

**KLASIFIKASI SENTIMEN MELALUI PERBANDINGAN SVM
DAN INDOBERT SERTA KATEGORISASI POLA KELUHAN
APLIKASI TRANSPORTASI PUBLIK**



TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat kelulusan Program Sarjana

RISKA ALIFIA PUTRI

15210004

**Program Studi Sains Data
Fakultas Teknologi Informasi
Universitas Nusa Mandiri
Jakarta
2026**

PERSEMBAHAN

Segala puji dan syukur saya panjatkan kepada Allah SWT atas segala rahmat, kesehatan, kekuatan, pertolongan, dan kemudahan yang diberikan hingga saya dapat menyelesaikan tugas akhir ini. Perjalanan ini tidak selalu mudah, tetapi setiap proses, doa, dan dukungan dari orang-orang terkasih telah mengajarkan saya untuk terus belajar dan bertumbuh. Dengan penuh rasa syukur, karya ini saya persembahkan kepada:

1. Ayah dan Mama yang senantiasa mencurahkan cinta, kasih sayang, doa yang tidak pernah putus, pengorbanan, kesabaran, serta segala dukungan yang telah diberikan hingga anak perempuannya dapat berada di titik ini. Terima kasih telah menjadi tempat pulang, tempat bercerita, dan alasan terbesar saya untuk terus berjuang. Semoga pencapaian kecil ini dapat menjadi salah satu bentuk kebahagiaan dan kebanggaan untuk Ayah dan Mama.
2. Adikku, atas perhatian, dukungan, dan semangat yang selalu diberikan. Semoga kita senantiasa saling mendukung dalam meraih impian masing-masing.
3. Para *support system*-ku, Ravi Romario Nugroho, Silvyana Puspa Anggraini, Riska Nurfajriah, Khaerunnisa Setya Putri, Natalia Susanthika, dan Ester Marysa, terima kasih telah hadir dalam setiap proses, menjadi tempat berbagi cerita, bertukar pikiran, saling menguatkan, dan menemani di masa-masa yang tidak mudah. Kehadiran kalian memberikan banyak warna selama perjalanan perkuliahan hingga penyusunan skripsi ini.
4. Teman-teman seperjuangan Program Studi Sains Data angkatan 2021 dan 2022, terima kasih atas kebersamaan, kerja sama, bantuan, pengalaman, dan kenangan selama menempuh pendidikan di Universitas Nusa Mandiri. Semoga langkah kita selanjutnya senantiasa dimudahkan dan diberikan kesuksesan.
5. Diriku sendiri, Riska Alifia Putri. *We did it!* Terima kasih telah bertahan, berusaha menyelesaikan apa yang telah dimulai, dan tidak menyerah ketika proses terasa berat. Skripsi ini bukan sekadar tentang gelar sarjana, melainkan bukti bahwa saya mampu melewati proses yang tidak selalu mudah. Teruslah percaya pada proses, belajar, bertumbuh, dan berani melangkah menuju impian berikutnya. Jangan pernah berhenti menjadi pribadi yang haus akan ilmu pengetahuan.

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini, saya:

Nama : Riska Alifia Putri
NIM : 15210004
Program Studi : Sains Data
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyatakan bahwa Tugas Akhir yang telah saya buat dengan judul: **“Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik”**, adalah asli (orisinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Tugas Akhir yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Universitas Nusa Mandiri** dicabut/dibatalkan.

Dibuat di : Jakarta
Pada tanggal: 30 Juli 2026
Yang Menyatakan,



Riska Alifia Putri

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Riska Alifia Putri
NIM : 15210004
Program Studi : Sains Data
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyetujui untuk memberikan ijin kepada pihak **Universitas Nusa Mandiri**, Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalti-Free Right*) atas karya ilmiah kami yang berjudul: **“Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik”**, beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak **Universitas Nusa Mandiri** berhak menyimpan, mengalih-media atau format-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak **Universitas Nusa Mandiri** segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 30 Juli 2026
Yang Menyatakan,



Riska Alifia Putri

PERSETUJUAN DAN PENGESAHAN TUGAS AKHIR

Tugas Akhir ini diajukan oleh:

Nama : Riska Alifia Putri

NIM 15210004

Program Studi : Sains Data

Fakultas : Teknologi Informasi

Jenjang : Strata Satu (S1)

Judul Tugas Akhir : **Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik**

Telah dipertahankan pada Periode I-2026 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Sains Data (S.Si.D) pada Program Sarjana Program Studi Sains Data Fakultas Teknologi Informasi di Universitas Nusa Mandiri.

Jakarta, 30 Juli 2026

PEMBIMBING TUGAS AKHIR

Dosen Pembimbing : **Ani Oktarini Sari, S.Kom, MMSI**



Ani Oktarini Sari
Ditandatangani Digital:
Ani Oktarini Sari
2026-08-20 16:47:15

DEWAN PENGUJI

Penguji I : **Tati Mardiana, M.Kom**



Tati Mardiana
Ditandatangani Digital:
Tati Mardiana
2026-08-20 09:30:29

Penguji II : **Susafa'ati, M.Kom**



Susafa'ati
Ditandatangani Digital:
Susafa'ati
2026-08-20 09:52:30

PEDOMAN PENGGUNAAN HAK CIPTA

Tugas Akhir yang berjudul “**Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik**” adalah hasil karya tulis asli **RISKA ALIFIA PUTRI** dan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku di lingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:

Nama : RISKA ALIFIA PUTRI
Alamat : Griya Cileungsi 3, Jawa Barat
No.Telp : 0895-4047-88488
E-mail : riskalifiaaa@gmail.com

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, karunia, dan pertolongan-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik”** sebagai salah satu syarat memperoleh gelar Sarjana pada Program Studi Sains Data, Universitas Nusa Mandiri.

Penyusunan Tugas Akhir ini tidak terlepas dari dukungan, bimbingan, dan doa dari berbagai pihak. Oleh karena itu, penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Rektor Universitas Nusa Mandiri.
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri.
3. Wakil Rektor II Bidang Non Akademik Universitas Nusa Mandiri.
4. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri.
5. Ibu Tati Mardiana, M.Kom, selaku Kepala Program Studi Sains Data, yang telah memberikan arahan, dukungan, serta masukan selama proses penyusunan Tugas Akhir ini.
6. Ibu Ani Oktarini Sari, S.Kom, MMSI., selaku Dosen Pembimbing, yang dengan penuh kesabaran telah memberikan bimbingan, arahan, dan masukan sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan baik.
7. Seluruh dosen dan staf Universitas Nusa Mandiri atas ilmu dan pengalaman, serta pelayanan yang telah diberikan selama masa perkuliahan.

Penulis menyadari penelitian ini masih memiliki keterbatasan sehingga kritik dan saran yang membangun sangat diharapkan. Penelitian ini diharapkan dapat menjadi referensi sekaligus memberikan gambaran penerapan klasifikasi sentimen dan kategorisasi pola keluhan pada aplikasi transportasi publik.

Jakarta, 30 Juli 2026

Penulis,



Riska Alifia Putri

ABSTRAK

Riska Alifia Putri (15210004), Klasifikasi Sentimen Melalui Perbandingan SVM dan IndoBERT serta Kategorisasi Pola Keluhan Aplikasi Transportasi Publik

Ulasan pengguna aplikasi TJ: Transjakarta di Google Play Store terus meningkat sehingga analisis sentimen secara manual menjadi tidak efisien untuk mengolah teks tidak terstruktur dalam jumlah besar. Selain itu, belum terdapat penelitian yang membandingkan performa *Support Vector Machine* (SVM) dan IndoBERT pada ulasan aplikasi tersebut serta mengkaji pola keluhan pengguna secara mendalam. Penelitian ini bertujuan membandingkan performa SVM dan IndoBERT dalam mengklasifikasikan sentimen ulasan serta mengidentifikasi tema keluhan yang dominan pada ulasan negatif. Penelitian menggunakan kerangka kerja *Cross Industry Standard Process for Data Mining* (CRISP-DM) terhadap 2.000 ulasan hasil *web scraping* yang dilabeli berdasarkan rating menjadi kelas positif dan negatif. SVM menggunakan representasi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan penanganan ketidakseimbangan kelas melalui *Synthetic Minority Over-sampling Technique* (SMOTE). Sementara itu, IndoBERT dibangun melalui *fine-tuning* model prelatih *indobert-base-pl* dengan pendekatan *class weighting*. Evaluasi kinerja menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* dengan pendekatan *Macro Average*. Identifikasi pola keluhan dilakukan melalui analisis frekuensi kata, *bigram*, dan *rule-based categorization*. Hasil penelitian menunjukkan bahwa IndoBERT memperoleh *Accuracy* 96,65%, *Recall Macro* 96,66%, dan *F1-Score Macro* 93,96%, lebih tinggi dibandingkan SVM yang masing-masing memperoleh 96,39%, 95,14%, dan 93,37%. Namun, *Precision Macro* IndoBERT sebesar 91,72% sedikit lebih rendah daripada SVM sebesar 91,80%. Tema *Bug & Update* Sistem menjadi keluhan paling dominan, yaitu 31,23% dari keseluruhan ulasan negatif. Temuan ini dapat menjadi masukan bagi pengembang untuk memprioritaskan perbaikan stabilitas aplikasi.

Kata Kunci: Klasifikasi Sentimen, *Support Vector Machine*, IndoBERT, TF-IDF, Pola Keluhan Pengguna

ABSTRACT

Riska Alifia Putri (15210004), *Sentiment Classification through SVM and IndoBERT Comparison, and Complaint Pattern Categorization in a Public Transportation Application*

User reviews of the TJ: Transjakarta application on the Google Play Store continue to increase, making manual sentiment analysis inefficient for processing large volumes of unstructured text. In addition, no studies have compared the performance of Support Vector Machine (SVM) and IndoBERT on reviews of this application while also examining user complaint patterns in depth. This study aims to compare the performance of SVM and IndoBERT in classifying review sentiment and to identify the dominant complaint themes in negative reviews. This study applies the Cross Industry Standard Process for Data Mining (CRISP-DM) framework to 2,000 reviews collected through web scraping and labeled as positive or negative based on their ratings. SVM uses Term Frequency-Inverse Document Frequency (TF-IDF) for feature representation, with class imbalance addressed using the Synthetic Minority Over-sampling Technique (SMOTE). Meanwhile, IndoBERT is developed by fine-tuning the pretrained indobert-base-p1 model using class weighting. Model performance is evaluated using Accuracy, Precision, Recall, and F1-Score with the Macro Average approach. Complaint patterns are identified through word frequency analysis, bigram analysis, and rule-based categorization. The results show that IndoBERT achieves an Accuracy of 96.65%, a Macro Recall of 96.66%, and a Macro F1-Score of 93.96%, outperforming SVM, which achieves 96.39%, 95.14%, and 93.37%, respectively. However, IndoBERT's Macro Precision of 91.72% is slightly lower than SVM's 91.80%. The Bug & System Update theme is the most dominant complaint, accounting for 31.23% of all negative reviews. These findings can serve as input for developers to prioritize improvements to application stability.

Keywords: *Sentiment Classification, Support Vector Machine, IndoBERT, TF-IDF, User Complaint Pattern*