

**ANALISIS SENTIMEN PADA TWITTER TERHADAP AKSI
BOIKOT PRODUK ISRAEL MENGGUNAKAN
BIDIRECTIONAL ENCODER REPRESENTATIONS FROM
TRANSFORMERS (BERT)**

SKRIPSI

**Diajukan untuk memenuhi persyaratan
dalam memperoleh gelar Sarjana Komputer
Program Studi Sistem Informasi**

Disusun oleh:



**ZILVI AZUS SRIYANTI
20082010057**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
SURABAYA**

2024

SKRIPSI

ANALISIS SENTIMEN PADA TWITTER TERHADAP AKSI BOIKOT
PRODUK ISRAEL MENGGUNAKAN BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS (BERT)

Disusun Oleh:

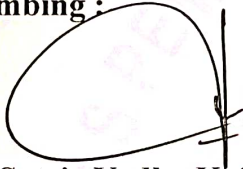
ZILVI AZUS SRIYANTI

20082010057

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi
Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas
Pembangunan Nasional "Veteran" Jawa Timur Pada Tanggal 16 Juli 2024

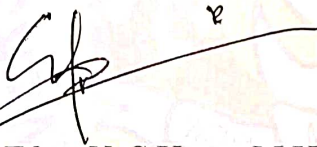
Pembimbing :

1.



Dhian Satria Yudha K, S.Kom., M.Kom.
NPT. 2 0119 86 052224 9

2.



Abdul Rezha Efrat N, S.Kom., M.Kom.
NIP. 19940929 2022031 008

Tim Penguji :

1.



Eka Dyar Wahyuni, S.Kom., M.Kom.
NIP. 19841201 2021212 005

2.



Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 2022032 007

3.



Anindo Saka Fitri, S.Kom., M.Kom.
NPT. 2 1219 93 032526 8

Mengetahui

Dekan Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur



Prof. Dr. Ir. Novirina Hendrasarie, M.T.
NIP. 19681126 199403 2 001

LEMBAR PENGESAHAN

**ANALISIS SENTIMEN PADA TWITTER TERHADAP AKSI BOIKOT
PRODUK ISRAEL MENGGUNAKAN BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS (BERT)**

Disusun Oleh:

ZILVI AZUS SRIYANTI

20082010057

**Telah disetujui Ujian Negara Lisan Gelombang Juli Periode 2024 pada
Tanggal 16 Juli 2024**

Menyetujui,

Dosen Pembimbing 1

Dosen Pembimbing 2

Dhian Satria Yudha N., S.Kom, M.Kom

Abdul Rezha Efrat N., S.Kom, M.Kom

NPT. 2 0119 86 052224 9

NIP. 19940929 2022031 008

Mengetahui,

**Ketua Program Studi Sistem Informasi
Fakultas Ilmu Komputer**

Universitas Pembangunan Nasional "Veteran" Jawa Timur

Agung Brastama Putra, S.Kom, M.Kom.

NIP. 19851124 2021211 003



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR FAKULTAS ILMU KOMPUTER

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Zilvi Azus Sriyanti

NPM : 20082010057

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 19 Juli 2024 dengan judul:

**"ANALISIS SENTIMEN PADA TWITTER TERHADAP AKSI BOIKOT
PRODUK ISRAEL MENGGUNAKAN BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS (BERT)"**

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.

Surabaya, 19 Juli 2024

Dosen penguji yang memeriksa revisi:

1. Eka Dyar Wahyuni, S.Kom., M.Kom.




NIP. 19841201 2021212 005

2. Reisa Permatasari, S.T., M.Kom.

NIP. 19920514 2022032 007

3. Anindo Saka Fitri, S.Kom., M.Kom.

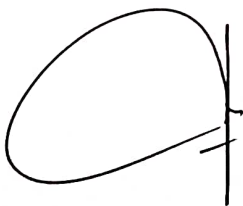
NPT. 2 1219 93 032526 8

{  }
{  }
{  }

Mengetahui,

Dosen Pembimbing 1

Dosen Pembimbing 2



Dhian Satria Yudha K, S.Kom, M.Kom.

NPT. 2 0119 86 052224 9



Abdul Rezha Efrat N, S.Kom, M.Kom.

NIP. 19940929 2022031 008



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR FAKULTAS ILMU KOMPUTER

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Zilvi Azus Sriyanti

NPM : 20082010057

Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi/Tugas Akhir sebagai berikut :

**"ANALISIS SENTIMEN PADA TWITTER TERHADAP AKSI BOIKOT
PRODUK ISRAEL MENGGUNAKAN BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS (BERT)"**

Bukan merupakan plagiat dari skripsi / Tugas Akhir / Penelitian orang lain dan bukan merupakan Produk / Software / Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun institusi pendidikan lain. Jika ternyata kemudian hari pernyataan ini terbukti tidak benar, maka saya bertanggung jawab penuh dan siap menerima segala kosekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 19 Juli 2024

Hormat Saya,



Zilvi Azus Sriyanti

NPM. 20082010057

Judul : **ANALISIS SENTIMEN PADA TWITTER
TERHADAP AKSI BOIKOT PRODUK ISRAEL
MENGUNAKAN BIDIRECTIONAL ENCODER
REPRESENTATIONS FROM TRANSFORMERS
(BERT)**

Pembimbing 1 : **Dhian Satria Yudha Kartika, S.Kom, M.Kom**

Pembimbing 2 : **Abdul Rezha Efrat Najaf, S.Kom, M.Kom**

ABSTRAK

Konflik yang telah terjadi antara Israel dan Palestina menjadi sumber perdebatan global. Konflik tersebut menyebabkan timbulnya aksi boikot terhadap produk yang terafiliasi Israel. Aksi tersebut menimbulkan reaksi positif maupun negatif dari masyarakat Indonesia yang diungkapkan melalui media sosial *Twitter*. Penelitian ini bertujuan untuk menganalisis sentimen terhadap aksi boikot produk Israel pada *Twitter* untuk memperoleh pandangan yang mendalam mengenai aksi boikot produk Israel. Penelitian ini melalui tahapan pengumpulan data dari *Twitter* dan tahap pemrosesan data *tweet*. Data yang telah melalui tahap pemrosesan sebanyak 7.881 data. Klasifikasi sentimen menggunakan model BERT untuk mengklasifikasikan sentimen positif dan sentimen negatif. Penelitian ini menggunakan pelabelan otomatis *Vader Lexicon* dan melakukan penyeimbangan kelas sebelum klasifikasi menggunakan *Random Oversampling*. Pengujian model dilakukan menggunakan *confusion matrix*. Nilai akurasi model yang didapatkan dalam penelitian ini yaitu 85%, presisi 85%, *recall* 85%, dan *f1-score* 85%.

Visualisasi ditampilkan dalam *website* menggunakan *streamlit*. Hasil uji validasi sistem mendapatkan nilai akurasi 76%. Sentimen yang banyak diungkapkan masyarakat yaitu sentimen positif yang mengarah pada dukungan untuk melakukan aksi boikot produk Israel. Hasil penelitian ini diharapkan dapat memberikan pemahaman respon publik terhadap aksi boikot produk Israel di Indonesia, serta disarankan untuk masa yang akan datang dapat menggunakan skenario model klasifikasi dan pengujian yang berbeda.

Kata kunci :

Analisis sentimen, boikot Israel, klasifikasi, BERT, *random oversampling*

KATA PENGANTAR

Puji dan syukur kehadiran Allah SWT, atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi yang merupakan persyaratan dalam menyelesaikan Program Studi S1/Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Dalam penyusunan skripsi ini tidak terlepas dari bantuan berbagai pihak, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua saya yang selalu memberikan do'a, motivasi, dan dukungan penuh untuk kelancaran penyelesaian skripsi ini.
2. Kakak dan adik saya yang senantiasa memberikan dukungan dalam pengerjaan skripsi ini.
3. Bapak Dhian Satria Yudha Kartika, S.Kom, M.Kom, selaku dosen pembimbing 1 dan Bapak Abdul Rezha Efrat Najaf, S.Kom, M.Kom, selaku dosen pembimbing 2 yang senantiasa memberikan bimbingan dan saran terkait penyelesaian skripsi ini serta motivasi semangat selama pengerjaan skripsi ini.
4. Bapak Agung Brastama Putra, S.Kom, M.Kom, selaku koordinator program studi Sistem Informasi yang telah memberikan motivasi dan dukungan dalam penyelesaian skripsi ini.
5. Seluruh Bapak/Ibu Dosen Sistem Informasi UPN “Veteran” Jawa Timur yang telah memberikan dukungan, motivasi, dan ilmu pengetahuan dari awal perkuliahan hingga selesainya skripsi ini.
6. Teman-teman ATENSI 2020, yang telah memberikan semangat dan dukungan dari awal perkuliahan hingga masa akhir perkuliahan.

7. Seluruh mentor dan teman-teman MSIB batch 3 di *Zenius Data Analytics* dan batch 4 di *PT Stechoq Robotika Indonesia Software Engineering Specialist* yang telah memberikan ilmunya dalam program MBKM terutama dalam bidang data mining dan web sehingga dapat diterapkan dalam pengerjaan skripsi ini.
8. Terima kasih kepada Shofi sebagai teman sekamar kos yang selalu mendengar keluh kesah dalam pengerjaan skripsi ini dan teman-teman VETERAN MUDA, Sasa, Nilam, Gisa, dan Novita yang turut memberikan dukungan dan menjadi teman seperjuangan skripsi.
9. Terima kasih kepada Mila dan teman-teman satu bimbingan yang senantiasa turut mendukung proses pengerjaan dan menjadi teman diskusi dalam penyelesaian skripsi.

Penulis menyadari bahwa dalam laporan skripsi ini masih banyak kekurangan. Namun, penulis berharap semoga skripsi ini dapat bermanfaat dalam perkembangan ilmu pengetahuan.

Surabaya, Juli 2024

Penulis

DAFTAR ISI

	Hal
ABSTRAK	ii
KATA PENGANTAR.....	iv
DAFTAR ISI	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9
2.1 Dasar Teori	9
2.1.1 Boikot.....	9
2.1.2 Analisis Sentimen.....	10
2.1.3 <i>Twitter</i>	11
2.1.4 <i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i> ...	12
2.1.5 <i>Text Mining</i>	14
2.1.6 <i>Natural Language Processing (NLP)</i>	15
2.1.7 <i>Preprocessing Text</i>	16
2.1.8 <i>Natural Language Toolkit (NLTK)</i>	17
2.1.9 Sastrawi.....	18

2.1.10 Vader	18
2.1.11 <i>Random Oversampling</i>	19
2.1.12 <i>Bidirectional Encoder Representations from Transformers (BERT)</i>	19
2.1.12.1 IndoBERT	21
2.1.12.2 <i>Hyperparameter</i>	23
2.1.13 <i>Confusion Matrix</i>	25
2.1.14 Streamlit	27
2.2 Penelitian Terdahulu	27
BAB III METODOLOGI PENELITIAN	33
3.1 Studi Literatur.....	34
3.2 Analisis Kebutuhan.....	35
3.2.1 Kebutuhan Data.....	35
3.2.2 Kebutuhan <i>Hardware dan Software</i>	35
3.3 Pengumpulan Data	35
3.4 <i>Preprocessing Text</i>	37
3.5 Penyaringan Data	39
3.6 Pelabelan Data.....	39
3.7 <i>Exploratory Data Analysis</i>	39
3.8 Klasifikasi.....	40
3.8.1 <i>Word Embedding</i>	40
3.8.2 Pembagian Data	40
3.8.3 Klasifikasi Sentimen Menggunakan BERT	41
3.9 Evaluasi	41
3.10 Visualisasi dan Validasi	41
3.11 Pembuatan Laporan	43

BAB IV HASIL DAN PEMBAHASAN.....	44
4.1 Analisis Kebutuhan.....	44
4.1.1 Analisis Kebutuhan Data.....	44
4.1.2 Kebutuhan <i>Hardware</i> dan <i>Software</i>	44
4.2 Pengumpulan Data	45
4.3 <i>Preprocessing Text</i>	49
4.4 Penyaringan Data	60
4.5 Pelabelan Data.....	61
4.6 <i>Exploratory Data Analysis</i>	63
4.7 Klasifikasi.....	72
4.7.1 <i>Word Embedding</i>	72
4.7.1 Pembagian Data	78
4.7.3 Klasifikasi Sentimen Menggunakan BERT	78
4.8 Evaluasi	80
4.9 Visualisasi dan Validasi	89
BAB V KESIMPULAN DAN SARAN.....	96
5.1 Kesimpulan.....	96
5.2 Saran.....	98
DAFTAR PUSTAKA	99
LAMPIRAN.....	103

DAFTAR TABEL

	Hal
Tabel 2. 1 Rentang <i>Compound Score Vader</i>	19
Tabel 2. 2 <i>Confusion Matrix Biner</i>	25
Tabel 2. 3 Penelitian Terdahulu.....	28
Tabel 4. 1 Contoh Hasil <i>Scrapping Data</i>	46
Tabel 4. 2 Contoh Hasil <i>Case Folding</i>	50
Tabel 4. 3 Contoh Hasil <i>Data Cleaning</i>	52
Tabel 4. 4 Contoh Hasil <i>Normalization</i>	55
Tabel 4. 5 Contoh Hasil <i>Stopwords Removal</i>	57
Tabel 4. 6 Contoh Hasil <i>Stemming</i>	59
Tabel 4. 7 Contoh Hasil Penerjemahan Teks.....	61
Tabel 4. 8 Hasil Pelabelan Data	63
Tabel 4. 9 Jumlah Data untuk Skenario Klasifikasi	64
Tabel 4. 10 Distribusi Pembagian <i>Training</i> dan <i>Testing Data</i> (80:20).....	78
Tabel 4. 11 Distribusi Pembagian <i>Training</i> dan <i>Testing Data</i> (70:30).....	78
Tabel 4. 12 <i>Classification Report</i>	81
Tabel 4. 13 Perbandingan Hasil Prediksi dan Label Asli.....	95

DAFTAR GAMBAR

	Hal
Gambar 2. 1 Proses CRISP-DM (Sumarauw, 2022)	13
Gambar 2. 2 Prosedur <i>Pre-Training</i> dan <i>Fine-tuning</i> BERT (Devlin dkk., 2019) 20	20
Gambar 2. 3 Representasi <i>Input</i> BERT (Devlin dkk., 2019).....	21
Gambar 3. 1 Alur Metodologi Penelitian	33
Gambar 3. 2 <i>Trend</i> Kata Kunci Boikot Produk Israel	36
Gambar 3. 3 Alur Preprocessing Text	37
Gambar 3. 4 Rancangan <i>Wireframe</i> Streamlit (Dashboard 1).....	42
Gambar 3. 5 Rancangan <i>Wireframe</i> Streamlit (Dashboard 2).....	42
Gambar 3. 6 Rancangan <i>Wireframe</i> Streamlit (Analisis Teks)	43
Gambar 3. 7 Rancangan <i>Wireframe</i> Streamlit (Analisis File CSV).....	43
Gambar 4. 1 <i>Source Code</i> Pengambilan Data	45
Gambar 4. 2 <i>Word Cloud</i> Sebelum Pemrosesan Data	49
Gambar 4. 3 <i>Source Code Case Folding</i>	50
Gambar 4. 4 <i>Source Code Data Cleaning</i>	52
Gambar 4. 5 <i>Source Code Normalization</i>	54
Gambar 4. 6 <i>Source Code Stopwords Removal</i>	57
Gambar 4. 7 <i>Source Code Stemming</i>	59
Gambar 4. 8 <i>Source Code</i> Pelabelan.....	62
Gambar 4. 9 Diagram Batang Persebaran Label.....	64
Gambar 4. 10 <i>Source Code Balance Data Random Over Sampling</i>	65
Gambar 4. 11 Diagram Batang Data Seimbang	65
Gambar 4. 12 Grafik Jumlah <i>Tweet</i> Per Bulan.....	66
Gambar 4. 13 Grafik Jumlah <i>Tweet</i> Per Tanggal	67
Gambar 4. 14 Grafik Distribusi Panjang <i>Tweet</i>	68
Gambar 4. 15 Grafik Frekuensi Kata	69
Gambar 4. 16 <i>Word Cloud</i>	70
Gambar 4. 17 <i>Word Cloud</i> Sentimen Positif.....	71
Gambar 4. 18 <i>Word Cloud</i> Sentimen Negatif.....	71
Gambar 4. 19 Ilustrasi Tokenisasi	72

Gambar 4. 20 <i>Source Code</i> Tokenisasi	73
Gambar 4. 21 Ilustrasi Penambahan [CLS] dan [SEP]	73
Gambar 4. 22 <i>Source Code</i> Menambahkan [CLS] dan [SEP]	73
Gambar 4. 23 Ilustrasi Mengubah Token ke ID	74
Gambar 4. 24 <i>Source Code</i> Mengubah Token ke ID	74
Gambar 4. 25 Ilustrasi Menambahkan <i>Padding</i>	75
Gambar 4. 26 <i>Source Code</i> Menambahkan <i>Padding</i>	75
Gambar 4. 27 Ilustrasi Menambahkan <i>Attention Mask</i>	76
Gambar 4. 28 <i>Source Code</i> Menambahkan <i>Attention Mask</i>	76
Gambar 4. 29 Ilustrasi Menambahkan <i>Positional Embedding</i>	77
Gambar 4. 30 <i>Source Code</i> Menambahkan <i>Positional Embedding</i>	77
Gambar 4. 31 <i>Source Code Training</i>	79
Gambar 4. 32 <i>Confusion Matrix</i> Skenario Terbaik <i>Dropout</i> 0.1 (80:20).....	82
Gambar 4. 33 Grafik Akurasi dan <i>Loss Training Validation Dropout</i> 0.1 Skenario Terbaik (80:20).....	83
Gambar 4. 34 <i>Confusion Matrix</i> Skenario Terbaik <i>Dropout</i> 0.3 (80:20).....	84
Gambar 4. 35 Grafik Akurasi dan <i>Loss Training Validation Dropout</i> 0.3 Skenario Terbaik (80:20).....	84
Gambar 4. 36 <i>Confusion Matrix</i> Skenario Terbaik <i>Dropout</i> 0.1 (70:30).....	86
Gambar 4. 37 Grafik Akurasi dan <i>Loss Training Validation Dropout</i> 0.1 Skenario Terbaik (70:30).....	86
Gambar 4. 38 <i>Confusion Matrix</i> Skenario Terbaik <i>Dropout</i> 0.3 (70:30).....	87
Gambar 4. 39 Grafik Akurasi dan <i>Loss Training Validation Dropout</i> 0.3 Skenario Terbaik (70:30).....	88
Gambar 4. 40 Tampilan Visualisasi Halaman <i>Dashboard</i> 1	90
Gambar 4. 41 Tampilan Visualisasi Halaman <i>Dashboard</i> 2	91
Gambar 4. 42 Fitur Filter Sentimen	92
Gambar 4. 43 Fitur Filter Tanggal.....	92
Gambar 4. 44 Fitur Filter Jumlah Kata Teratas	93
Gambar 4. 45 Halaman Validasi dengan <i>Input Text</i>	93
Gambar 4. 46 Halaman Validasi dengan <i>File CSV</i>	94

DAFTAR LAMPIRAN

Lampiran 1 : *Source Code* Visualisasi dengan Streamlit