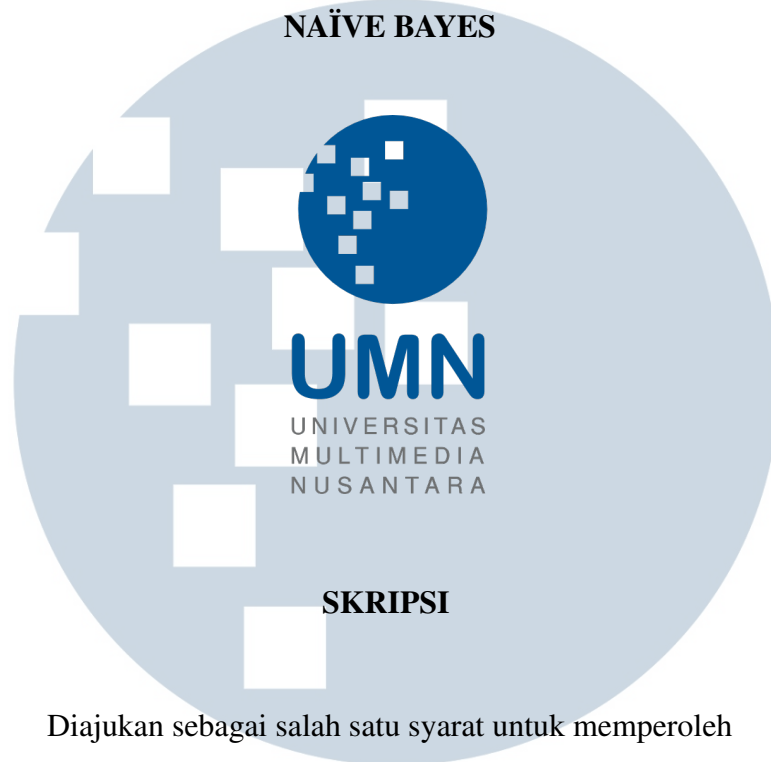


**ANALISIS SENTIMEN PENGGUNA TWITTER DI INDONESIA  
TERHADAP KENAIKAN HARGA BBM MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**



**SKRIPSI**

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

**Joel Alfa Nugraha**

**00000036934**

**UMN**

**UNIVERSITAS**

**MULTIMEDIA**

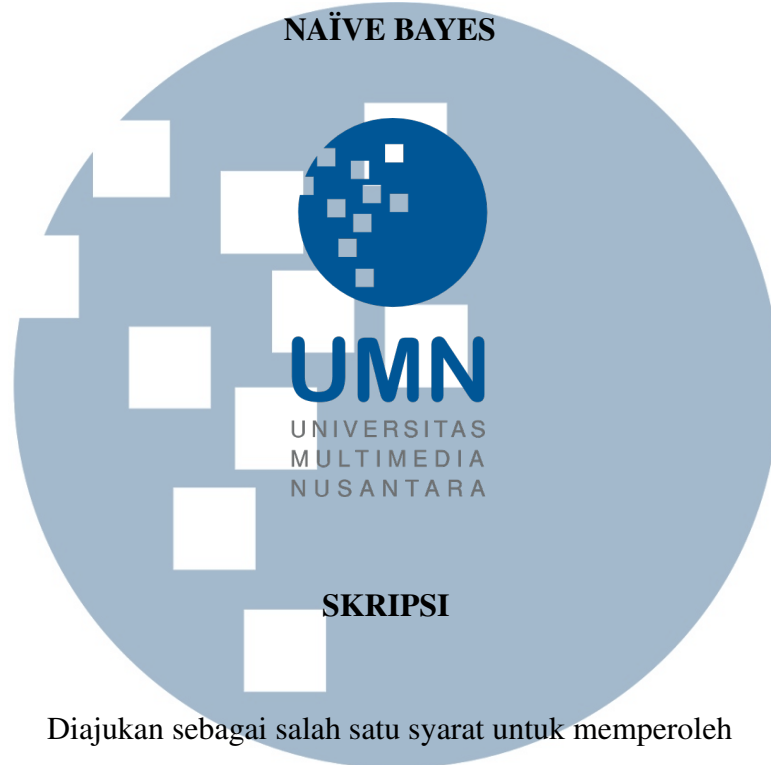
**NUSANTARA**

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

**TANGERANG**

**2023**

**ANALISIS SENTIMEN PENGGUNA TWITTER DI INDONESIA  
TERHADAP KENAIKAN HARGA BBM MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**



**SKRIPSI**

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

**Joel Alfa Nugraha**

**0000036934**

**UMN**

**UNIVERSITAS**

**MULTIMEDIA**

**NUSANTARA**

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

**TANGERANG**

**2023**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Joel Alfa Nugraha

Nomor Induk Mahasiswa : 00000036934

Program Studi : Informatika

Skripsi dengan judul:

**Analisis Sentimen Pengguna Twitter di Indonesia Terhadap Kenaikan Harga BBM Menggunakan Algoritma Naïve Bayes**

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.

Tangerang, 16 Juni 2023



16/6/2023  
(Joel Alfa Nugraha)

UMM  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## HALAMAN PENGESAHAN

Skripsi dengan judul  
**ANALISIS SENTIMEN PENGGUNA TWITTER DI INDONESIA  
TERHADAP KENAIKAN HARGA BBM MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**

oleh  
Nama : Joel Alfa Nugraha  
NIM : 00000036934  
Program Studi : Informatika  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 27 Juni 2023  
Pukul 15.00 s/d 16.01 dan dinyatakan  
**LULUS**

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Dr. Maria Irmina Prasetiyowati, S.Kom.,  
M.T.)  
NIDN: 0725057201

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 : Joel Alfa Nugraha  
NIM : 00000036934  
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 PENGGUNA TWITTER DI INDONESIA  
TERHADAP KENAIKAN HARGA BBM MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**

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, 16 Juni 2023  
Yang menyatakan



Joel Alfa Nugraha

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**Halaman Persembahan / Motto**

”Perjalanan seribu mil dimulai dengan satu langkah.”

Lao Tzu



**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 Pengguna Twitter di Indonesia Terhadap Kenaikan Harga BBM Menggunakan Algoritma Naïve Bayes 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. Bapak 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 pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Orang Tua, teman dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

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

Tangerang, 16 Juni 2023



Joel Alfa Nugraha

**ANALISIS SENTIMEN PENGGUNA TWITTER DI INDONESIA  
TERHADAP KENAIKAN HARGA BBM MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**

Joel Alfa Nugraha

**ABSTRAK**

Bahan bakar minyak merupakan salah satu sektor energi yang penting dalam perekonomian nasional. Namun pada 3 September 2022, harga bahan bakar minyak mengalami kenaikan di Indonesia. Hal ini mengundang banyak opini dan unggahan yang disalurkan ke platform-platform media sosial, salah satunya Twitter. Oleh karena itu, dilakukan penelitian yang bertujuan untuk melakukan klasifikasi sentimen dari opini masyarakat yang menggunakan Twitter sebagai media sosial terhadap kenaikan harga bahan bakar minyak. Untuk mencapai tingkat akurasi tertinggi, sentimen dibagi menjadi dua, yaitu sentimen positif dan sentimen negatif. Model yang digunakan untuk melakukan klasifikasi merupakan model dengan algoritma Naïve Bayes dengan Lexicon inset untuk melabeli sentimen data dan TF-IDF Vectorizer untuk mengukur relevansi kata. Dari penelitian ini didapatkan hasil klasifikasi terbaik pada model Complement Naïve Bayes yang memperoleh total akurasi sebesar 82% dengan *precision* sebesar 89%, *recall* sebesar 88%, dan f-1 score sebesar 89%.

**Kata kunci:** *Machine Learning, Naïve Bayes, Analisis sentimen, Python, Kenaikan Harga bbm*





# Sentiment Analysis of Twitter Users in Indonesia About Fuel Price Increases Using the Naïve Bayes Algorithm

Joel Alfa Nugraha

## ABSTRACT

Fuels are one of the important energy sectors in the national economy. However, the prices have increased from September 3, 2022 in Indonesia. This matter invites opinions and posts that channeled through social media, including Twitter. Therefore, this research has been conducted for doing sentiment classifications from people's opinions who have been using Twitter as their social media about the fuel price increases. The sentiments are divided into two parts: positive and negative. The models that have been used for the classification are models with the Naïve Bayes algorithm with Lexicon inset for labeling data sentiment and TF-IDF Vectorizer to determine the relevance of those words in the data. The best data classifications has been obtained by the Complement Naïve Bayes model with the accuracy total of 82%, precicion of 89%, recall of 88%, and f-1 score of 89%.

**Keywords:** *Machine Learning, Naïve Bayes, Sentiment analysis, Python, Fuel price increases*



## 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 GAMBAR . . . . .	xi
DAFTAR TABEL . . . . .	xii
DAFTAR LAMPIRAN . . . . .	xiii
BAB 1 PENDAHULUAN . . . . .	1
1.1 Latar Belakang Masalah . . . . .	1
1.2 Rumusan Masalah . . . . .	2
1.3 Batasan Masalah . . . . .	2
1.4 Tujuan Penelitian . . . . .	3
1.5 Manfaat Penelitian . . . . .	3
1.6 Sistematika Penulisan . . . . .	3
BAB 2 LANDASAN TEORI . . . . .	5
2.1 Kenaikan Harga BBM . . . . .	5
2.2 Twitter . . . . .	6
2.3 Analisis Sentimen . . . . .	6
2.4 Text Mining . . . . .	6
2.5 Text Pre-processing . . . . .	6
2.6 Klasifikasi . . . . .	8
2.7 Algoritma Naïve Bayes . . . . .	8
2.8 Data Splitting . . . . .	9
2.9 Confusion Matrix . . . . .	9
BAB 3 METODOLOGI PENELITIAN . . . . .	11
3.1 Alur Metodologi Penelitian . . . . .	11
3.1.1 Studi Literatur . . . . .	11
3.1.2 Pengumpulan Data . . . . .	12
3.1.3 Perancangan Sistem . . . . .	12
3.1.4 Pengujian dan Evaluasi . . . . .	12
3.2 Rancangan Sistem . . . . .	12
3.2.1 Text Mining . . . . .	13
3.2.2 Text Pre-Processing . . . . .	13
3.2.3 TF-IDF . . . . .	15
3.2.4 Data Splitting . . . . .	15
3.2.5 Penerapan Model Multinomial Naïve Bayes dan Complement Naïve Bayes . . . . .	15
3.2.6 Pengujian Model . . . . .	16
3.2.7 Evaluasi . . . . .	17
BAB 4 HASIL DAN DISKUSI . . . . .	18
4.1 Spesifikasi Sistem . . . . .	18
4.2 Implementasi Sistem . . . . .	18
4.2.1 Text Mining . . . . .	18

4.2.2	Text Pre-Processing . . . . .	19
4.2.3	TF-IDF . . . . .	29
4.2.4	Data Splitting . . . . .	30
4.2.5	Multinomial Naïve Bayes . . . . .	30
4.2.6	Complement Naïve Bayes . . . . .	31
4.3	Uji Coba dan Evaluasi . . . . .	32
4.3.1	Uji Coba . . . . .	32
4.3.2	Evaluasi . . . . .	35
BAB 5	SIMPULAN DAN SARAN . . . . .	40
5.1	Simpulan . . . . .	40
5.2	Saran . . . . .	40
	DAFTAR PUSTAKA . . . . .	41

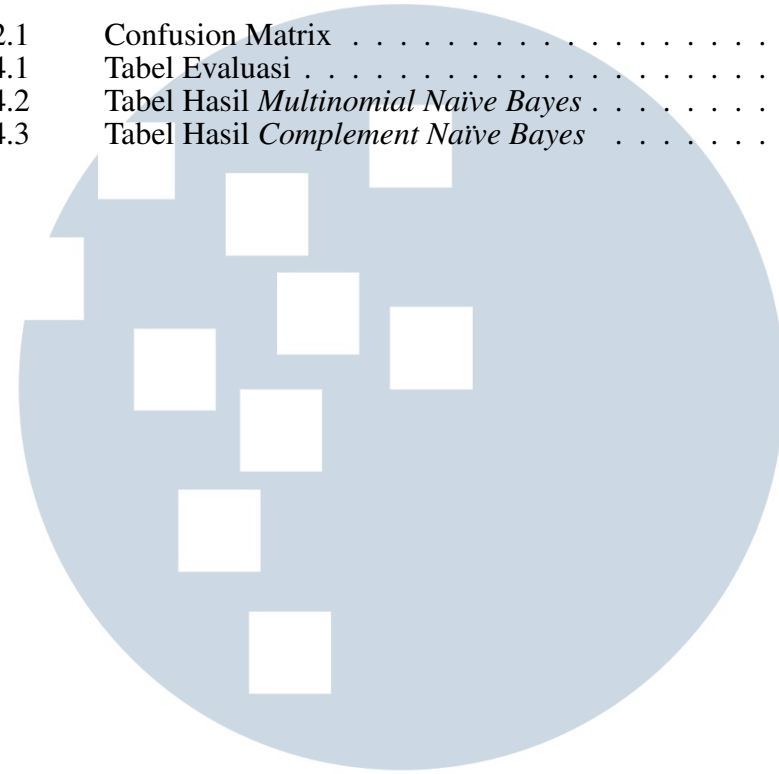


## DAFTAR GAMBAR

Gambar 2.1	Grafik Kenaikan Harga BBM PT Pertamina(Persero) . . . .	5
Gambar 2.2	Diagram <i>Text Pre-Processing</i> . . . . .	7
Gambar 3.1	Diagram alir metodologi penelitian . . . . .	11
Gambar 3.2	Diagram Alir Rancangan Sistem . . . . .	13
Gambar 3.3	Diagram Alir <i>Text Pre-Processing</i> . . . . .	14
Gambar 3.4	Diagram Alir Cara Kerja Model <i>Naïve Bayes</i> . . . . .	16
Gambar 4.1	Potongan Kode Text Mining . . . . .	19
Gambar 4.2	Potongan Kode Import Data . . . . .	19
Gambar 4.3	Hasil dari Kode pada Gambar 4.2 . . . . .	19
Gambar 4.4	Potongan Kode untuk Menghapus Data Duplikat . . . . .	19
Gambar 4.5	Potongan Kode untuk Memeriksa jika Ada Data yang Bersifat <i>NaN</i> . . . . .	20
Gambar 4.6	Potongan Kode untuk Menghapus Data Duplikat dan Memaparkan Hasilnya . . . . .	20
Gambar 4.7	Hasil dari Kode Pada Gambar 4.6 . . . . .	20
Gambar 4.8	Potongan Kode untuk Menghilangkan <i>URL</i> . . . . .	21
Gambar 4.9	Potongan Kode untuk <i>Define Functions</i> yang Dieksekusi pada Tahap Selanjutnya . . . . .	23
Gambar 4.10	Potongan Kode Perbaikan Kata "BBM" . . . . .	23
Gambar 4.11	Hasil dari Potongan Kode Gambar 4.10 . . . . .	23
Gambar 4.12	Potongan Kode Eksekusi Function <i>clean_tweets</i> . . . . .	24
Gambar 4.13	Hasil dari Potongan Kode Pada Gambar 4.12 . . . . .	24
Gambar 4.14	Potongan Kode Deteksi Bahasa . . . . .	24
Gambar 4.15	Hasil dari Potongan Kode pada Gambar 4.14 . . . . .	25
Gambar 4.16	Potongan Kode Deteksi Jumlah Bahasa . . . . .	25
Gambar 4.17	Hasil dari kode pada Gambar 4.16 . . . . .	25
Gambar 4.18	Potongan Kode untuk Menterjemahkan . . . . .	26
Gambar 4.19	Hasil Terjemahan <i>Library</i> "Googletrans" . . . . .	26
Gambar 4.20	Potongan Kode untuk <i>Stemming</i> . . . . .	26
Gambar 4.21	Potongan Kode untuk <i>Labelling</i> . . . . .	27
Gambar 4.22	Potongan Kode untuk Skoring Sentimen . . . . .	28
Gambar 4.23	Hasil Skoring . . . . .	28
Gambar 4.24	Potongan kode untuk <i>labelling</i> . . . . .	29
Gambar 4.25	Potongan Kode untuk Membuang Duplikasi . . . . .	29
Gambar 4.26	Potongan Kode Pembobotan TF-IDF . . . . .	30
Gambar 4.27	Potongan Kode Data Splitting . . . . .	30
Gambar 4.28	Potongan Kode <i>Multinomial Naïve Bayes</i> . . . . .	31
Gambar 4.29	Potongan Kode <i>Complement Naïve Bayes</i> . . . . .	32
Gambar 4.30	Total Data . . . . .	32
Gambar 4.31	Hasil Pengujian <i>Multinomial Naïve Bayes</i> . . . . .	33
Gambar 4.32	<i>Confusion Matrix</i> untuk <i>Multinomial Naïve Bayes</i> . . . . .	33
Gambar 4.33	Hasil Pengujian <i>Complement Naïve Bayes</i> . . . . .	34
Gambar 4.34	<i>Confusion Matrix</i> untuk <i>Complement Naïve Bayes</i> . . . . .	34

## DAFTAR TABEL

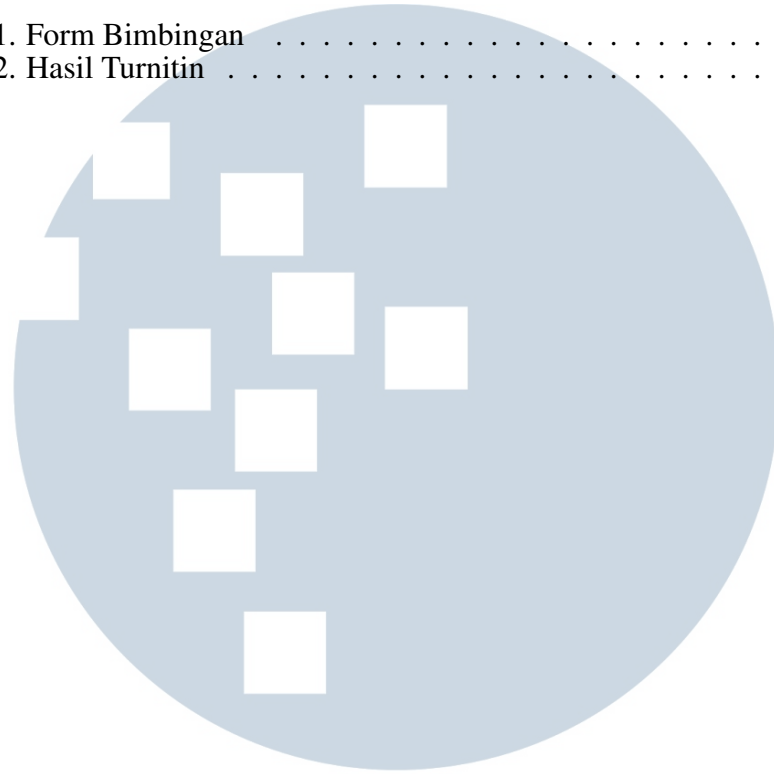
Tabel 2.1	Confusion Matrix . . . . .	9
Tabel 4.1	Tabel Evaluasi . . . . .	35
Tabel 4.2	Tabel Hasil <i>Multinomial Naïve Bayes</i> . . . . .	36
Tabel 4.3	Tabel Hasil <i>Complement Naïve Bayes</i> . . . . .	37



UMMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan .....	44
Lampiran 2. Hasil Turnitin .....	46



# UMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA