



**IMPLEMENTASI SENTIMEN ANALISIS PADA TWITTER
MENGUNAKAN BERT TERHADAP KOMENTAR
KENAIKAN HARGA BEBERAPA KOMODITAS DALAM DUA
BAHASA**

LAPORAN SKRIPSI

**UNIVERSITAS
MERCU BUANA**
MUHAMMAD FARHAN
41518110084

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2024**

HALAMAN JUDUL



**IMPLEMENTASI SENTIMEN ANALISIS PADA TWITTER
MENGUNAKAN BERT TERHADAP KOMENTAR
KENAIKAN HARGA BEBERAPA KOMODITAS DALAM DUA
BAHASA**

LAPORAN SKRIPSI

**UNIVERSITAS
MUHAMMAD FARHAN
41518110084
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2024**

LEMBAR PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Farhan
NIM : 41518110084
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi Sentimen Analisis Pada Twitter Menggunakan BERT Terhadap Komentar Kenaikan Harga Beberapa komoditas Dalam Dua Bahasa

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 5 Februari 2024



(Muhammad Farhan)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Farhan
NIM : 41518110084
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi Sentimen Analisis Pada Twitter Menggunakan BERT Terhadap Komentar Kenaikan Harga Beberapa komoditas Dalam Dua Bahasa

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Anis Cherid, SE, MTI.
NIDN : 0328127203
Ketua Penguji : Dr. Ir. Eliyani
NIDN : 0321026901
Penguji 1 : Dr. Hadi Santoso, S.Kom.,
M.Kom
NIDN : 0225067701
Penguji 2 : Saruni Dwiasnati, ST., MM
NIDN : 0325128802



MERCU BUANA

Jakarta, 5 Februari 2024

Mengetahui,

Dekan



Dr. Bambang Jokonowo, S.Si., MTI

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas rahmat dan karunia-Nya saya dapat menyelesaikan Laporan Akhir ini. Penulisan Laporan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk memperoleh gelar Sarjana Ilmu Komputer Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan hingga penyusunan Tugas Akhir ini, akan sangat sulit bagi saya untuk menyelesaikan Laporan Akhir ini. Oleh karena itu, saya ingin mengucapkan terima kasih kepada:

1. Bapak Anis Cherid, SE, MTI selaku Dosen Pembimbing yang telah memberikan waktu, tenaga, dan pikirannya untuk mengarahkan saya dalam penyusunan tugas akhir ini.
2. Dr. Ir. Eliyani selaku ketua penguji, Dr. Hadi Santoso, S.Kom., M.Kom dan Saruni Dwiasnati, ST., MM selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
3. Kedua orang tua ibu dan ayah serta adik yang selalu mendoakan dan memberikan dukungan.
4. Semua pihak yang telah banyak membantu dalam penyelesaian skripsi ini yang tidak dapat penulis sebutkan satu persatu.
5. Kepada Stussy yang telah menjadi peliharaan saya dan menemani saya saat masa-masa penyelesaian Tugas Akhir.
6. Kepada Nova Aulia sebagai adik cina yang paling supportif dalam keuangan saya, dan makanan enak untuk memenuhi kebutuhan gizi saya sebagai mahasiswa.
7. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all these hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, I wanna thank me for always being a giver and trying to give more than I receive. I wanna thank me for trying do more right than wrong, I wanna thank me for just being me all times.*

Akhir kata, semoga Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Tugas Akhir ini membawa manfaat bagi perkembangan ilmu pengetahuan

Jakarta, 5 Februari 2024



Muhammad Farhan

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Farhan
NIM : 41518110084
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi Sentimen Analisis Pada Twitter Menggunakan BERT Terhadap Komentar Kenaikan Harga Beberapa komoditas Dalam Dua Bahasa

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 5 Februari 2024

Yang menyatakan



(Muhammad Farhan)

UNIVERSITAS
MERCU BUANA

ABSTRAK

Nama : Muhammad Farhan
NIM : 41518110084
Program Studi : Teknik Informatika
Judul : Implementasi Sentimen Analisis Pada Twitter
Menggunakan BERT Terhadap Komentar Kenaikan Harga
Beberapa komoditas Dalam Dua Bahasa.
Pembimbing : Anis Cherid, SE, MTI.

Twitter merupakan platform media sosial terkemuka dengan lebih dari 1 miliar pengguna di Google PlayStore. Ini merupakan tempat buat berbagi komentar, data, serta berhubungan. Peningkatan harga komoditas semacam bahan bakar, bahan pokok, serta yang lain kerap dibahas di sini. Ketersediaan pangan sangat berarti, serta riset ini memanfaatkan pendekatan baru ialah model BERT guna menganalisis sentimen di Twitter. Kami menguji model pre- training BERT, khususnya mBERT, pada komentar Tweet tentang peningkatan harga. Kami mendeteksi kalau model fine- tuned mBERT bahasa Indonesia mempunyai akurasi besar(95%) dengan hyperparameter tertentu. Tata cara pelabelan data mempengaruhi hasil, dengan pelabelan berbasis TextBlob bahasa Indonesia memberikan akurasi terbaik. Model mBERT dapat memprediksi sentimen dengan baik, serta hasil terbaik diperoleh dengan akurasi 95% pada data bahasa Indonesia memakai hyperparameter yaitu Learning Rate 0.00002, Batch Size 32, jumlah Epoch 5 dan waktu pelatihan 10 menit.

Kata kunci: Analisis sentimen, BERT, Twitter, NLP

UNIVERSITAS
MERCU BUANA

ABSTRACT

Name : Muhammad Farhan
Student Number : 41518110084
Study Program : Engineering Informatics
Title : Implementation of Sentiment Analysis on Twitter Using BERT to Comment on Price Increases for Several Commodities in Two Languages
Counsellor : Anis Cherid, SE, MTI.

Twitter is a leading social media platform with over 1 billion users on Google PlayStore. This is a place to share feedback, data and connect. Here we often discuss increases in the prices of raw materials such as fuel, basic materials and others. Food availability is very important and this research uses a new approach, namely the BERT model, to analyze sentiment on Twitter. We tested BERT pre-training models, specifically mBERT, on Tweet comments about price increases. We found that the optimized Indonesian mBERT model had great accuracy (95%) with some hyperparameters. The data labeling procedure influences the results, with TextBlob-based Indonesian labeling providing the highest accuracy. The mBERT model can predict sentiment well and the best results are achieved with 95% accuracy on Indonesian language data using hyperparameters, i.e. learning rate 0.00002, batch size 32, number of epochs 5 and training time 10 minutes.

Keywords: Sentiment Analysis, BERT, Twitter, NLP

UNIVERSITAS
MERCU BUANA

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERNYATAAN KARYA SENDIRI.....	ii
LEMBAR PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Penelitian	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Penelitian Terdahulu	7
2.2 Teori Pendukung	16
BAB III METODOLOGI PENELITIAN	28
3.1 Jenis Penelitian.....	28
3.2 Tahapan Penelitian.....	28
BAB IV HASIL DAN PEMBAHASAN	38
4.1 Dataset.....	38
4.2 Pre-Processing.....	39
4.3 Pembuatan Model	42
4.4 Pengujian dan Evaluasi	44

BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan	57
5.2 Saran.....	58
DAFTAR PUSTAKA	59
LAMPIRAN.....	62
Lampiran 1. Bimbingan TA	62
Lampiran 2. Surat Pernyataan Luaran Tugas Akhir	63
Lampiran 3. Bukti Submit Artikel Ilmiah	64
Lampiran 4. Naskah Artikel Jurnal	65
Lampiran 5. Curriculum Vitae	67
Lampiran 6. Surat Pernyataan HKI.....	69



UNIVERSITAS
MERCU BUANA

DAFTAR TABEL

Tabel 2.1 Penelitian terkait	7
Tabel 3.1 Distribusi label berbasis TextBlob pada bahasa indonesia	33
Tabel 3.2 Distribusi label berbasis TextBlob pada bahasa melayu.....	33
Tabel 3.3 Hyperparameter yang digunakan penelitian ini	35
Tabel 4.1 Deskripsi dataset	38
Tabel 4.2 Hasil pre-processing dataset.....	40
Tabel 4.3 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk.....	45
Tabel 4.4 Hasil fine-tuning dua model BERT pelabelan berbasis Textblob pada bahasa melayu berdasarkan seluruh produk.....	45
Tabel 4. 5 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk menggunakan Epoch 2	45
Tabel 4. 6 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan seluruh produk menggunakan Epoch 2.....	45
Tabel 4. 7 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk menggunakan Optimizer Adafactor.....	46
Tabel 4. 8 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan seluruh produk menggunakan Optimizer Adafactor.....	46
Tabel 4. 9 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan produk.....	46
Tabel 4. 10 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan produk	46
Tabel 4. 11 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa.....	52
Tabel 4. 12 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa menggunakan Epoch 2	52

Tabel 4. 13 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa menggunakan Optimizer Adafactor	52
Tabel 4. 14 Hasil classification report Model BERT pelabelan berbasis Textblob bahasa Indonesia pada setiap produk	53
Tabel 4. 15 Hasil classification report Model BERT pelabelan berbasis Textblob bahasa Melayu pada setiap produk	53
Tabel 4. 16 Hasil sentiment prediction dua model BERT pelabelan berbasis Textblob	53
Tabel 4. 17 Perbandingan hasil dengan penelitian lain.....	56
Tabel 4.1 Deskripsi dataset	38
Tabel 4.2 Hasil pre-processing dataset.....	40
Tabel 4.3 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk.....	45
Tabel 4.4 Hasil fine-tuning dua model BERT pelabelan berbasis Textblob pada bahasa melayu berdasarkan seluruh produk.....	45
Tabel 4. 5 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk menggunakan Epoch 2	45
Tabel 4. 6 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan seluruh produk menggunakan Epoch 2.....	45
Tabel 4. 7 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan seluruh produk menggunakan Optimizer Adafactor.....	46
Tabel 4. 8 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan seluruh produk menggunakan Optimizer Adafactor.....	46
Tabel 4. 9 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Indonesia berdasarkan produk.....	46
Tabel 4. 10 Hasil fine-tuning model BERT pelabelan berbasis TextBlob pada bahasa Melayu berdasarkan produk	46
Tabel 4. 11 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa.....	52

Tabel 4. 12 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa menggunakan Epoch 2	52
Tabel 4. 13 Hasil classification report Model BERT pelabelan berbasis Textblob dengan dua bahasa menggunakan Optimizer Adafactor	52
Tabel 4. 14 Hasil classification report Model BERT pelabelan berbasis Textblob bahasa Indonesia pada setiap produk	53
Tabel 4. 15 Hasil classification report Model BERT pelabelan berbasis Textblob bahasa Melayu pada setiap produk	53
Tabel 4. 16 Hasil sentiment prediction dua model BERT pelabelan berbasis Textblob	53
Tabel 4. 17 Perbandingan hasil dengan penelitian lain.....	56



DAFTAR GAMBAR

Gambar 2. 1 Transformer	18
Gambar 2. 2 Bagaimana cara kerja BERT	19
Gambar 2. 3 Arsitektur model BERT _{BASE}	20
Gambar 2. 4 Arsitektur model BERT _{LARGE}	21
Gambar 2. 5 Arsitektur model BERT lainnya.....	22
Gambar 2. 6 Input data representation BERT	23
Gambar 2. 7 Confusion matrix.....	24
Gambar 2. 8 Gambaran umum model CRISP-DM	25
Gambar 3. 1 Tahapan penelitian ini	28
Gambar 3. 2 Tahapan pre-processing.....	31
Gambar 3. 3 Pelabelan data berbasis TextBlob	31
Gambar 3. 4 Arsitektur model BERT untuk penelitian ini	34
Gambar 4. 1 Proses coding tahap Pre-processing dataset	40
Gambar 4. 2 Hasil distribusi sentimen keseluruhan pelabelan berbasis TextBlob	41
Gambar 4. 3 Tahap coding pelabelan sentimen dan translate berbasis TextBlob.	42
Gambar 4. 4 Tahap coding menentukan hyperparameter untuk fine-tuning	43
Gambar 4. 5 Prosen input formatting menggunakan fungsi sebelumnya pada keseluruhan data.....	43
Gambar 4. 6 Hasil training mBERT batch size 16 pelabelan berbasis TextBlob berbahasa indonesia	47
Gambar 4. 7 Hasil training mBERT batch size 32 pelabelan berbasis TextBlob berbahasa indonesia	47
Gambar 4. 8 Hasil training mBERT batch size 16 pelabelan berbasis TextBlob berbahasa melayu	48
Gambar 4. 9 Hasil training mBERT batch size 32 pelabelan berbasis TextBlob berbahasa melayu	48
Gambar 4. 10 Hasil training mBERT batch size 16 pelabelan berbasis TextBlob berbahasa indonesia menggunakan epoch 2	49

Gambar 4. 11 Hasil training mBERT batch size 32 pelabelan berbasis TextBlob berbahasa indonesia menggunakan epoch 2	49
Gambar 4. 12 Hasil training mBERT batch size 16 pelabelan berbasis TextBlob berbahasa Melayu menggunakan epoch 2.....	50
Gambar 4. 13 Hasil training mBERT batch size 32 pelabelan berbasis TextBlob berbahasa Melayu menggunakan epoch 2.....	50
Gambar 4. 14 Hasil nilai probability mBERT review text berbahasa indonesia tabel 4.13	54
Gambar 4. 15 Hasil nilai probability mBERT review text berbahasa melayu tabel 4.13.....	54
Gambar 4. 16 Hasil Wordcloud pada sentimen positif dalam text berbahasa indonesia	54
Gambar 4. 17 Hasil Wordcloud pada sentimen negatif dalam text berbahasa indonesia	55
Gambar 4. 18 Hasil Wordcloud pada sentimen netral dalam text berbahasa indonesia	55

DAFTAR LAMPIRAN

Lampiran 1. Bimbingan TA	62
Lampiran 2. Surat Pernyataan Luaran Tugas Akhir.....	63
Lampiran 3. Lampiran Naskah Artikel Jurnal.....	64
Lampiran 4. Curriculum Vitae	65
Lampiran 5. Surat Pernyataan HAKI	67
Lampiran 6. Sertifikat BNSP	69
Lampiran 7. Form Revisi Dosen Penguji.....	70



UNIVERSITAS
MERCU BUANA