



**OPTIMALISASI SP4N LAPOR! : INTEGRASI BERT DAN LSTM UNTUK
PENINGKATAN ANALISIS SENTIMEN PUBLIK**

LAPORAN TUGAS AKHIR

DIAZ THAUFIQURAHMAN
41521010088

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA

2025



**OPTIMALISASI SP4N LAPOR! : INTEGRASI BERT DAN LSTM UNTUK
PENINGKATAN ANALISIS SENTIMEN PUBLIK**

LAPORAN TUGAS AKHIR

DIAZ THAUFIQURAHMAN

41521010088

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 : Diaz Thaufiqurahman
NIM : 41521010088
Program Studi : Teknik Informatika
Judul Laporan Skripsi : “Optimalisasi SP4N LAPOR!: Integrasi BERT dan LSTM Untuk Peningkatan Analisis Sentimen Publik”

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, 19 Juli 2025

UNIVERSITAS
MERCU BUANA



Diaz Thaufiqurahman

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Diaz Thaufiqurahman
NIM : 41521010088
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Optimalisasi SP4N LAPOR!: Integrasi BERT dan LSTM Untuk Peningkatan Analisis Sentimen Publik

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 : Dwi Anindyani Rochmah, S.T.,
M.TI.
NIDN : 011057801
Ketua Penguji : Wawan Gunawan, S.Kom., M.T.,
M.Kom.
NIDN : 0424108104
Penguji 1 : Harni Kusniyati, S.T., M.Kom.
NIDN : 0324068101
Penguji 2 : Ida Farida, S.T., M.Kom.
NIDN : 0324018301

()

()

()

()

Jakarta, 19 Juli 2025

Mengetahui,

Dekan

Ketua Program Studi


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


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

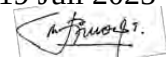
KATA PENGANTAR

Puji syukur kehadiran 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. Ibu Dwi Anindyani Rochmah, ST,MTI. selaku dosen pembimbing MPTI yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Ibu Harni Kusniyati, S.T., M.Kom selaku Dosen Penguji 1 dan Ibu Ida Farida, ST,M.Kom sebagai Dosen Penguji 2 Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana..
7. 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, 19 Juli 2025



Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : DIAZ THAUFIQURAHMAN
NIM : 41521010088
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Optimalisasi SP4N LAPOR!: Integrasi BERT dan LSTM Untuk Peningkatan Analisis Sentimen Publik

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, 19 Juli 2024

Yang menyatakan,

UNIVERSITAS
MERCU BUANA



Diaz Thaufiqurahman

ABSTRAK

Nama : Diaz Thaufiqurahman
NIM : 41521010088
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Optimalisasi SP4N LAPOR!: Integrasi BERT dan LSTM Untuk Peningkatan Analisis Sentimen Publik
Dosen Pembimbing : Dwi Anindyani Rochmah, ST,MTI

Pemerintah Indonesia terus mendorong transformasi pelayanan publik berbasis e-government melalui aplikasi SP4N LAPOR!. Namun, tantangan kualitas layanan dan dominasi keluhan pengguna masih menjadi isu utama, tercermin dari rating rendah serta ulasan negatif di Google Play Store. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi SP4N LAPOR! secara sistematis dan empiris menggunakan pendekatan deep learning kombinasi BERT dan LSTM. Data ulasan periode 2020–2025 dikumpulkan melalui web scraping, menghasilkan 1.376 data yang kemudian diproses melalui tahapan cleansing, tokenizing, normalisasi, stopword removal, dan stemming, serta penandaan spam berbasis heuristik tanpa menghapus data untuk menguji robustness model. Pelabelan sentimen otomatis menggunakan IndoBERT dari library Hugging Face menunjukkan dominasi sentimen negatif (62,2%), diikuti sentimen positif (21,7%) dan netral (16,1%). Proses pelatihan model dilakukan dengan batch size 16 selama maksimal 20 epoch dengan implementasi early stopping (patience=2) untuk mencegah overfitting. Regularisasi diterapkan melalui dropout pada LSTM dengan dropout rate 0.2, dan optimasi model menggunakan Adam optimizer dengan learning rate 0.001 dan weight decay 1e-5. Eksperimen pada varian dataset ALL DATA dan NON-SPAM dengan berbagai skenario pembagian data menunjukkan bahwa model BERT+LSTM konsisten meraih akurasi tinggi 93–94% pada ALL DATA dan hingga 95% pada NON-SPAM, dengan precision dan recall tertinggi mencapai 0.99 pada kelas negatif.

Kata kunci: Analisis Sentimen, SP4N LAPOR!, BERT, LSTM, Deep Learning, e-Government, Ulasan Pengguna, Early Stopping.

ABSTRACT

Nama : Diaz Thaufiqurahman
NIM : 41521010088
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Optimalisasi SP4N LAPOR!: Integrasi BERT dan LSTM Untuk Peningkatan Analisis Sentimen Publik
Dosen Pembimbing : Dwi Anindyani Rochmah, ST,MTI

The Indonesian government continues to encourage the transformation of e-government-based public services through the SP4N LAPOR! application. However, service quality challenges and the dominance of user complaints are still major issues, reflected in low ratings and negative reviews on the Google Play Store. This research aims to systematically and empirically analyze the sentiment of users of the SP4N LAPOR! application using a combination of BERT and LSTM deep learning approaches. Review data for the period 2020-2025 was collected through web scraping, resulting in 1,376 data which was then processed through the stages of cleansing, tokenizing, normalization, stopword removal, and stemming, as well as heuristic-based spam tagging without deleting data to test the robustness of the model. Automatic sentiment labeling using IndoBERT from the Hugging Face library shows the dominance of negative sentiment (62.2%), followed by positive sentiment (21.7%) and neutral (16.1%). The model training process was conducted with a batch size of 16 for a maximum of 20 epochs with the implementation of early stopping (patience=2) to prevent overfitting. Regularization is applied through dropout on the LSTM with a dropout rate of 0.2, and model optimization using Adam optimizer with a learning rate of 0.001 and weight decay of 1e-5. Experiments on variants of ALL DATA and NON-SPAM datasets with various data sharing scenarios show that the BERT+LSTM model consistently achieves high accuracy of 93-94% on ALL DATA and up to 95% on NON-SPAM, with the highest precision and recall reaching 0.99 on the negative class.

Keywords: Sentiment Analysis, SP4N LAPOR!, BERT, LSTM, Deep Learning, e-Government, User Reviews, Early Stopping.

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Penelitian.....	5
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terdahulu	7
2.2 Teori Utama	18
2.1.1 Analisis Sentimen.....	18
2.1.2 Text Mining.....	19
2.1.3 SP4N-LAPOR! (Sistem Pengelolaan Pengaduan Pelayanan Publik Nasional - Layanan Aspirasi dan Pengaduan Online Rakyat).....	19
2.1.4 Bidirectional Encoder Representations from Transformers (BERT).....	20

2.1.5 Long Short-Term Memory (LSTM).....	23
2.3 Teori Pendukung	25
2.2.1 Pelayanan Publik.....	25
2.2.2 Natural Language Processing (NLP)	25
2.2.3 Confusion Matrix	26
2.4 Gap Penelitian	28
2.4.1. Summarize.....	28
2.4.2 Synthesize	28
2.4.3 Comparison	28
2.4.4 Contribution	28
BAB III METODE PENELITIAN	29
3.1 Pendekatan Penelitian	29
3.2 Desain Penelitian.....	29
3.3 Subjek Penelitian.....	32
3.4 Instrumen Penelitian.....	32
3.5 Teknik Pengumpulan Data.....	33
3.6 Analisis Data	33
3.7 Prosedur Penelitian.....	38
3.7.1 Identifikasi Masalah	38
3.7.2 Studi Literatur	39
3.7.3 Pengumpulan Data	39
3.7.4 Preprocessing	39
3.7.5 Pelabelan Data.....	40
3.7.6 Pembagian Dataset	40
3.7.7 Pemodelan dengan Algoritma BERT dan LSTM	40
3.7.8 Evaluasi Model.....	41
3.7.9 Kesimpulan	42
3.8 Evaluasi Hasil Penelitian.....	42
BAB IV HASIL DAN PEMBAHASAN	43
4.1. Dataset.....	43

4.2. Preprocessing	44
4.2.1. Data Cleansing	44
4.2.2. Tokenizing.....	45
4.2.3. Normalization.....	46
4.2.4. Stopword Removing (Filtering)	46
4.2.5. Stemming	47
4.3. Identifikasi dan Filtrasi Ulasan Spam	47
4.4. Labelling dengan Hugging Face	49
4.5. Pembagian data	51
4.6. Pemodelan dengan BERT + LSTM	51
4.7. Evaluasi Kinerja Model.....	53
4.7.1 Evaluasi ALL DATA Data Split 70:30.....	54
4.7.2 Evaluasi ALL DATA Data Split 80:20.....	60
4.7.3 Evaluasi ALL DATA Data Split 90:10.....	66
4.7.4 Evaluasi NON-SPAM Data Split 70:30.....	72
4.7.5 Evaluasi NON-SPAM Data Split 80:20.....	78
4.7.6 Evaluasi NON-SPAM Data Split 90:10.....	84
4.7.7 Perbandingan Data Split.....	90
4.8. Visualisasi	91
4.8.1. Visualisasi Tren Sentimen.....	91
4.8.2. Frekuensi Sentimen Aplikasi pada Tahun 2020.....	94
4.8.3. Frekuensi Sentimen Aplikasi pada Tahun 2021.....	95
4.8.4. Frekuensi Sentimen Aplikasi pada Tahun 2022.....	96
4.8.5. Frekuensi Sentimen Aplikasi pada Tahun 2023.....	97
4.8.6. Frekuensi Sentimen Aplikasi pada Tahun 2024.....	98
4.8.7. Frekuensi Sentimen Aplikasi pada Tahun 2025.....	99
4.8.8. Frekuensi Sentimen Aplikasi sepanjang Tahun 2020-2025.....	100
4.7.8 Frekuensi Sentimen Berdasarkan Ulasan Aplikasi dan Layanan.....	101
BAB V PENUTUP.....	103
5.1 Kesimpulan	103
5.2 Saran.....	105

DAFTAR PUSTAKA.....	106
LAMPIRAN.....	110



DAFTAR TABEL

Tabel 2. 1 Penelitian terkait	7
Tabel 3. 1 Perangkat lunak.....	33
Tabel 3. 2 Data sentimen publik SP4N LAPOR!	34
Tabel 4. 1 Daftar data yang dihapus.....	44
Tabel 4. 2 Contoh hasil pelabelan dengan BERT	50
Tabel 4. 3 Tabel hasil perhitungan analisis per kelas (ALLDATA 70:30)	57
Tabel 4. 4 Tabel hasil perhitungan analisis per kelas (ALLDATA 80:20)	63
Tabel 4. 5 Tabel hasil perhitungan analisis per kelas (ALLDATA 90:10)	69
Tabel 4. 6 Tabel hasil perhitungan analisis per kelas (NON-SPAM 70:30)	75
Tabel 4. 7 Tabel hasil perhitungan analisis per kelas (NON-SPAM 80:20)	81
Tabel 4. 8 Tabel hasil perhitungan analisis per kelas (NON-SPAM 90:10)	87
Tabel 4. 9 Hasil Perbandingan Evaluasi Untuk Setiap Varian dan Data Split.....	90



DAFTAR GAMBAR

Gambar 1. 1 Rating Penilaian Layanan Aplikasi SP4N LAPOR!	2
Gambar 2. 1 Proses Text mining.....	19
Gambar 2. 2 Arsitektur Pre-training dan fine-tuning BERT.....	21
Gambar 2. 3 Arsitektur Transformers	21
Gambar 2. 4 Representasi masukan BERT	22
Gambar 2. 5 Arsitektur LSTM.....	23
Gambar 2. 6 Confusion Matrix	27
Gambar 3. 1 Tahapan penelitian	38
Gambar 4. 1 Dataset.....	43
Gambar 4. 2 Hasil Data Cleansing.....	44
Gambar 4. 3 Hasil Tokenizing	45
Gambar 4. 4 Hasil Normalization	46
Gambar 4. 5 Hasil Stopword Removing (Filtering).....	46
Gambar 4. 6 Hasil Stemming.....	47
Gambar 4. 7 Hasil Distribusi Ulasan Spam dan NON-Spam	48
Gambar 4. 8 Hasil Distribusi Labelling	50
Gambar 4. 9 Hasil Training Log dan Learning Curve Model LSTM (ALL DATA 70:30)	54
Gambar 4. 10 Hasil Evaluasi Klasifikasi Model (ALL DATA 70:30)	55
Gambar 4. 11 Hasil Training Log dan Learning Curve Model LSTM (ALL DATA 80:20)	60
Gambar 4. 12 Hasil Evaluasi Klasifikasi Model (ALL DATA 80:20)	61
Gambar 4. 13 Hasil Training Log dan Learning Curve Model LSTM (ALL DATA 90:10)	66
Gambar 4. 14 Hasil Evaluasi Klasifikasi Model (ALL DATA 90:10)	67
Gambar 4. 15 Hasil Training Log dan Learning Curve Model LSTM (NON-SPAM 70:30)	72
Gambar 4. 16 Hasil Evaluasi Klasifikasi Model (NON-SPAM 70:30)	73
Gambar 4. 17 Hasil Training Log dan Learning Curve Model LSTM (NON-SPAM 80:20)	78

Gambar 4. 18 Hasil Evaluasi Klasifikasi Model (NON-SPAM 80:20)	79
Gambar 4. 19 Hasil Training Log dan Learning Curve Model LSTM (NON-SPAM 90:10)	84
Gambar 4. 20 Hasil Evaluasi Klasifikasi Model (NON-SPAM 90:10)	85
Gambar 4. 21 Hasil Visualisasi Tren Sentimen	92
Gambar 4. 22 Pola frekuensi sentimen aplikasi pada tahun 2020	94
Gambar 4. 23 Pola frekuensi sentimen aplikasi pada tahun 2021	95
Gambar 4. 24 Pola frekuensi sentimen aplikasi pada tahun 2022	96
Gambar 4. 25 Pola frekuensi sentimen aplikasi pada tahun 2023	97
Gambar 4. 26 Pola frekuensi sentimen aplikasi pada tahun 2024	98
Gambar 4. 27 Pola frekuensi sentimen aplikasi pada tahun 2025	99
Gambar 4. 28 Pola frekuensi sentimen aplikasi sepanjang Tahun 2020-2025 ...	100
Gambar 4. 29 Pola frekuensi sentiment berdasarkan ulasan mengenai aplikasi dan layanan	101



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	110
Lampiran 2 Curriculum Vitae	111
Lampiran 3 Surat Pernyataan HAKI.....	112
Lampiran 4 Sertifikat BNSP	114
Lampiran 5 Form Revisi Dosen Penguji	115
Lampiran 6 Hasil Cek Turnitin	117
Lampiran 7 Halaman Persetujuan	118

