



Perbandingan Analisis Sentimen Twitter Terkait Pelayanan BPJS Dan Non-BPJS Kesehatan Menggunakan Algoritma BERT

LAPORAN TUGAS AKHIR

HENGKI PRAYOGO

41521010038

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2025



Perbandingan Analisis Sentimen Twitter Terkait Pelayanan BPJS Dan Non-BPJS Kesehatan Menggunakan Algoritma BERT

LAPORAN TUGAS AKHIR

HENGKI PRAYOGO

41521010038

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : HENGKI PRAYOGO
NIM : 41521010038
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Perbandingan Analisis Sentimen Twitter Terkait Pelayanan BPJS Kesehatan Dan Non-BPJS Kesehatan Menggunakan Algoritma BERT

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 Proposal Penelitian saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 10 Juli 2025



Hengki Prayogo

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : HENGKI PRAYOGO
NIM : 41521010038
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Analisis Sentimen Twitter
Terkait Pelayanan BPJS Kesehatan Dan Non-
BPJS Kesehatan Menggunakan Algoritma
BERT

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 : Dwiki Jatikusumo, S.Kom, M.Kom.
NIDN : 0301128903
Ketua Penguji : Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701
Penguji 1 : Muhammad Rifqi, S.Kom, M.Kom
NIDN : 0301067101
Penguji 2 : Ilham Nugraha, S.Kom, M.Sc
NIDN : 307098904

()
()
()
()

MERCU BUANA

Jakarta, 21 Juli 2025

Mengetahui,

Dekan



Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Dwiki Jatikusumo selaku dosen pembimbing MPTI yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensuport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana..
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 21 Juli 2025


HENGKI PRAYOGO

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Hengki Prayogo
NIM : 41521010038
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Analisis Sentimen Twitter Terkait Pelayanan BPJS Kesehatan Dan Non-BPJS Kesehatan Menggunakan Algoritma BERT

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, 21 Juli 2024

Yang menyatakan,



(Hengki Prayogo)

ABSTRAK

Nama : HENGKI PRAYOGO

NIM : 41521010038

Program Studi : Teknik Informatika

Judul Proposal Penelitian : Perbandingan Analisis Sentimen Twitter Terkait BPJS Kesehatan Dan Non-BPJS kesehatan Menggunakan Algoritma BERT

Dosen Pembimbing : Dwiki Jatikusumo, S.Kom, M.Kom

Analisis sentimen adalah alat penting untuk memahami persepsi publik terhadap berbagai isu, termasuk pelayanan kesehatan. Dalam konteks Indonesia, BPJS memegang peran vital dalam penyediaan layanan kesehatan, namun memicu beragam kritik dan dukungan di media sosial, khususnya Twitter. Penelitian ini bertujuan menganalisis sentimen dari data Twitter yang berkaitan dengan pelayanan kesehatan BPJS dan *non-BPJS* menggunakan **algoritma machine learning terkini, yaitu BERT (*Bidirectional Encoder Representations from Transformers*)**. Data dikumpulkan melalui Twitter API menggunakan kata kunci yang relevan, menghasilkan ribuan *tweet* yang merepresentasikan sentimen publik seputar perawatan kesehatan. Prosedur praproses digunakan untuk membersihkan data dari gangguan, menghapus duplikat, melakukan *tokenization* pada format teks, dan mengubahnya menjadi representasi yang dapat dipahami oleh algoritma. BERT adalah model pembelajaran mendalam yang memanfaatkan arsitektur transformator untuk memperoleh pemahaman yang lebih mendalam tentang konteks kata-kata dalam sebuah kalimat. BERT dievaluasi berdasarkan berbagai metrik kinerja, termasuk **akurasi, presisi, recall, dan F1-score**, untuk menentukan efektivitasnya dalam mengklasifikasikan sentimen positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa BERT memiliki keunggulan signifikan dalam pemahaman konteks dan akurasi tinggi dalam klasifikasi sentimen. Temuan dari penelitian ini tidak hanya memberikan wawasan berharga bagi pemangku kebijakan dan praktisi kesehatan dalam memahami opini masyarakat terhadap pelayanan BPJS dan *non-BPJS*, tetapi juga menunjukkan potensi penggunaan algoritma canggih dalam analisis data sosial. Penelitian ini diharapkan dapat menjadi dasar bagi studi lebih lanjut mengenai sentimen publik dan perbaikan pelayanan kesehatan di Indonesia.

Kata Kunci: Analisis Sentimen, BERT, Pelayanan Kesehatan, Data Twitter

ABSTRACT

Nama : HENGKI PRAYTOGO
NIM : 41521010038
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Perbandingan Analisis Sentimen Twitter Terkait BPJS Kesehatan Dan Non-BPJS kesehatan Menggunakan Algoritma BERT
Dosen Pembimbing : Dwiki Jatikusumo, S.Kom, M.Kom

Analisis sentimen adalah alat penting untuk memahami persepsi publik terhadap berbagai isu, termasuk pelayanan kesehatan. Dalam konteks Indonesia, BPJS memegang peran vital dalam penyediaan layanan kesehatan, namun memicu beragam kritik dan dukungan di media sosial, khususnya Twitter. Penelitian ini bertujuan menganalisis sentimen dari data Twitter yang berkaitan dengan pelayanan kesehatan BPJS dan *non-BPJS* menggunakan **algoritma machine learning terkini, yaitu BERT (*Bidirectional Encoder Representations from Transformers*)**. Data dikumpulkan melalui Twitter API menggunakan kata kunci yang relevan, menghasilkan ribuan *tweet* yang merepresentasikan sentimen publik seputar perawatan kesehatan. Prosedur praproses digunakan untuk membersihkan data dari gangguan, menghapus duplikat, melakukan *tokenization* pada format teks, dan mengubahnya menjadi representasi yang dapat dipahami oleh algoritma. BERT adalah model pembelajaran mendalam yang memanfaatkan arsitektur transformator untuk memperoleh pemahaman yang lebih mendalam tentang konteks kata-kata dalam sebuah kalimat. BERT dievaluasi berdasarkan berbagai metrik kinerja, termasuk **akurasi, presisi, recall, dan F1-score**, untuk menentukan efektivitasnya dalam mengklasifikasikan sentimen positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa BERT memiliki keunggulan signifikan dalam pemahaman konteks dan akurasi tinggi dalam klasifikasi sentimen. Temuan dari penelitian ini tidak hanya memberikan wawasan berharga bagi pemangku kebijakan dan praktisi kesehatan dalam memahami opini masyarakat terhadap pelayanan BPJS dan *non-BPJS*, tetapi juga menunjukkan potensi penggunaan algoritma canggih dalam analisis data sosial. Penelitian ini diharapkan dapat menjadi dasar bagi studi lebih lanjut mengenai sentimen publik dan perbaikan pelayanan kesehatan di Indonesia.

Kata Kunci: Analisis Sentimen, BERT, Pelayanan Kesehatan, Data Twitter

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN.....	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	v
ABSTRAK.....	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABLE	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	1
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 Teori Utama.....	5
2.1.1 Analisis Sentimen	5
2.1.2 BERT (<i>Bidirectional Encoder Representations from Transformers</i>).....	6
2.2 Teori Pendukung	7
2.2.1 Preprocessing Data Twitter.....	8
2.2.2 Evaluasi Model	8
2.3 Penelitian Terdahulu.....	8
2.4 Gap Penelitian	55
BAB III METODE PENELITIAN	57
3.1 Pendekatan Penelitian.....	57
3.2 Desain Penelitian	58
3.3 Subjek Penelitian	59
3.4 Instrumen Penelitian	59

3.5 Teknik Pengumpulan Data	60
3.6 Analisis Data	63
Proses Perhitungan Prediksi	64
Evaluasi Performa Algoritma	64
Loss Function (Fungsi Kerugian)	65
3.7 Prosedur Penelitian	65
BAB IV HASIL DAN PEMBAHASAN	67
4.1 Deskripsi Dataset dan Pra-pemrosesan Data	67
4.2 Pengumpulan Data	67
4.3 Pra-pemrosesan Data	69
4.4 Implementasi Model BERT	77
4.5 Hasil Evaluasi Performa Model	81
4.6 Analisis Sentimen Twitter Terkait Pelayanan BPJS Kesehatan	86
4.7 Analisis Sentimen Twitter Terkait Pelayanan Non-BPJS Kesehatan	88
4.8 Perbandingan Analisis Sentimen Twitter BPJS Kesehatan dan Non-BPJS Kesehatan	91
4.9 Perbandingan Hasil Analisis Sentimen: Model BERT, AI Generatif (Gemini), dan Analisis Human (secara Manual)	93
4.9.1 Hasil Analisis Sentimen oleh Analisis Human (secara Manual)	93
4.9.2 Hasil Analisis Sentimen oleh AI Generatif (Gemini)	94
4.9.3 Hasil Analisis Sentimen oleh Model BERT	94
4.9.4 Perbandingan Distribusi Sentimen Antar Metode	94
BAB V KESIMPULAN DAN SARAN	99
5.1 Kesimpulan	99
5.2 Saran	99
DAFTAR PUSTAKA	101
LAMPIRAN	104

DAFTAR TABLE

No table of figures entries found.

Table 4. 1 Hasil Pra-Pemrosesan	73
Table 4. 2 Log Pelatihan	81
Table 4. 3 Cofusio Matrix	83
Table 4. 4 Weighted Average.....	85
Table 4. 5Komperatif Distribusi Sentimen	91
Table 4. 6 Perbandingan Distribusi sentimen Berdasarkan Metode Analisis	95



DAFTAR GAMBAR

Gambar 3. 1 Klasifikasi Sentimen	57
Gambar 3. 2 Desain Penelitian.....	58
Gambar 3. 3 Token.....	60
Gambar 3. 4 Tweet Harvest	61
Gambar 3. 5 Crowling BPJS Kesehatan	61
Gambar 3. 6 Crowling Non-BPJS.....	61
Gambar 3. 7 Crowling Layanan BPJS	62
Gambar 3. 8 Crowling Pelayanan Kesehatan.....	62
Gambar 3. 9 Rumah sakit BPJS	62
Gambar 3. 10 Vertifikasi dan Navigasi Direktori	62
Gambar 3. 11 Penggabungan Dan Penyimpanan Data Akhir	63
Gambar 3. 12 Alur Penelitian	66
Gambar 4. 1 Instalasi Library.....	69
Gambar 4. 2 Output Instlasi Libarry	70
Gambar 4. 3 Upload CSV	70
Gambar 4. 4 Import Pandas Dan Muat Data.....	70
Gambar 4. 5 Output Import Pandas Dan Muat Data.....	71
Gambar 4. 6 Cleaning Data.....	71
Gambar 4. 7 Output Cleaning Data.....	72
Gambar 4. 8 Distribusi Panjang Teks Setelah Cleaning	72
Gambar 4. 9 Tokenizer.....	74
Gambar 4. 10 Output Tokenizer.....	74
Gambar 4. 11 Perlabelan Sentimen.....	75
Gambar 4. 12 Output Perlabelan Sentimen.....	75
Gambar 4. 13 Grafik Perlabelasn Sentimen.....	76
Gambar 4. 14 Output Perlabelan Sentimen.....	76
Gambar 4. 15 Insialisasi Model Bert.....	77
Gambar 4. 16 Output Inisialisasi Model Bert	77
Gambar 4. 17 Pelatihan Model Bert.....	78
Gambar 4. 18 Output pelatihan Model Bert.....	78

Gambar 4. 19 Training Arguments	79
Gambar 4. 20 Output Training Arguments	79
Gambar 4. 21 Pelatihan Model Dengan Trainer	80
Gambar 4. 22 Output Pelatiha Model Dengan Trainer	80
Gambar 4. 23 Evaluasi Performa Model.....	82
Gambar 4. 24 Evaluasi Performa Model.....	82
Gambar 4. 25 Confusion Matrik	83
Gambar 4. 26 Sentimen PJS Kesehatan	86
Gambar 4. 27 Grafik Distribusi Sentimen BPJS Kesehatan	87
Gambar 4. 28 Analisis Sentimen Per Kategori	89
Gambar 4. 29 Output Analisis Sentimen Per Kategori	89
Gambar 4. 30 Grafik Distribusi Sentimen Non-BPJS Kesehatan.....	90
Gambar 4. 31 Grafik Perbandingan Distribusi Sentimen Twitter.....	92
Gambar 4. 32 Grafik Distribusi Sentimen Antar Metode	95

