

**ANALISIS TRANSFER LEARNING INDOBERT DENGAN RANDOM FOREST
PADA KLASIFIKASI SENTIMEN LAYANAN KESEHATAN BPJS**

TESIS

Oleh:

**STEVEN CIAM
NIM. 231232161**



**PROGRAM STUDI S-2 TEKNOLOGI INFORMASI
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2025**

**ANALYSIS OF TRANSFER LEARNING INDOBERT WITH RANDOM FOREST
FOR SENTIMENT CLASSIFICATION OF BPJS HEALTHCARE SERVICES**

THESIS

By:

**STEVEN CIAM
ID NUMBER. 231232161**



**MAJOR OF S-2 INFORMATION TECHNOLOGY
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2025**

LEMBARAN PENGESAHAN

ANALISIS TRANSFER LEARNING INDOBERT DENGAN RANDOM FOREST PADA KLASIFIKASI SENTIMEN LAYANAN KESEHATAN BPJS

TESIS

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Magister
Program Studi S-2 Teknologi Informasi

Oleh:

STEVEN CIAM
NIM. 231232161

Disetujui Oleh:

Dosen Pembimbing I,


Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D.

Medan, 23 Agustus 2025
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-2 Teknologi Informasi


Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-2 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Steven Ciam

NIM : 231232161

Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul dan tempat penelitian sebagai berikut:

Judul Tesis *) : Analisis Transfer Learning IndoBERT Dengan Random Forest Pada Klasifikasi Sentimen Layanan Kesehatan BPJS

Tempat Penelitian :

Alamat Tempat Penelitian :

No. Telp. Tempat Penelitian :

Sehubungan dengan Tesis tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 06 Agustus 2025

Saya yang membuat pernyataan,

 
Steven Ciam

STRACT

Social media has become a primary platform for Indonesian citizens to express their opinions on public services such as BPJS Kesehatan, generating data rich in public perception. Sentiment analysis of these comments is essential for systematically understanding public opinion; however, accurate approaches like fine-tuning IndoBERT require substantial computational resources, which may not be readily available to all institutions. This study aims to compare two IndoBERT-based approaches full end-to-end fine-tuning and transfer learning using Random Forest for sentiment classification of Indonesian-language public comments. Using a labeled dataset of 3,059 comments, both models were evaluated based on accuracy, F1-score, and computational efficiency. The results show that the IndoBERT + Random Forest model achieved 98% accuracy and a weighted F1-score of 0.98, outperforming the fine-tuned model (92% and 0.92, respectively), while being significantly more time-efficient (33 seconds vs. 206 seconds). In addition to superior performance, the hybrid model offers greater interpretability and is more suitable for deployment in environments with limited computational infrastructure. These findings suggest that lightweight IndoBERT-based transfer learning strategies offer a practical solution for NLP-driven public opinion monitoring in Indonesia's public service sector.

Keywords: Sentiment Analysis, IndoBERT, Transfer Learning, Random Forest, Digital Public Ser

ABSTRAK

Media sosial telah menjadi sarana utama masyarakat Indonesia untuk menyampaikan opini mengenai layanan publik seperti BPJS Kesehatan, menghasilkan data yang kaya akan persepsi publik. Analisis sentimen terhadap komentar-komentar ini penting untuk memahami opini masyarakat secara sistematis, namun pendekatan akurat seperti fine-tuning IndoBERT memerlukan sumber daya komputasi besar yang belum tentu tersedia di semua institusi. Penelitian ini bertujuan membandingkan dua pendekatan IndoBERT *fine-tuning end-to-end* dan *transfer learning* dengan *Random Forest* dalam klasifikasi sentimen komentar publik berbahasa Indonesia. Dengan menggunakan 3.059 data komentar yang telah diberi label, kedua model dievaluasi berdasarkan akurasi, *F1-score*, dan efisiensi waktu. Hasil menunjukkan bahwa model IndoBERT + *Random Forest* mencapai akurasi 98% dan *F1-score* tertimbang 0.98, mengungguli model *fine-tuning* (92% dan 0.92), serta jauh lebih efisien dari sisi waktu (33 detik dengan 206 detik). Selain performa yang tinggi, model *hybrid* juga lebih mudah diinterpretasikan dan lebih layak diterapkan dalam sistem dengan keterbatasan komputasi. Temuan ini menunjukkan bahwa strategi *transfer learning* ringan berbasis IndoBERT dapat menjadi solusi praktis untuk pemantauan opini publik berbasis NLP di sektor layanan publik Indonesia.

Kata Kunci: Analisis Sentimen, IndoBERT, *Transfer Learning*, *Random Forest*, Layanan Publik Digital

KATA PENGANTAR

Penulis mengucapkan puji syukur kepada Tuhan Yang Maha Esa, atas karunia dan bimbingan-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Tesis ini berjudul **“Analisis Transfer Learning IndoBERT Dengan Random Forest Pada Klasifikasi Sentimen Layanan Kesehatan BPJS”** dan disusun sebagai syarat kelulusan untuk memperoleh Gelar Magister Komputer (M.Kom) tahap pendidikan Strata II di Program Studi Teknologi Informasi Universitas Mikroskil Medan.

Pada kesempatan ini, penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D. , selaku Dosen Pembimbing I.
2. Ibu Dr. Sofiana Nurjanah, S.Kom., M.T.I. , selaku Dosen Pembimbing II.
3. Bapak Hardy, S.Kom., M.SC., PH.D. ,selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.Kom., M.T.I ,selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D. ,selaku Ketua Program Studi S-2 Teknologi Informasi Fakultas Informatika Universitas Mikroskil Medan.
6. Para Dosen di Universitas Mikroskil, baik S-1 maupun S-2
7. Secara Khusus, penulis sampaikan terimakasih kepada Orang tua, Keluarga, famili dan sanak saudara penulis yang telah memberikan dukungan moril dan materil.
8. Secara Khusus, penulis sampaikan terimakasih kepada Dina Maria Sasnita Turnip yang telah memberikan dukungan moril dan materil.
9. Seluruh Teman-teman sekelas dan rekan-rekan mahasiswa/i Program Studi S-2 Teknologi Informasi yang selalu memberikan semangat dan bantuan dalam penulisan Tesis ini.

Dalam penyusunan Tesis ini, penulis menyadari sepenuhnya bahwa dalam penulisan skripsi ini masih terdapat banyak kekurangan dan masih jauh dari sempurna. Oleh karena itu, setiap saran yang bersifat membangun akan diterima penulis dengan senang hati.

Medan, 06 Agustus 2025

Penulis,


Steven Ciam

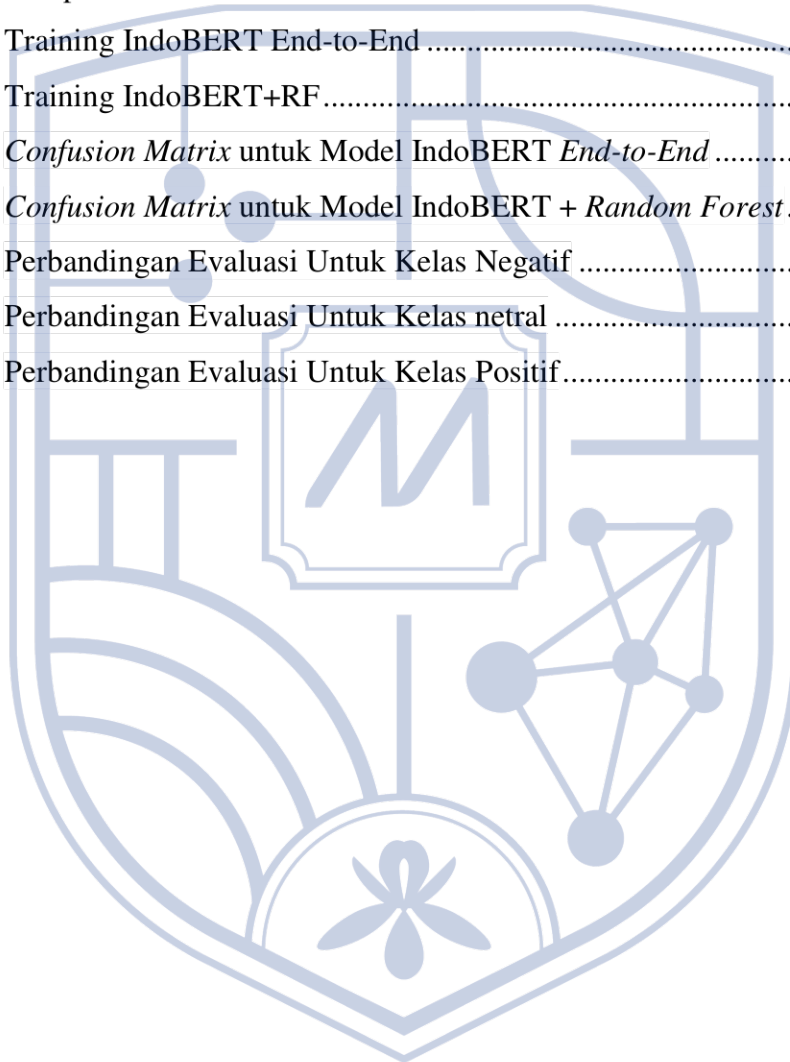
DAFTAR ISI

ABSTRACT.....	i
ABSTRAK.....	ii
KATA PENGANTAR	iii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan.....	3
1.4 Manfaat	3
1.5 Ruang Lingkup	4
BAB II KAJIAN LITERATUR.....	5
2.1 Tinjauan Pustaka	5
2.1.1 Analisis Sentimen	5
2.1.2 Pemrosesan Bahasa Alami (<i>Natural Language Processing / NLP</i>)	8
2.1.3 Transformers dan Model IndoBERT	9
2.1.4 Transfer Learning dalam NLP	12
2.1.5 <i>Random Forest</i>	13
2.2 Penelitian Terkait.....	15
2.3 Kerangka Konseptual.....	23
BAB III METODOLOGI PENELITIAN.....	26
3.1 Desain Penelitian	26
3.2 Pengumpulan Data	26
3.3 Preprocessing Data.....	27
3.4 Perancangan Model.....	27
3.5 Evaluasi Hasil	28
3.6 Implementasi Teknis.....	29
3.7 Perbandingan dan Analisis	29

3.8 Alur Proses Penelitian.....	30
BAB IV HASIL DAN PEMBAHASAN.....	33
4.1 Hasil.....	33
4.1.1 Pengumpulan Data.....	33
4.1.2 Preprocessing Data.....	34
4.1.3 Pelatihan Model.....	35
4.1.4 Evaluasi Model.....	38
4.2 Pembahasan.....	54
4.2.1 Interpretasi Hasil.....	55
4.2.2 Relevansi Hasil terhadap Tujuan Penelitian.....	57
4.2.3 Keterbatasan Penelitian.....	59
4.2.4 Hubungan dengan Studi Sebelumnya.....	62
BAB V KESIMPULAN DAN SARAN.....	66
5.1 Kesimpulan.....	66
5.2 Saran.....	67
DAFTAR PUSTAKA.....	69
DAFTAR RIWAYAT HIDUP.....	73

DAFTAR GAMBAR

Gambar 2. 1 Proses Analisis Sentimen.....	5
Gambar 2. 2 Proses NLP	8
Gambar 2. 3 Kerangka Konseptual.....	23
Gambar 3. 1 Alur Proses Penelitian	30
Gambar 4. 1 Tampilan Dataset Awal	34
Gambar 4. 2 Tampilan Dataset Akhir.....	35
Gambar 4. 3 Training IndoBERT End-to-End	37
Gambar 4. 4 Training IndoBERT+RF.....	38
Gambar 4. 5 <i>Confusion Matrix</i> untuk Model IndoBERT <i>End-to-End</i>	41
Gambar 4. 6 <i>Confusion Matrix</i> untuk Model IndoBERT + <i>Random Forest</i>	45
Gambar 4. 7 Perbandingan Evaluasi Untuk Kelas Negatif	48
Gambar 4. 8 Perbandingan Evaluasi Untuk Kelas netral	49
Gambar 4. 9 Perbandingan Evaluasi Untuk Kelas Positif.....	51



TABEL

Tabel 2. 1 Ringkasan Penelitian Terdahulu.....	15
Tabel 2. 2 Perbandingan Konseptual.....	25
Tabel 4. 1 Hasil Evaluasi Model IndoBERT <i>End-to-End</i>	39
Tabel 4. 2 Hasil Evaluasi Model IndoBERT + <i>Random Forest</i>	43
Tabel 4. 3 Perbandingan Metrik Evaluasi antara Dua Pendekatan.....	47
Tabel 4. 4 Perbandingan IndoBERT <i>End-to-End</i> vs. IndoBERT + <i>Random Forest</i>	55
Tabel 4. 5 Ringkasan Keterbatasan Penelitian	60
Tabel 4. 6 Hubungan Penelitian Ini dengan Studi Terdahulu.....	62

