

**ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN PPN 12%
PADA MEDIA SOSIAL X DENGAN MODEL HIBRIDA INDOBERT–
BILSTM**

SKRIPSI

Oleh:

BRYANT ALFRONSO PURBA

NIM. 221113288

ROY JANNES SIMBOLON

NIM. 221113506

SAMUEL NATALINO SITORUS

NIM. 221111771



**PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

PUBLIC SENTIMENT ANALYSIS ON THE 12% VAT POLICY ON X SOCIAL MEDIA USING A HYBRID INDOBERT–BILSTM MODEL

UNDERGRADUATE THESIS

By:

BRYANT ALFRONSO PURBA

ID NUMBER. 221113288

ROY JANNES SIMBOLON

ID NUMBER. 221113506

SAMUEL NATALINO SITORUS

ID NUMBER. 221111771



**UNDERGRADUATE PROGRAM OF COMPUTER SCIENCE
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

**ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN PPN 12%
PADA MEDIA SOSIAL X DENGAN MODEL HIBRIDA INDOBERT-
BILSTM**

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

BRYANT ALFRONSO PURBA

NIM. 221113288

ROY JANNES SIMBOLON

NIM. 221113506

SAMUEL NATALINO SITORUS

NIM. 221111771

Disetujui Oleh:

Dosen Pembimbing I,

Dosen Pembimbing II,

Gunawan, S.Kom., M.T.I.

Irpan Adiputra Pardosi, S.Kom., M.TI.

Medan, 14 Juli 2026

Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika,



Carles Juliandy, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Bryant Alfronso Purba

NIM : 221113288

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

Judul Skripsi : Analisis Sentimen Publik terhadap Kebijakan PPN 12% pada Media Sosial X dengan Model Hibrida IndoBERT-BiLSTM

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi 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 Skripsi 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 Skripsi 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 Skripsi 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, 1 Juli 2026

Saya yang membuat pernyataan,



Bryant Alfronso Purba

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Roy Jannes Simbolon

NIM : 221113506

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

Judul Skripsi : Analisis Sentimen Publik terhadap Kebijakan PPN 12% pada Media Sosial X dengan Model Hibrida IndoBERT-BiLSTM

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi 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 Skripsi 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 Skripsi 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 Skripsi 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, 1 Juli 2026

Saya yang membuat pernyataan,


Roy Jannes Simbolon

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Samuel Natalino Sitorus

NIM : 221111771

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

Judul Skripsi : Analisis Sentimen Publik terhadap Kebijakan PPN 12% pada Media Sosial X dengan Model Hibrida IndoBERT-BiLSTM

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi 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 Skripsi 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 Skripsi 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 Skripsi 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, 1 Juli 2026

Saya yang membuat pernyataan,



Samuel Natalino Sitorus

ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN PPN 12% PADA MEDIA SOSIAL X DENGAN MODEL HIBRIDA INDOBERT-BILSTM

Abstrak

Kebijakan Pajak Pertambahan Nilai (PPN) sebesar 12% yang mulai berlaku pada Januari 2025 memicu perdebatan luas di media sosial X, sehingga diperlukan model yang mampu menganalisis dan mengklasifikasikan sentimen publik secara akurat dari data teks bebas berbahasa Indonesia yang berjumlah besar. Penelitian ini mengimplementasi model hibrida IndoBERT-BiLSTM melalui proses fine-tuning untuk mengklasifikasikan sentimen ke dalam tiga kategori, yaitu Negatif, Netral, dan Positif. Dataset terdiri dari 1.913 tweet yang dikumpulkan pada periode Oktober 2024 hingga Maret 2025 menggunakan kata kunci #TolakPPN12Persen, dianotasi oleh tiga anotator dengan nilai Fleiss' Kappa sebesar 0,9001, dan dibagi dengan rasio 80:10:10. Eksperimen dilakukan terhadap 15 konfigurasi dari tiga skenario penanganan ketidakseimbangan kelas, yaitu Baseline, Class Weight, dan Focal Loss, dengan F1-macro sebagai metrik evaluasi utama. Hasil terbaik diperoleh pada skenario Focal Loss konfigurasi K4 dengan Test Accuracy 87,50% dan Test F1-macro 83,63%, mengungguli skenario Baseline (81,48%) dan Class Weight (80,03%). Analisis sentimen menunjukkan dominasi sentimen Negatif sebesar 62,2% sepanjang periode pengamatan dengan puncak penolakan pada Desember 2024, bertepatan dengan pengumuman kepastian implementasi kebijakan. Model terbaik diimplementasikan ke dalam aplikasi web berbasis Streamlit untuk mendukung prediksi sentimen secara praktis dan interaktif.

Kata kunci: analisis sentimen, IndoBERT, BiLSTM, PPN 12%, media sosial

Abstract

The 12% Value Added Tax (VAT) policy enacted in January 2025 triggered widespread debate on social media platform X, necessitating a model capable of accurately analyzing and classifying public sentiment from large volumes of unstructured Indonesian text. This study implements a hybrid IndoBERT-BiLSTM model through fine-tuning to classify sentiment into three categories: negative, neutral, and positive. The dataset consists of 1,913 tweets collected between October 2024 and March 2025 using the keyword #TolakPPN12Persen, annotated by three annotators with a Fleiss' Kappa of 0.9001, and split at an 80:10:10 ratio. Experiments were conducted across 15 configurations derived from three class imbalance handling scenarios Baseline, Class Weight, and Focal Loss with F1-macro as the primary evaluation metric. The best performance was achieved by the Focal Loss scenario with configuration K4, yielding a Test Accuracy of 87.50% and Test F1-macro of 83.63%, outperforming Baseline (81.48%) and Class Weight (80.03%). Sentiment analysis reveals that negative sentiment dominates at 62.2% throughout the observation period, peaking in December 2024 coinciding with the official announcement of the policy implementation. The best model was deployed as a Streamlit-based web application to support practical and interactive sentiment prediction.

Keywords: sentiment analysis, IndoBERT, BiLSTM, VAT 12%, social media

KATA PENGANTAR

Ucapan syukur yang sedalam-dalamnya penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas limpahan kasih, penyertaan, serta anugerah-Nya sehingga skripsi yang berjudul "Analisis Sentimen Publik terhadap Kebijakan PPN 12% pada Media Sosial X dengan Model Hibrida IndoBERT–BiLSTM" dapat diselesaikan dengan baik.

Skripsi ini disusun sebagai salah satu persyaratan akademik untuk memperoleh gelar Sarjana pada Program Studi S-1 Teknik Informatika, Fakultas Informatika, Universitas Mikroskil. Selain sebagai bentuk pemenuhan syarat kelulusan, penyusunan skripsi ini juga menjadi sarana bagi penulis untuk menerapkan ilmu yang telah diperoleh selama masa perkuliahan, serta wawasan tambahan bagi para pembaca.

Proses penyusunan skripsi ini tidak terlepas dari berbagai tantangan. Namun, berkat doa, dukungan, bimbingan, serta bantuan dari berbagai pihak, setiap tahapan penelitian dapat diselesaikan dengan baik. Oleh karena itu, dengan penuh rasa hormat dan penghargaan, penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Gunawan, S.Kom., M.T.I., selaku Dosen Pembimbing I yang telah meluangkan waktu, memberikan arahan, bimbingan, masukan, serta motivasi kepada penulis selama proses penelitian dan penyusunan skripsi ini.
2. Bapak Irpan Adiputra Pardosi, S.Kom., M.TI., selaku Dosen Pembimbing II, atas kesediaan, kesabaran, serta perhatian dalam memberikan saran, koreksi, dan bimbingan yang sangat berarti sehingga skripsi ini dapat diselesaikan dengan baik.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil, yang telah memberikan kesempatan dan fasilitas kepada penulis untuk menempuh pendidikan hingga penyelesaian studi di Universitas Mikroskil.
4. Bapak Sunaryo Winardi, S.Kom., M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil, yang telah memberikan dukungan dan fasilitas akademik kepada penulis selama menempuh pendidikan hingga penyelesaian skripsi ini.
5. Bapak Carles Juliandy, S.Kom., M.Kom., selaku Ketua Program Studi S-1 Teknik Informatika Fakultas Informatika Universitas Mikroskil, atas dukungan, arahan, dan pelayanan akademik yang telah diberikan selama penulis menjalani proses perkuliahan.
6. Bapak Kastro E. Purba dan Ibu T. E. Arta Novelia Pakpahan selaku orang tua yang senantiasa memberikan doa, kasih sayang, motivasi, perhatian, serta dukungan baik

secara moral maupun material selama proses perkuliahan hingga penyelesaian skripsi ini. Ucapan terima kasih juga disampaikan kepada saudara kandung, yaitu Kevin Denovan Purba dan Luis Enrico Purba, atas semangat dan dukungan yang selalu diberikan. Selain itu, penulis mengucapkan terima kasih kepada rekan satu tim penyusunan skripsi, yaitu Roy Jannes Simbolon dan Samuel Natalino Sitorus, atas kerja sama, dedikasi, serta kontribusi yang diberikan selama proses penelitian dan penyusunan skripsi.

7. Bapak Sabar Simbolon dan Ibu Loide Simamora selaku orang tua yang senantiasa memberikan doa, kasih sayang, motivasi, serta dukungan tanpa henti selama masa perkuliahan hingga penyelesaian skripsi ini. Penulis juga mengucapkan terima kasih kepada saudara kandung, yaitu Sabetria Simbolon, Albert Simbolon, Zuliver Simbolon dan Tio Bennita Simbolon, atas perhatian, dukungan, dan semangat yang selalu diberikan. Penghargaan yang sama juga disampaikan kepada rekan satu tim penyusunan skripsi, yaitu Bryant Alfronso Purba dan Samuel Natalino Sitorus, atas kerja sama, kebersamaan, serta kontribusi yang telah diberikan dalam menyelesaikan penelitian ini.
8. Bapak Darwin Sitorus dan Ibu Fransiska Diah selaku orang tua yang senantiasa mendoakan, memberikan kasih sayang, dukungan, serta motivasi kepada penulis dalam setiap proses penyelesaian studi. Ucapan terima kasih juga disampaikan kepada saudara kandung, yaitu Gabriella Agusta dan Moses Oktavino, atas doa dan dukungan yang selalu menyertai. Selain itu, penulis menyampaikan apresiasi kepada rekan satu tim penyusunan skripsi, yaitu Bryant Alfronso Purba dan Roy Jannes Simbolon, atas kerja sama, kekompakan, serta kontribusi yang diberikan selama proses penelitian dan penyusunan skripsi ini.
9. Seluruh dosen dan tenaga kependidikan Universitas Mikroskil, yang telah memberikan ilmu pengetahuan, pengalaman, bimbingan, serta pelayanan akademik yang sangat bermanfaat kepada penulis selama menempuh pendidikan.

Penulis menyadari bahwa skripsi ini masih memiliki berbagai kekurangan dan keterbatasan. Oleh sebab itu, kritik dan saran yang bersifat membangun sangat diharapkan sebagai bahan evaluasi untuk penyempurnaan karya ilmiah ini di masa yang akan datang.

Akhir kata, penulis berharap semoga skripsi ini dapat memberikan manfaat, menambah wawasan, serta menjadi salah satu referensi yang berguna bagi pengembangan ilmu pengetahuan.

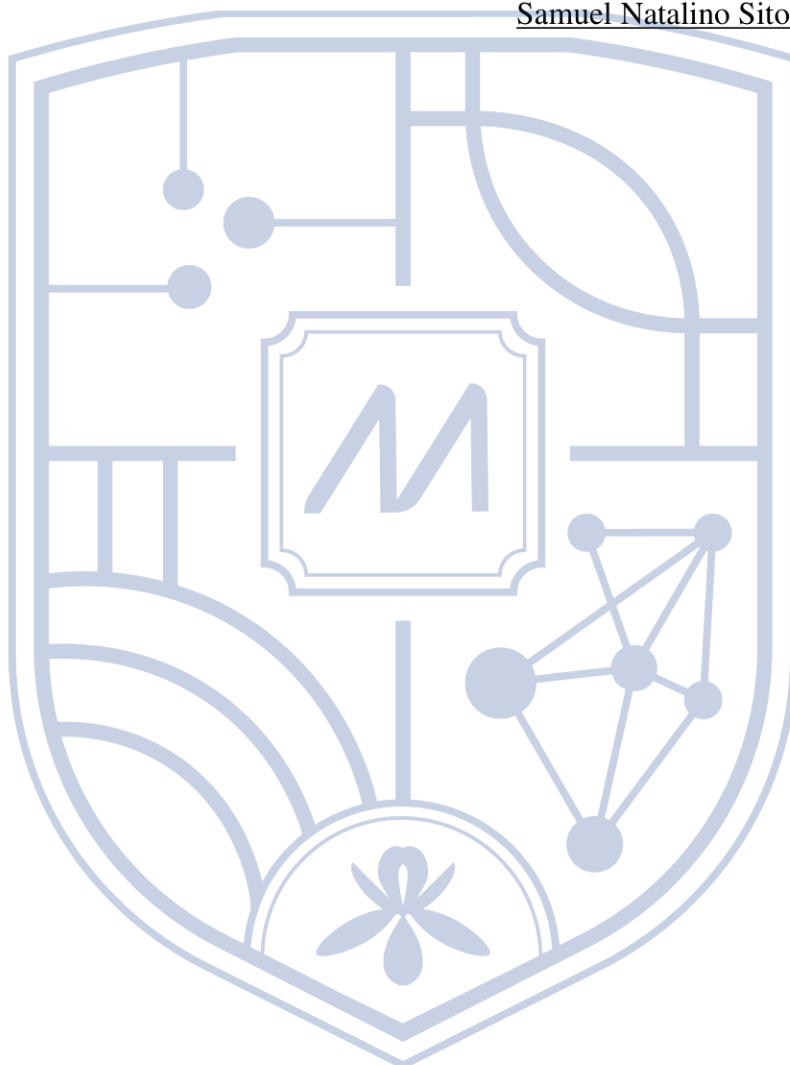
Medan, 1 Juli 2026

Penulis,

Bryant Alfronso Purba

Roy Jannes Simbolon

Samuel Natalino Sitorus



DAFTAR ISI

Abstrak	i
KATA PENGANTAR	ii
DAFTAR ISI	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
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.....	3
BAB II KAJIAN LITERATUR.....	5
2.1 Kebijakan Pajak Pertambahan Nilai (PPN) 12%	5
2.2 Media Sosial.....	6
2.3 Platform X (Twitter)	7
2.4 Kecerdasan Buatan.....	8
2.5 Pemrosesan Bahasa Alami (<i>Natural Language Processing</i>)	9
2.6 Analisis Sentimen	9
2.7 Pembelajaran Mesin Tradisional.....	10
2.8 Pembelajaran Mendalam (<i>Deep Learning</i>)	11
2.8.1 BERT dan IndoBERT	11
2.8.2 LSTM dan BiLSTM.....	14
2.8.3 Tokenisasi dan <i>Padding</i>	16
2.8.4 <i>Fine Tunning</i>	17
2.8.5 Hyperparameter.....	18
2.8.6 Model Hibrida IndoBERT–BiLSTM	19
2.9 <i>Data Crawling</i>	22
2.9.1 <i>Tweets Harvest</i>	23
2.10 Pelabelan Data Manual	23
2.10.1 Inter-Annotator Agreement dan Fleiss Kappa	24
2.11 <i>Exploratory Data Analysis</i> (EDA).....	26
2.12 Prapemrosesan Data (<i>Text Preprocessing</i>)	27

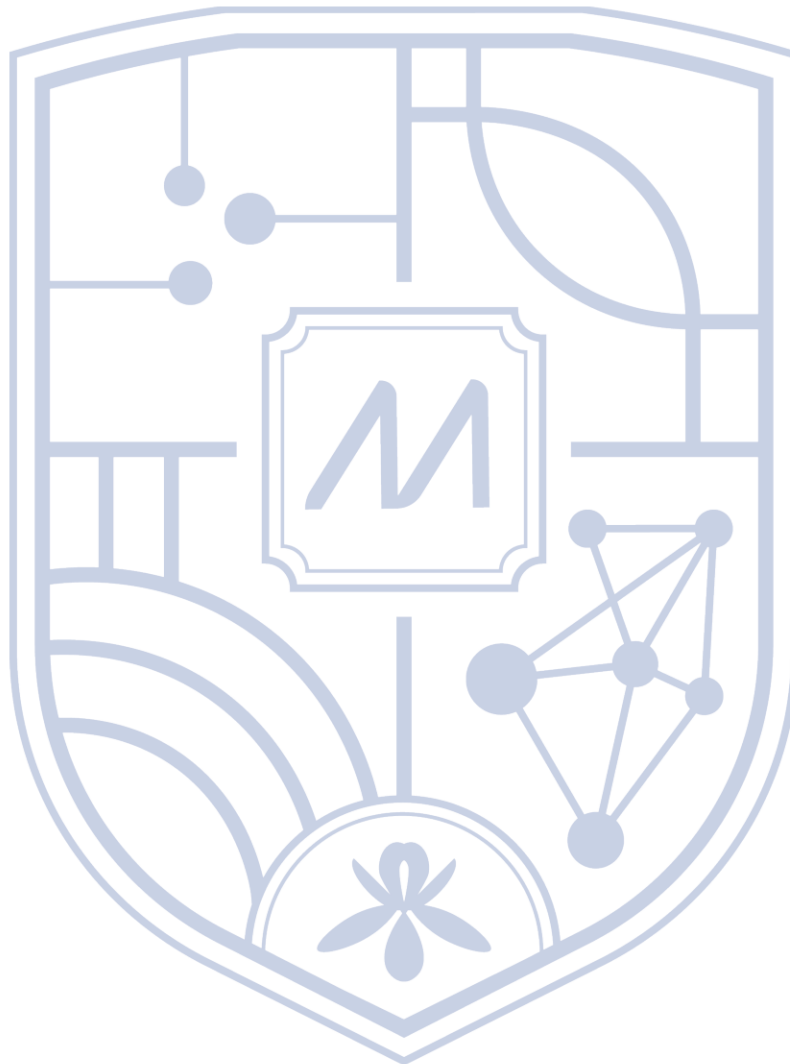
2.12.1 <i>Case Folding</i>	28
2.12.2 <i>Noise Removal</i>	28
2.12.3 <i>Demojize</i>	29
2.12.4 <i>Handling Repetitive Characters</i>	29
2.12.5 <i>Slang Transformation</i>	30
2.13 Penanganan Ketidakseimbangan Kelas (<i>Class Imbalance</i>)	30
2.13.1 <i>Class Weight</i>	31
2.13.2 <i>Focal Loss</i>	31
2.14 Pembagian Data (<i>Data Splitting</i>)	32
2.15 Evaluasi Performa Model	33
2.15.1 <i>Confusion matrix</i>	33
2.15.2 Metrik Evaluasi	34
2.16 Python dan <i>Library</i> Pendukung	35
2.17 Penelitian Terdahulu	36
BAB III TAHAPAN PELAKSANAAN	40
3.1 Kerangka Tahapan Pelaksanaan	40
3.2 Pengumpulan Dataset	42
3.3 Pelabelan Data (<i>Labeling</i>)	44
3.4 Prapemrosesan Data (<i>Text Preprocessing</i>)	46
3.5 Analisis Data Eksploratif	63
3.6 Penanganan Ketidakseimbangan Kelas (<i>Class Imbalance Handling</i>)	64
3.7 Pembagian Dataset	66
3.8 Perancangan dan Implementasi Model Hibrida IndoBERT-BiLSTM	67
3.8.1 Tokenisasi dan Padding	68
3.8.2 Ekstraksi Fitur	73
3.8.3 Arsitektur Model	80
3.9 Fine-tuning	81
3.10 Evaluasi Model	82
3.11 Analisis Kebutuhan Sistem	85
3.11.1 Kebutuhan Fungsional	85
3.11.2 Kebutuhan Non-Fungsional	88
3.12 Perancangan <i>Website</i>	89
3.12.1 Navigasi <i>Website</i>	90
3.12.2 Perancangan Antarmuka Pengguna (<i>User Interface</i>)	90
BAB IV HASIL DAN PEMBAHASAN	95

4.1 Hasil	95
4.1.1 Hasil Pelatihan Model	97
4.1.2 Hasil Pengujian dan Evaluasi Model	102
4.1.3 Hasil Pengujian dan Evaluasi Menggunakan Data Lapangan.....	110
4.1.4 Perbandingan Antar Skenario	117
4.1.5 Implementasi Antarmuka Website.....	121
4.2 Pembahasan.....	126
4.2.1 Evaluasi Kinerja Model Klasifikasi Sentimen	127
4.2.2 Kemampuan Model terhadap Konteks Semantik dan Hubungan Sekuensial...	127
4.2.3 Analisis Sentimen Publik terhadap Kebijakan PPN 12%	129
4.2.4 Implikasi Temuan Penelitian	134
BAB V PENUTUP	136
5.1 Kesimpulan	136
5.2 Saran.....	136
DAFTAR PUSTAKA	138
DAFTAR RIWAYAT HIDUP	146
DAFTAR RIWAYAT HIDUP	148
DAFTAR RIWAYAT HIDUP	149

DAFTAR GAMBAR

Gambar 2.1 Ragam Platform Media Sosial [35]	7
Gambar 2.2 Platform X [36].....	7
Gambar 2.3 Arsitektur Model BERT[43].....	11
Gambar 2.4 Arsitektur Model IndoBERT [45]	12
Gambar 2.5 Arsitektur Model LSTM [43]	14
Gambar 2.6 Arsitektur Model BiLSTM [45]	15
Gambar 2.7 Tahapan Model Hibrida IndoBERT-BiLSTM [18][26]	20
Gambar 3.1 Flowchart Tahapan Pelaksanaan.....	41
Gambar 3.2 Tahapan Prapemrosesan Data.....	47
Gambar 3.3 Tahapan Implementasi Model IndoBERT-BiLSTM	68
Gambar 3.4 Proses Tokenisasi Dan Padding.....	73
Gambar 3.5 Use Case Diagram Sistem Analisis Sentimen Berbasis Web.....	86
Gambar 3.6 Tampilan Halaman <i>Input Text</i>	91
Gambar 3.7 Tampilan Upload File dan Preview Dataset	92
Gambar 3.8 Tampilan Hasil Analisis Dataset	93
Gambar 3.9 Tampilan Insight dan Visualisasi Dataset.....	94
Gambar 4.1 Kurva Pelatihan Konfigurasi Terbaik Baseline (K1).....	99
Gambar 4.2 Kurva Pelatihan Konfigurasi Terbaik Class Weight (K4).....	100
Gambar 4.3 Kurva Pelatihan Konfigurasi Terbaik Focal Loss (K4).....	102
Gambar 4.4 <i>Confusion matrix</i> Konfigurasi Terbaik Baseline (K1)	104
Gambar 4.5 <i>Confusion matrix</i> Konfigurasi Terbaik Class Weight (K4)	107
Gambar 4.6 <i>Confusion matrix</i> Konfigurasi Terbaik Focal Loss (K4).....	109
Gambar 4.7 Perbandingan Test Accuracy dan F1-Macro Model Terbaik Setiap Skenario	118
Gambar 4.8 F1-Score per Kelas Sentimen Model Terbaik Setiap Skenario	119
Gambar 4.9 Halaman Input Teks.....	121
Gambar 4.10 Hasil Prediksi Sentimen Teks.....	122
Gambar 4.11 Halaman Upload Dataset	123
Gambar 4.12 Preview Dataset yang Berhasil di Unggah	123
Gambar 4.13 Hasil Klasifikasi Dataset.....	124
Gambar 4.14 Halaman Insight dan Visualisasi Sentimen	125

Gambar 4.15 Halaman Word Cloud dan Top Words	126
Gambar 4.16 Frekuensi Kata Setiap Kelas	130
Gambar 4.17 Visualisasi Word Cloud	131
Gambar 4.18 Analisis Bigram	131
Gambar 4.19 Analisis Trigram	132
Gambar 4.20 Analisis Tren Sentimen Berdasarkan Waktu	133



DAFTAR TABEL

Tabel 2.1 Nilai Interpretasi Tingkat Kesepakatan [61]	26
Tabel 2.2 Aturan Proses Case Folding [54].....	28
Tabel 2.3 Komponen Noise Removal [54].....	28
Tabel 2.4 Contoh Transformasi Emoji dengan Proses Demojize [55]	29
Tabel 2.5 Penanganan Karakter Berulang [54]	30
Tabel 2.6 Contoh Normalisasi Kata Tidak Baku [11]	30
Tabel 2.7 Representasi <i>Confusion matrix</i> 3x3 [63]	33
Tabel 2.8 Penelitian Terdahulu.....	39
Tabel 3.1 Hasil Crawling Data	44
Tabel 3.2 Hasil Kesepakatan Anotator	45
Tabel 3.3 Distribusi Label Hasil Pelabelan Dataset	45
Tabel 3.4 Contoh Hasil Pelabelan per Kelas Sentimen.....	45
Tabel 3.5 Hasil Case Folding	48
Tabel 3.6 Penghapusan URL.....	49
Tabel 3.7 Penghapusan Mention (@).....	51
Tabel 3.8 Penghapusan Hashtag.....	53
Tabel 3.9 Penghapusan HTML Entity	55
Tabel 3.10 Penghapusan Tanda Baca dan Karakter Khusus	56
Tabel 3.11 Demojize.....	58
Tabel 3.12 Penerapan Penanganan Karakter Berulang	60
Tabel 3.13 Penerapan Slang Transformation	62
Tabel 3.14 Hasil Perhitungan <i>Class Weight</i>	65
Tabel 3.15 Perbandingan CE Loss vs Focal Loss ($\gamma=2,0$) untuk Berbagai Nilai Probabilitas	66
Tabel 3.16 Hasil Pembagian Dataset.....	67
Tabel 3.17 Hasil Tokenisasi IndoBERT WordPiece	68
Tabel 3.18 Contoh Hasil <i>Padding</i>	71
Tabel 3.19 Output Tokenizer.....	72
Tabel 3.20 Nilai Embedding Token (5 dari 768 Dimensi)	74
Tabel 3.21 Hasil Output BiLSTM (5 dari 1024 Dimensi).....	75
Tabel 3.22 Hasil raw attention score	76

Tabel 3.23 Bobot Attention Setiap Token	78
Tabel 3.24 Hasil Logits Klasifikasi Sentimen	79
Tabel 3.25 Detail Arsitektur Model IndoBERT-BiLSTM	81
Tabel 3.26 Analisis Kebutuhan Fungsional Sistem	86
Tabel 3.27 Skenario Use Case Upload Dataset	87
Tabel 3.28 Skenario Use Case Upload Dataset	87
Tabel 3.29 Analisis Kebutuhan Non-Fungsional Sistem	88
Tabel 3.30 Struktur Navigasi Sistem	90
Tabel 3.31 Komponen dan Fungsi Tampilan Halaman Input Text	90
Tabel 3.32 Komponen dan Fungsi Tampilan Upload File dan Preview Dataset	91
Tabel 3.33 Komponen dan Fungsi Tampilan Hasil Analisis Dataset	92
Tabel 3.34 Komponen dan Fungsi Tampilan Insight dan Visualisasi	93
Tabel 4.1 Konfigurasi Eksperimen Hyperparameter [82]	95
Tabel 4.2 Ringkasan Seluruh Hasil Eksperimen	96
Tabel 4.3 Hasil Eksperimen Skenario Baseline	97
Tabel 4.4 Hasil Eksperimen Skenario Class Weight	99
Tabel 4.5 Hasil Eksperimen Skenario Focal Loss	101
Tabel 4.6 Sample Data Uji Skenario Baseline	102
Tabel 4.7 Classification Report Konfigurasi Terbaik Baseline (K1)	105
Tabel 4.8 Sampel Data Uji Skenario Class Weight	105
Tabel 4.9 Classification Report Konfigurasi Terbaik Class Weight (K4)	107
Tabel 4.10 Sampel Uji Data Skenario Focal Loss	108
Tabel 4.11 Classification Report Konfigurasi Terbaik Focal Loss (K4)	110
Tabel 4.12 Pengujian dengan Teks Baru	111
Tabel 4.13 Pengujian Model terhadap Berbagai Karakteristik Bahasa pada Data Teks ...	115
Tabel 4.14 Perbandingan Model Terbaik Skenario	118
Tabel 4.15 Perbandingan F1-Score per Kelas Sentimen Antar Skenario	119
Tabel 4.16 Statistik Panjang Teks per Kelas	129
Tabel 4.17 Volume <i>Tweet</i> per Kelas Berdasarkan Waktu	133