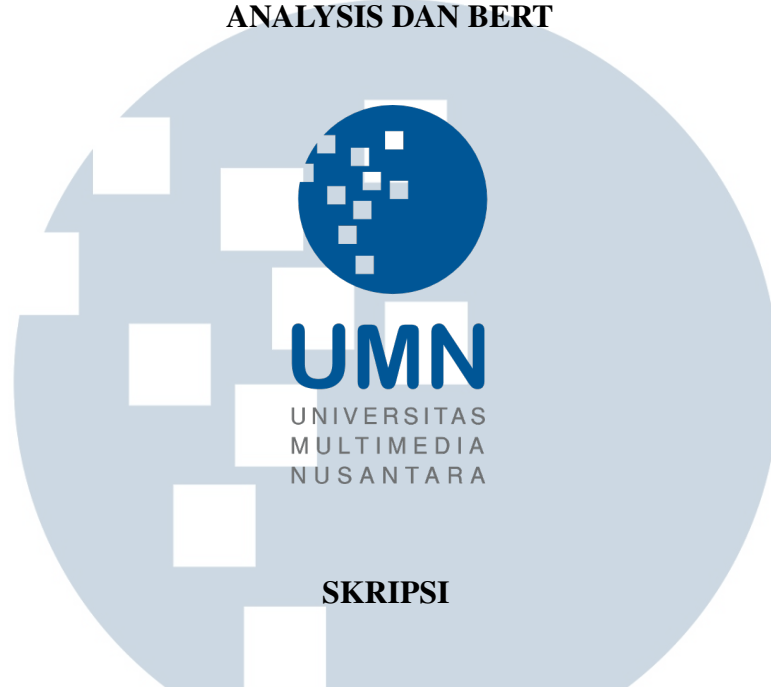
JONATHAN OCTAVIEN

00000044495

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**PENGEMBANGAN U-TAPIS UNTUK DETEKSI KESALAHAN
PENGUNAAN KATA KONJUNGSI MENGGUNAKAN SEMANTIC
ANALYSIS DAN BERT**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

JONATHAN OCTAVIEN

00000044495

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Jonathan Octavien

Nomor Induk Mahasiswa : 00000044495

Program Studi : Informatika

Skripsi dengan judul:

Pengembangan U-Tapis untuk Deteksi Kesalahan Penggunaan Kata Konjungsi Menggunakan *Semantic Analysis* dan *BERT*

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

UMM

Tangerang, 10 Januari 2025



(Jonathan Octavien)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

PENGEMBANGAN U-TAPIS UNTUK DETEKSI KESALAHAN PENGUNAAN KATA KONJUNGSI MENGGUNAKAN SEMANTIC ANALYSIS DAN BERT

oleh

Nama : Jonathan Octavien
NIM : 00000044495
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika


Telah diujikan pada hari Jumat, 10 Januari 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

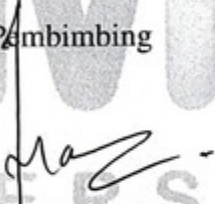
Ketua Sidang


(Dennis Gunawan, S.Kom., M.Sc.)
NIDN: 0320059001

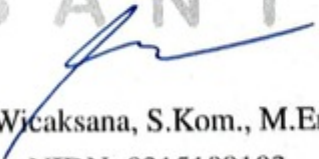
Penguji


(Wirawan Istiono, S.Kom., M.Kom.)
NIDN: 0313048304

Pembimbing


(Marlinda Vasty Overbeek, S.Kom., M.Kom.)
NIDN: 0818038501

Ketua Program Studi Informatika,


(Arya Wicaksana, S.Kom., M.Eng.Sc.)
NIDN: 0315109103

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan dibawah ini:

Nama : Jonathan Octavien
NIM : 00000044495
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : PENGEMBANGAN U-TAPIS UNTUK DETEKSI
KESALAHAN PENGGUNAAN KATA
KONJUNGSI MENGGUNAKAN SEMANTIC
ANALYSIS DAN BERT

Menyatakan dengan sesungguhnya bahwa saya bersedia* (**pilih salah satu**):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 20 Januari 2025


(Jonathan Octavien)

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

”Blessed are those who have not seen and yet have believed.”

John 20:29 (EXB)



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Pengembangan U-Tapis untuk Deteksi Kesalahan Penggunaan Kata Konjungsi Menggunakan Semantic Analysis dan BERT dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., sebagai Pembimbing utama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya Skripsi ini.
5. Orang Tua dan adik saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan Skripsi ini.
6. Teman-teman terdekat saya yang telah memberikan bantuan dukungan moral, sehingga penulis dapat menyelesaikan Skripsi ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 10 Januari 2025



Jonathan Octavien

**PENGEMBANGAN U-TAPIS UNTUK DETEKSI KESALAHAN
PENGUNAAN KATA KONJUNSI MENGGUNAKAN SEMANTIC
ANALYSIS DAN BERT**

Jonathan Octavien

ABSTRAK

Teknologi selalu berkembang dalam berbagai bidang, terutama dalam bidang bahasa. *Natural Language Processing* (NLP) merupakan cara untuk mesin memahami bahasa alami manusia dan menggunakan pemahaman tersebut untuk membantu manusia. U-TAPIS adalah salah satu aplikasi yang dikembangkan untuk mendeteksi kesalahan bahasa pada artikel berita. Salah satu modul pada U-TAPIS adalah pendeteksi kesalahan penggunaan kata konjungsi. Penggunaan kata konjungsi perlu diperhatikan karena penggunaan kata konjungsi yang salah dapat membuat pembaca melewati tujuan utama dari sebuah artikel dan mempengaruhi pembaca dalam memahami isi artikel tersebut. Modul yang dikembangkan dalam penelitian ini menggunakan metode *Semantic Analysis* untuk melakukan deteksi penggunaan kata konjungsi dan model BERT untuk menghasilkan saran perbaikan. Hasil pengujian untuk deteksi penggunaan kata konjungsi mendapatkan nilai *F1-score* untuk 150, 200, dan 250 artikel berita secara berturut-turut sebesar 87,03%, 87,22%, dan 87,54%. Hasil pengujian saran perbaikan dengan benar mendapatkan nilai sebesar 78,00%. Saran untuk penelitian selanjutnya adalah menambahkan aturan deteksi untuk memperluas kata konjungsi yang dapat dideteksi atau mencoba teknik OpenIE untuk membuat sistem OpenIE untuk Bahasa Indonesia serta mencari metode lain untuk saran perbaikan agar dapat memperbaiki struktur kalimat.

Kata kunci: BERT, Deteksi, Kata konjungsi, Kesalahan penggunaan, *Semantic Analysis*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

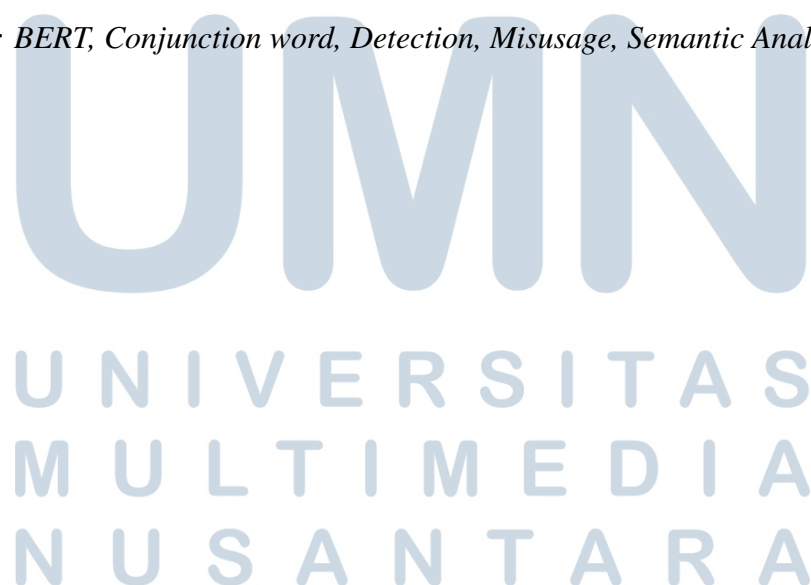
***DEVELOPMENT OF U-TAPIS FOR CONJUNCTION WORD MISUSAGE
DETECTION USING SEMANTIC ANALYSIS AND BERT***

Jonathan Octavien

ABSTRACT

Technology continues to advance in various fields, particularly in language. Natural Language Processing (NLP) is a method for machines to understand human natural language and use that understanding to assist humans. U-TAPIS is an application developed to detect language errors in news articles. One of the modules in U-TAPIS is designed to detect errors in the use of conjunction words. The use of conjunction words must be carefully considered because incorrect use of conjunction words can cause readers to miss the main purpose of an article and affect their understanding of its content. The module developed in this research uses the Semantic Analysis method to detect the use of conjunction words and the BERT model to provide improvement suggestions. Testing results for conjunction usage detection achieved an F1-score of 89.50% for testing on 50 articles and 87.75% for testing on 100 news articles. The testing results for correct improvement suggestions achieved a score of 78.00%. Suggestions for future research include adding detection rules to expand the range of conjunction words that can be detected, exploring OpenIE techniques to develop an OpenIE system for Bahasa Indonesia, and finding alternative methods for improvement suggestions to enhance sentence structure correction.

Keywords: *BERT, Conjunction word, Detection, Misusage, Semantic Analysis*



DAFTAR ISI

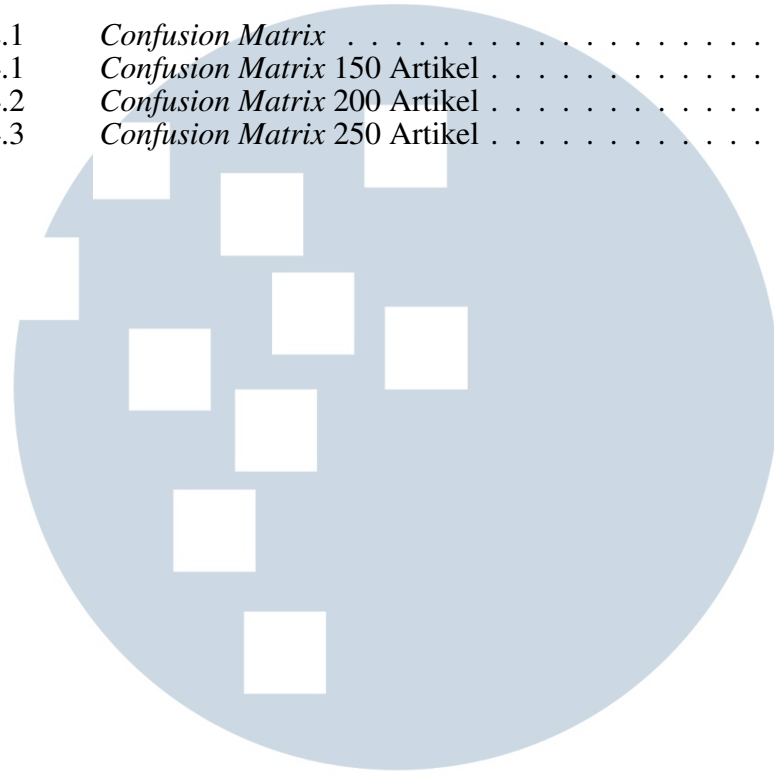
HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	7
2.1 Kata Konjungsi	7
2.1.1 Konjungsi Koordinatif	7
2.1.2 Konjungsi Subordinatif	8
2.1.3 Konjungsi Korelatif	10
2.1.4 Konjungsi Antarkalimat	10
2.2 Natural Language Processing	11
2.2.1 Semantic Analysis	11
2.2.2 Model Transformer	13
2.3 Confusion Matrix	15
BAB 3 METODOLOGI PENELITIAN	16
3.1 Studi Literatur	16
3.2 Pengumpulan dan Praproses Data Artikel	16
3.3 Pembuatan <i>rule</i> penggunaan kata konjungsi	17
3.4 Perancangan algoritma deteksi kesalahan penggunaan kata konjungsi	18
3.5 Implementasi model BERT untuk saran perbaikan kata konjungsi	19
3.6 Perancangan API	20
3.7 Pengujian dan Evaluasi	21
BAB 4 HASIL DAN DISKUSI	22
4.1 Spesifikasi Sistem	22
4.2 Pengumpulan Data dan Praproses Data	22
4.3 Proses Deteksi Penggunaan Kata Konjungsi	24
4.3.1 Penyaringan Kata Konjungsi	24
4.3.2 Deteksi Penggunaan Kata Konjungsi menggunakan <i>Semantic Analysis</i>	25
4.3.3 Menampilkan Hasil Deteksi Kata Konjungsi	27

4.3.4	Menghasilkan Saran Perbaikan Kata menggunakan model BERT	28
4.4	Implementasi Tampilan Antarmuka	29
4.5	Pengujian Sistem	32
4.5.1	<i>Confusion Matrix</i>	33
4.5.2	Saran Perbaikan	34
BAB 5	SIMPULAN DAN SARAN	35
5.1	Simpulan	35
5.2	Saran	35
	DAFTAR PUSTAKA	36



DAFTAR TABEL

Tabel 2.1	<i>Confusion Matrix</i>	15
Tabel 4.1	<i>Confusion Matrix</i> 150 Artikel	33
Tabel 4.2	<i>Confusion Matrix</i> 200 Artikel	33
Tabel 4.3	<i>Confusion Matrix</i> 250 Artikel	34



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 2.1	Alur proses <i>Semantic Analysis</i> secara umum [1]	12
Gambar 2.2	Alur proses <i>Semantic Analysis</i> untuk Bahasa Inggris [2]	13
Gambar 2.3	Model arsitektur Transformer [3]	14
Gambar 2.4	Prosedur <i>pre-training</i> dan <i>fine-tuning</i> untuk BERT [4]	15
Gambar 3.1	Diagram alur metodologi penelitian	16
Gambar 3.2	Daftar aturan penggunaan kata konjungsi	17
Gambar 3.3	Daftar kata konjungsi dan aturan yang harus dipenuhi	18
Gambar 3.4	Diagram alur algoritma deteksi kesalahan penggunaan kata konjungsi	19
Gambar 3.5	Diagram alur algoritma saran perbaikan kata konjungsi	20
Gambar 3.6	Diagram alur API deteksi penggunaan kata konjungsi	21
Gambar 4.1	Halaman awal	29
Gambar 4.2	Menyalin URL API	30
Gambar 4.3	Menyalin cURL API	30
Gambar 4.4	Tampilan awal untuk halaman deteksi penggunaan kata konjungsi	30
Gambar 4.5	Menaruh berita dan tekan tombol "Periksa"	31
Gambar 4.6	Tampilan setelah menekan tombol "Periksa"	32
Gambar 4.7	Tampilan setelah menekan tombol "Perbaiki" dan "Copy"	32



DAFTAR KODE

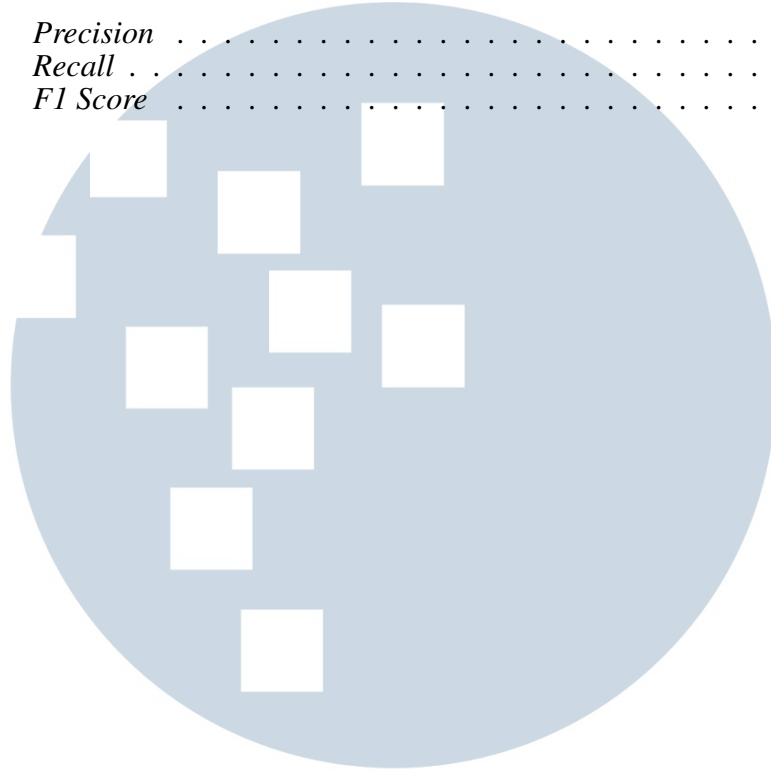
Kode 4.1	Proses Tokenisasi Artikel	23
Kode 4.2	Penyaringan Kata Konjungsi	24
Kode 4.3	Proses Pemeriksaan Kalimat dengan Rule Konjungsi	25
Kode 4.4	Menampilkan Hasil Deteksi Kata Konjungsi	27
Kode 4.5	Proses Pembuatan Saran Perbaikan Menggunakan BERT	28



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

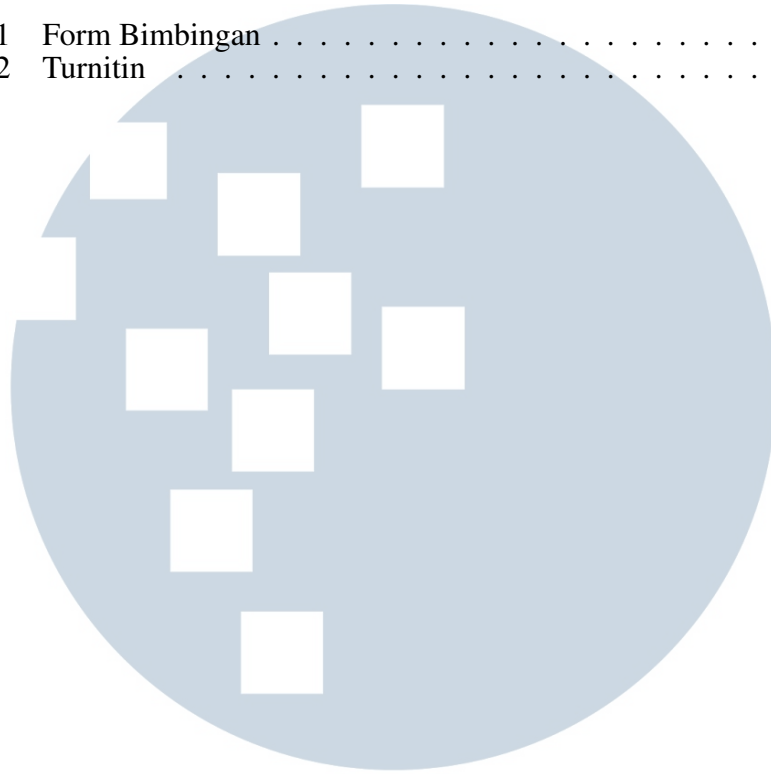
Rumus 2.1	<i>Precision</i>	15
Rumus 2.2	<i>Recall</i>	15
Rumus 2.3	<i>F1 Score</i>	15



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	39
Lampiran 2	Turnitin	42



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA