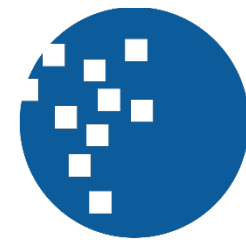


**Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan
IndoBERTweet: Perbandingan Kinerja dengan Model Klasik**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Skripsi

Ezra Mathian Lomar

00000059178

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan
IndoBERTweet: Perbandingan Kinerja dengan Model Klasik**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

Ezra Mathian Lomar

00000059178

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Ezra Mathian Lomar

Nomor Induk Mahasiswa : 00000059178

Program Studi : Sistem Informasi

Skripsi dengan judul:

Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan IndoBERTweet:

Perbandingan Kinerja dengan Model Klasik

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 28 November 2025



Ezra Mathian Lomar

UNIVERSITAS
MULTIMEDIA
NUSANTARA

LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Ezra Mathian Lomar
NIM : 00000059178
Program Studi : Sistem Informasi
Judul Skripsi : Analisis Sentimen Pelanggan Fore Coffee di Twitter
dengan IndoBERTweet: Perbandingan Kinerja
dengan Model Klasik

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas/Project/Tugas Akhir*(coret salah satu) sebagai berikut (beri tanda centang yang sesuai):

- ☒ Menggunakan AI sebagaimana diizinkan untuk membantu dalam menghasilkan ide-ide utama saja
- ☐ Menggunakan AI sebagaimana diizinkan untuk membantu menghasilkan teks pertama saja
- ☐ Menggunakan AI untuk menyempurnakan sintaksis dan tata bahasa untuk pengumpulan tugas
- ☐ Karena tidak diizinkan: Tidak menggunakan bantuan AI dengan cara apa pun dalam pembuatan tugas

Saya juga menyatakan bahwa:

- (1) Menyerahkan secara lengkap dan jujur penggunaan perangkat AI yang diperlukan dalam tugas melalui Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)
- (2) Saya mengakui bahwa saya telah menggunakan bantuan AI dalam tugas saya baik dalam bentuk kata, paraphrase, penyertaan ide atau fakta penting yang disarankan oleh AI dan saya telah menyantumkan dalam sitasi serta referensi

Terlepas dari pernyataan di atas, tugas ini sepenuhnya merupakan karya saya sendiri

Tangerang, 28 November 2025
Yang Menyatakan,



(Ezra Mathian Lomar)

HALAMAN PERSETUJUAN

Skripsi dengan judul

Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan IndoBERTweet:

Perbandingan Kinerja dengan Model Klasik

Oleh

Nama : Ezra Mathian Lomar
NIM : 00000059178
Program Studi : Sistem Informasi
Fakultas : Teknik Informatika

Telah disetujui untuk diajukan pada
Sidang Ujian Skripsi Universitas Multimedia Nusantara
Tangerang, 28 November 2025

Pembimbing



Dr. Irmawati, S.Kom., M.M.S.I.
NIDN: 0805097703/NIK: 081431

Ketua Program Studi Sistem Informasi



Ririn Ikana Desanti, S.Kom., M.Kom.
NIDN: 0313058001/NIK: 051314

UMIN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan IndoBERTweet:
Perbandingan Kinerja dengan Model Klasik

Oleh

Nama : Ezra Mathian Lomar
NIM : 00000059178
Program Studi : Sistem Informasi
Fakultas : Teknik Informatika

Telah diujikan pada hari Kamis, 11 Desember 2025

Pukul 15.30 s.d 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang



Ririn Ikana Desanti, S.Kom., M.Kom.
NIDN: 0313058001/NIK: 051314

Penguji



Prof. Dr. Friska Natalia, S.Kom., M.T., Ph.D.,
NIDN: 0306056102

Pembimbing



Dr. Irmawati, S.Kom., M.M.S.I.
NIDN: 0805097703/NIK: 081431

Ketua Program Studi Sistem Informasi



Ririn Ikana Desanti, S.Kom., M.Kom.
NIDN: 0313058001/NIK: 051314

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Ezra Mathian Lomar
NIM : 00000059178
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Pelanggan Fore Coffee di
Twitter dengan IndoBERTweet: Perbandingan
Kinerja dengan Model Klasik

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih salah satu):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 28 November 2025



(Ezra Mathian Lomar)

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat, rahmat, dan karunia-Nya sehingga saya dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan IndoBERTweet: Perbandingan Kinerja dengan Model Klasik.”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata 1 (S1) pada Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Saya menyadari bahwa penyusunan tugas akhir ini tidak terlepas dari bimbingan, dukungan, serta bantuan dari berbagai pihak yang telah berperan penting dalam proses penelitian hingga terselesaikannya skripsi ini. Oleh karena itu, pada kesempatan ini saya ingin menyampaikan rasa terima kasih dan penghargaan yang sebesar-besarnya kepada:

1. Bapak Dr. Andrey Andoko, M.Sc., Ph.D., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Dr. Irmawati, S.Kom., M.M.S.I., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan moral dan material yang tidak pernah putus selama proses penyusunan skripsi ini sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga penulisan karya ilmiah ini dapat memberikan manfaat, wawasan, dan inspirasi bagi pembaca serta dapat menjadi referensi bagi penelitian selanjutnya

Tangerang, 28 November 2025



Ezra Mathian Lomar

Analisis Sentimen Pelanggan Fore Coffee di Twitter dengan IndoBERTweet: Perbandingan Kinerja dengan Model Klasik

Ezra Mathian Lomar

ABSTRAK

Perkembangan media sosial seperti Twitter telah memberikan wadah bagi pelanggan untuk menyampaikan opini dan pengalaman mereka terhadap suatu produk atau layanan secara terbuka. Informasi tersebut dapat dimanfaatkan oleh perusahaan untuk mengetahui persepsi pelanggan terhadap mereknya melalui analisis sentimen. Fore Coffee sebagai salah satu merek kopi lokal Indonesia juga banyak menerima ulasan dari pengguna Twitter, namun data yang bersifat tidak terstruktur dan menggunakan bahasa informal memerlukan pendekatan analisis berbasis machine learning dan deep learning. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja beberapa algoritma dalam melakukan analisis sentimen terhadap ulasan pelanggan Fore Coffee di Twitter. Data dikumpulkan menggunakan tweet-harvest sebanyak 10.426 tweet dan melalui tahapan preprocessing seperti case folding, cleansing, stopword removal, dan stemming. Pelabelan sentimen dilakukan secara otomatis menggunakan metode keyword-based lexicon. Penelitian ini menggunakan metode CRISP-DM hingga tahap evaluation dengan menerapkan empat algoritma, yaitu Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), dan IndoBERTweet. Hasil penelitian menunjukkan bahwa IndoBERTweet memberikan hasil terbaik dengan akurasi 91,46% dan F1-macro 0,88, melampaui model klasik seperti SVM (87,58%), Logistic Regression (86,05%), dan Naïve Bayes (80,68%). Analisis sentimen menunjukkan bahwa ulasan pelanggan Fore Coffee di Twitter didominasi oleh sentimen positif sebesar 58,24%, diikuti netral 24,93%, dan negatif 16,83%. Dengan demikian, IndoBERTweet terbukti paling efektif untuk analisis sentimen berbahasa Indonesia di Twitter dan menunjukkan citra positif pelanggan terhadap Fore Coffee di media sosial.

Kata Kunci: Analisis Sentimen, Machine Learning, IndoBERTweet, Fore Coffee, Twitter

Fore Coffee Customer Sentiment Analysis on Twitter with IndoBERTweet: Performance Comparison with Classical Model

Ezra Mathian Lomar

ABSTRACT (English)

The rapid growth of social media platforms such as Twitter provides a medium for customers to express their opinions and experiences regarding products and services openly. These opinions can be analyzed to understand customer perceptions of a brand through sentiment analysis. Fore Coffee, as one of Indonesia's leading local coffee brands, receives numerous reviews on Twitter; however, the unstructured and informal nature of the data requires a computational approach using machine learning and deep learning methods. Therefore, this study aims to compare the performance of several algorithms in classifying customer sentiment toward Fore Coffee on Twitter. Data were collected using Tweet Harvest, totaling 10,426 tweets, and processed through several preprocessing stages, including case folding, cleansing, stopword removal, and stemming. Sentiment labeling was performed automatically using a keyword-based lexicon method. The research employed the CRISP-DM methodology up to the evaluation stage, applying four algorithms: Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), and IndoBERTweet. The results indicate that IndoBERTweet achieved the highest performance with 91.46% accuracy and an F1-macro score of 0.88, outperforming classical models such as SVM (87.58%), Logistic Regression (86.05%), and Naïve Bayes (80.68%). Sentiment distribution analysis revealed that customer reviews were dominated by positive sentiments (58.24%), followed by neutral (24.93%) and negative (16.83%). Thus, IndoBERTweet is proven to be the most effective model for Indonesian-language sentiment analysis, indicating that customer perception toward Fore Coffee on Twitter is generally positive.

Keywords: *Sentiment Analysis, Machine Learning, IndoBERTweet, Fore Coffee, Twitter*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN.....	iii
KATA PENGANTAR.....	vii
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian.....	5
1.4.1 Tujuan Penelitian	5
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan	5
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu	8
2.2 Analisis Gap Penelitian Terdahulu.....	14
2.3 Teori yang berkaitan.....	14
2.3.1 Analisis Sentimen	14
2.3.2 Twitter sebagai Sumber Data	15
2.4 Framework dan Algoritma yang digunakan	15
2.4.1 Cross-Industry Standard Process for Data Mining (CRISP-DM)	16
2.4.2 Naïve Bayes	18
2.4.3 Support Vector Machine (SVM).....	19
2.4.4 Logistic Regression.....	20
2.4.5 IndoBERTweet	22
2.5 Tools dan Software yang digunakan	23

BAB III METODOLOGI PENELITIAN	27
3.1 Gambaran Umum Objek Penelitian.....	27
3.1.1 Ulasan Pelanggan Fore Coffee	27
3.2 Metode Penelitian.....	28
3.2.1 Alur Penelitian.....	31
3.3 Teknik Pengumpulan Data.....	34
3.4 Teknik Pengolahan Data	35
3.5 Teknik Pengujian dan Evaluasi.....	37
BAB IV ANALISIS DAN HASIL PENELITIAN	42
4.1 Analisa Masalah	42
4.1.1 Analisa Proses Bisnis.....	42
4.1.2 Tujuan Bisnis.....	43
4.2 Data Collection	44
4.3 Data Preparation	46
4.3.1 Case Folding dan Text Cleaning	46
4.3.2 Stopword Removal dan Stemming.....	48
4.3.3 Sentiment Labeling.....	52
4.3.4 Tahap Splitting dan Penerapan Cost-Sensitive Learning.....	55
4.3.5 TF-IDF Vectorization dan Representasi Data	60
4.4 Modeling.....	62
4.4.1 Hasil dan Analisis Model Naïve Bayes.....	62
4.4.2 Hasil dan Analisis Model SVM / LinearSVC.....	65
4.4.3 Hasil dan Analisis Model Logistic Regression	69
4.4.4 Hasil dan Analisis Model IndoBERTweet.....	72
4.4.4.1 Distribusi Sentimen Hasil IndoBERTweet.....	77
4.5 Perbandingan Hasil dan Pengujian Model.....	78
BAB V SIMPULAN DAN SARAN	80
5.1 Simpulan	80
5.2 Rekomendasi.....	81
5.3 Saran.....	82
DAFTAR PUSTAKA	84
LAMPIRAN.....	87

DAFTAR TABEL

Table 2.1 Penelitian Terdahulu	12
Table 4.1 Contoh Proses Cleaning	47
Table 4.2 Contoh Stopword Removal	49
Table 4.3 Contoh Stemming	50



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	32
Gambar 4.1 Tahap awal pengumpulan data Twitter menggunakan Tweet Harvest di Google Colab	45
Gambar 4.2 Tahap Text Cleaning dan Case Folding	46
Gambar 4.3 Hasil Setelah Text Cleaning dan Case Folding.....	48
Gambar 4.4 Stopword Removal dan Stemming.....	51
Gambar 4.5 Hasil Setelah Stopword Removal dan Stemming	52
Gambar 4.6 Hasil pelabelan otomatis menggunakan pendekatan <i>keyword-based sentiment labeling</i>	54
Gambar 4.7 Distribusi sentiment tweet ulasan pelanggan Fore Coffee.....	55
Gambar 4.8 Konfigurasi Kolom untuk model Machine Learning dan IndoBERTweet.....	57
Gambar 4.9 Proses Splitting Data	58
Gambar 4.10 Distribusi Label Train vs Test.....	58
Gambar 4.11 Setup Cost-Sensitive Learning pada Data Training	59
Gambar 4.12 TF - IDF	61
Gambar 4.13 Hasil MultinomialNB.....	63
Gambar 4.14 Confusion Matrix Naïve Bayes	64
Gambar 4.15 Hasil Model LinearSVC / SVM.....	66
Gambar 4.16 Confusion Matrix LinearSVC / SVM	67
Gambar 4.17 Hasil Evaluasi Model Logistic Regression	69
Gambar 4.18 Confusion Matrix Logistic Regression	71
Gambar 4.19 Hasil IndoBERTweet	74
Gambar 4.20 Confusion Matrix IndoBERTweet	75
Gambar 4.21 Distribusi Sentimen Pelanggan Fore Coffee di Twitter	77
Gambar 4.22 Summary Hasil.....	78

DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	87
---	----



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA