



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

**ANALISIS SENTIMEN TWITTER TERHADAP VAKSINASI
COVID-19 DOSIS KE-4 DENGAN ALGORITMA
LONG SHORT-TERM MEMORY**



SKRIPSI

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

Grivia Trifosa Iskandar

00000036794

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

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

TANGERANG

2023

**ANALISIS SENTIMEN TWITTER TERHADAP VAKSINASI
COVID-19 DOSIS KE-4 DENGAN ALGORITMA
LONG SHORT-TERM MEMORY**



SKRIPSI

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

Grivia Trifosa Iskandar

0000036794

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

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

TANGERANG

2023

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Grivia Trifosa Iskandar
Nomor Induk Mahasiswa : 00000036794
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Twitter terhadap Vaksinasi Covid-19 Dosis Ke-4 dengan Algoritma Long Short-Term Memory

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, 6 Juni 2023



UNIVERSITAS
MULTIMEDIA
NUSANTARA

(Grivia Trifosa Iskandar)

HALAMAN PENGESAHAN

Skripsi dengan judul

**ANALISIS SENTIMEN TWITTER TERHADAP VAKSINASI
COVID-19 DOSIS KE-4 DENGAN ALGORITMA
LONG SHORT-TERM MEMORY**

oleh

Nama : Grivia Trifosa Iskandar
NIM : 00000036794
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Jumat, 16 Juni 2023

Pukul 15.00 s/d 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

(Marlinda Vasty Overbeek, S.Kom.,
M.Kom.)

NIDN: 0818038501

Penguji

(Adhi Kusnadi, S.T, M.Si.)

NIDN: 0303037304

Pembimbing

(Alethea Suryadibrata, S.Kom., M.Eng.)

NIDN: 0322099201

Ketua Program Studi Informatika,

(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Grivia Trifosa Iskandar
NIM : 00000036794
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN TWITTER TERHADAP VAKSINASI
COVID-19 DOSIS KE-4 DENGAN ALGORITMA
LONG SHORT-TERM MEMORY**

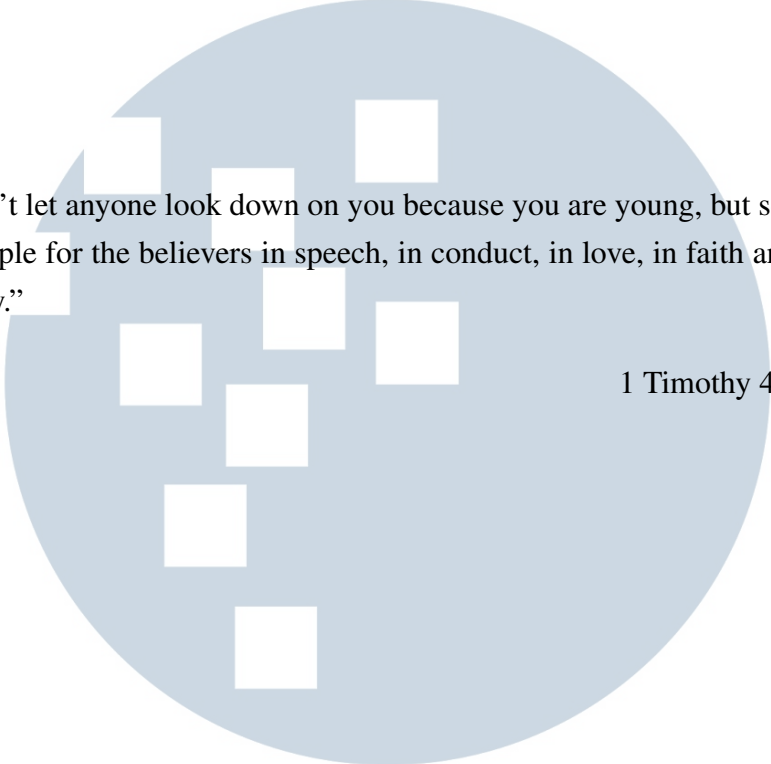
Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 6 Juni 2023
Yang menyatakan

UNIVERSITAS
MULTIMEDIA
NUSANTARA


Grivia Trifosa Iskandar

Halaman Persembahan / Motto



”Don’t let anyone look down on you because you are young, but set an example for the believers in speech, in conduct, in love, in faith and in purity.”

1 Timothy 4:12 (NIV)

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: Analisis Sentimen Twitter terhadap Vaksinasi Covid-19 Dosis Ke-4 dengan Algoritma Long Short-Term Memory 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. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Alethea Suryadibrata, S.Kom., M.Eng., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya skripsi ini.
5. Orang Tua, kakak, dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan skripsi ini.
6. Anas, Cebe, Lala, Iren, Ivana, Michelle, Olin, dan Ferin selaku teman terdekat yang telah memberikan dukungan selama masa perkuliahan dan juga selama masa penyusunan laporan skripsi.
7. Angel, Audric, Cinhya, dan Joel beserta seluruh teman seperjuangan, yang telah memotivasi dan membantu penulis selama penyusunan laporan skripsi dan keseluruhan aktivitas perkuliahan.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, karena adanya keterbatasan ilmu serta pengalaman yang dimiliki. Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 6 Juni 2023



Grivia Trifosa Iskandar



**ANALISIS SENTIMEN TWITTER TERHADAP VAKSINASI
COVID-19 DOSIS KE-4 DENGAN ALGORITMA
LONG SHORT-TERM MEMORY**

Grivia Trifosa Iskandar

ABSTRAK

Penyebaran virus Covid-19 telah menyebar di seluruh dunia termasuk di Indonesia, membuat masyarakat gempar dan mendorong pemerintah untuk melakukan vaksinasi massal. Di Indonesia, program vaksinasi Covid-19 telah diberlakukan sejak tahun 2021 dan saat ini mulai memasuki dosis ke-4. Sejak Agustus 2022, para petugas kesehatan di Indonesia sudah menerima vaksin tersebut untuk meningkatkan antibodi dari banyaknya varian virus Covid-19, salah satunya yaitu varian Omicron. Wacana vaksin Covid-19 dosis ke-4 untuk masyarakat umum menimbulkan respon dari berbagai kalangan, dan salah satu media untuk mengekspresikan respon tersebut adalah media sosial Twitter. Untuk itu, penelitian ini dilakukan untuk membuat suatu model yang dapat melakukan klasifikasi sentimen dari berbagai opini yang ditemukan di Twitter tentang dosis ke-4 vaksinasi Covid-19 di Indonesia yang terdiri dari sentimen positif, negatif, dan netral. Model klasifikasi yang akan digunakan adalah *deep learning* dengan menggunakan metode *Long Short-Term Memory* (LSTM) dengan bantuan *lexicon* InSet sebagai kamus untuk membantu proses pelabelan data dan teknik *resampling* sebagai teknik dalam mengatasi ketidakseimbangan dataset. Dari hasil penelitian, didapatkan performa model yang dapat melakukan klasifikasi terbaik dilakukan dengan dataset yang menggunakan teknik *oversampling* serta menggunakan 1 *layer* LSTM dan epoch sebanyak 10, dengan akurasi sebesar 86%, *precision* 86%, *recall* 86%, dan *F1-Score* 86%.

Kata kunci: *Analisis Sentimen, Deep Learning, LSTM, Twitter, Vaksinasi Covid-19*

UNIVERSITAS
MULTIMEDIA
NUSANTARA

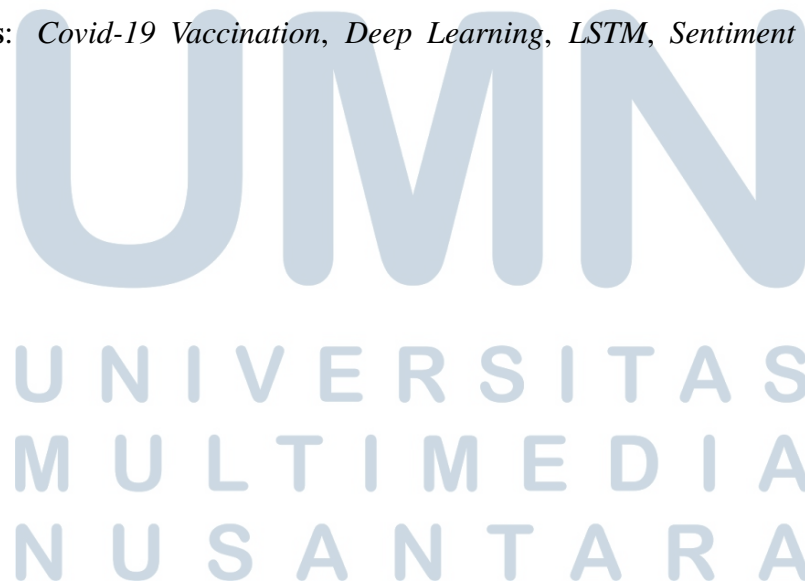
Sentiment Analysis of Twitter of the 4th Dose of Covid-19 Vaccination using Long Short-Term Memory Algorithm

Grivia Trifosa Iskandar

ABSTRACT

The spread of Covid-19 Virus has spread all around the world, including Indonesia, making people uproar and encourage the government to carry out mass vaccinations. In Indonesia, Covid-19 vaccination program has began in 2021 and now has entered it's fourth dose. Since August 2022, the health workers in Indonesia have received this vaccine to boost up the antibody from many Covid-19 virus varians, one of which is the Omicron. The discourse on the 4th dose of the Covid-19 vaccine has spark responses from various groups, and one of the medium to express those responses is Twitter. For that, this research was conducted to create a model that could do sentiment classification from various opinions found in Twitter about the fourth dose of Covid-19 vaccination in Indonesia which consists of positive, negative, and netral label. The classification model used is deep learning with Long Short-Term Memory (LSTM) method, with Lexicon InSet as the dictionary for labelling process and resampling method as the technique for handling the imbalanced dataset. From the research results, it was found that the performance of the model that could perform the best classification was done with the dataset that use oversampling method, with 1 LSTM layer and 10 epoch, with an accuracy of 86%, precision 86%, recall 86%, and F1-Score 86%.

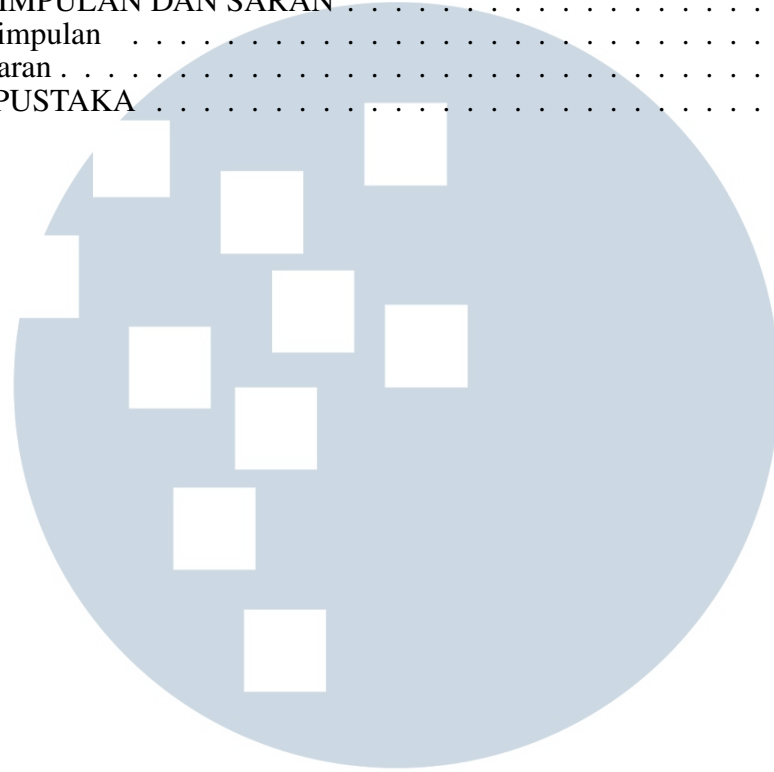
Keywords: *Covid-19 Vaccination, Deep Learning, LSTM, Sentiment Analysis, Twitter,*



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	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Vaksin Covid-19	5
2.2 Twitter	5
2.3 Analisis Sentimen	6
2.4 <i>Recurrent Neural Network</i>	6
2.5 <i>Long Short-Term Memory</i>	8
2.6 <i>Text Pre-Processing</i>	11
2.7 <i>Confusion Matrix</i>	11
BAB 3 METODOLOGI PENELITIAN	13
3.1 Metodologi Penelitian	13
3.1.1 Studi Literatur	13
3.1.2 Pengumpulan Data	13
3.1.3 Pemrograman Sistem	13
3.1.4 Pengujian dan Evaluasi	13
3.1.5 Penulisan Laporan	14
3.2 Perancangan Sistem	14
3.2.1 Gambaran Umum Sistem	14
3.2.2 Pengumpulan Data	15
3.2.3 <i>Text Preprocessing</i>	16
3.2.4 Pelabelan Data	17
3.2.5 Pembuatan Model	18
3.2.6 Evaluasi Model	18
BAB 4 HASIL DAN DISKUSI	19
4.1 Spesifikasi Sistem	19
4.2 Implementasi Sistem	19
4.2.1 Pengambilan Data	19
4.2.2 Pre-processing Dataset	21
4.2.3 Pembagian Data	32
4.2.4 Pembuatan Model	32

4.2.5	Evaluasi Model	35
4.3	Skenario Pengujian dan Evaluasi	35
BAB 5	SIMPULAN DAN SARAN	39
5.1	Simpulan	39
5.2	Saran	39
	DAFTAR PUSTAKA	40



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 2.1	Arsitektur RNN	6
Gambar 2.2	Penjabaran Arsitektur RNN	7
Gambar 2.3	Alur Proses RNN	7
Gambar 2.4	<i>Output</i> Alur Proses RNN	8
Gambar 2.5	Arsitektur <i>Long Short-Term Memory</i>	10
Gambar 3.1	<i>Flowchart</i> Gambaran Umum Sistem	14
Gambar 3.2	<i>Flowchart</i> Text <i>Preprocessing</i>	16
Gambar 3.3	<i>Flowchart</i> Pelabelan Data	17
Gambar 3.4	<i>Flowchart</i> Pembuatan Model	18
Gambar 4.1	Daftar <i>library</i> untuk <i>scrap</i> data	19
Gambar 4.2	Kode <i>Scraping</i> Data Berdasarkan <i>Keyword</i> Menggunakan Snsrape	20
Gambar 4.3	Kode <i>Scraping</i> Data Berdasarkan <i>Tweet</i> Menggunakan Snsrape	20
Gambar 4.4	Hasil Pengambilan Data dengan Menggunakan Snsrape	21
Gambar 4.5	Implementasi Kode untuk <i>Preprocessing</i>	22
Gambar 4.6	Hasil dari Noise Removal	22
Gambar 4.7	Kode untuk Mengubah Kata Slang	23
Gambar 4.8	Hasil dari Mengubah Kata Slang	23
Gambar 4.9	Implementasi Kode untuk Proses <i>Stopwords & Tokenization</i>	24
Gambar 4.10	Hasil <i>Stopwords & Tokenizing</i>	24
Gambar 4.11	Implementasi Kode untuk Proses <i>Stemming</i>	25
Gambar 4.12	Hasil <i>Stemming</i>	25
Gambar 4.13	Penghapusan Data	26
Gambar 4.14	Implementasi Kode untuk <i>Labelling</i> Data	27
Gambar 4.15	Hasil Data Setelah Proses <i>Labeling</i>	28
Gambar 4.16	Total Data Setelah Proses <i>Labeling</i>	28
Gambar 4.17	<i>Word Cloud</i> Sentimen Positif	28
Gambar 4.18	<i>Word Cloud</i> Sentimen Netral	29
Gambar 4.19	<i>Word Cloud</i> Sentimen Negatif	29
Gambar 4.20	Kode Proses <i>Oversampling</i>	30
Gambar 4.21	Total Setiap Kelas Sentimen setelah <i>Oversampling</i>	30
Gambar 4.22	Kode Proses <i>Undersampling</i>	31
Gambar 4.23	Total Setiap Kelas Sentimen setelah <i>Undersampling</i>	31
Gambar 4.24	Proses <i>Encoding dan Padding</i>	32
Gambar 4.25	Kode untuk Memisahkan Data <i>Train</i> , Data <i>Test</i> , dan Data <i>Validation</i>	32
Gambar 4.26	Jaringan Arsitektur dengan 1 <i>Layer LSTM</i>	33
Gambar 4.27	Jaringan Arsitektur dengan 2 <i>Layer LSTM</i>	34
Gambar 4.28	Pelatihan Model LSTM	35
Gambar 4.29	Perhitungan <i>Accuracy, Precision, Recall</i> , dan <i>F1-Score</i>	35
Gambar 4.30	Kode untuk Menampilkan <i>Confusion Matrix</i>	35
Gambar 4.31	<i>Confusion Matrix</i> Model Terbaik pada Data Awal	36
Gambar 4.32	<i>Confusion Matrix</i> Model Terbaik pada Data <i>Undersampling</i>	37
Gambar 4.33	<i>Confusion Matrix</i> Model Terbaik pada Data <i>Oversampling</i>	38

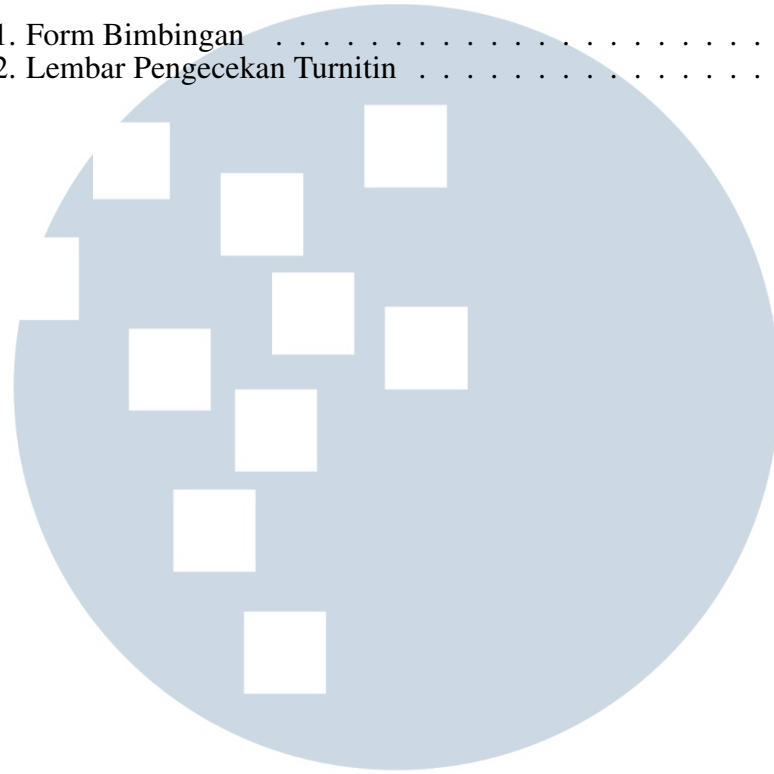
DAFTAR TABEL

Tabel 2.1	Tabel <i>Confusion Matrix</i>	11
Tabel 4.1	Tabel Nilai <i>Precision, Recall, F1-Score, dan Accuracy</i> menggunakan Data Awal	36
Tabel 4.2	Tabel Nilai <i>Precision, Recall, F1-Score, dan Accuracy</i> menggunakan Data <i>Undersampling</i>	37
Tabel 4.3	Tabel Nilai <i>Precision, Recall, F1-Score, dan Accuracy</i> menggunakan Data <i>Oversampling</i>	38



DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	42
Lampiran 2. Lembar Pengecekan Turnitin	44



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA