

**IMPLEMENTASI DAN ANALISIS SENTIMEN PADA ULASAN
APLIKASI GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA RoBERTa**

SKRIPSI



**Oleh:
Brian Natanael Nainggolan
210210086**

**PROOGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2025**

**IMPLEMENTASI DAN ANALISIS SENTIMEN PADA ULASAN
APLIKASI GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA RoBERTa**

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar sarjana**



**Oleh:
Brian Natanael Nainggolan
210210086**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2025**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Brian Natanael Nainggolan

NPM : 210210086

Fakultas : Teknik dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa “**Skripsi**” yang saya buat dengan judul:

IMPLEMENTASI DAN ANALISIS SENTIMEN PADA ULASAN APLIKASI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA RoBERTa

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan penulis, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip di dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah skripsi ini digugurkan dan gelar akademik yang penulis peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapa pun.

Batam, 18 Juli 2025



Brian Natanael Nainggolan

210210086

**IMPLEMENTASI DAN ANALISIS SENTIMEN PADA ULASAN
APLIKASI GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA RoBERTa**

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar sarjana**

**Oleh
Brian Natanael Nainggolan
210210086**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, 18 Juli 2025


Koko Handoko, S.Kom., M.Kom.
Pembimbing

ABSTRAK

Pesatnya pertumbuhan jumlah ulasan pengguna aplikasi Instagram di *Google Play Store* menimbulkan tantangan dalam mengolah dan memahami sentimen secara cepat dan akurat. Proses pengolahan manual menjadi tidak efisien seiring bertambahnya volume data. Penelitian ini bertujuan untuk mengembangkan *dashboard* analisis sentimen berbasis model *Indonesian-RoBERTa* yang mampu secara otomatis mengumpulkan, menganalisis, dan menampilkan data sentimen pengguna. Proses dimulai dengan pengambilan data ulasan dari *Google Play Store* menggunakan *library google_play_scraper*, yang kemudian diproses dalam *dashboard* untuk dianalisis menggunakan model *Indonesian-RoBERTa* yang telah *fine-tuning* sebelumnya. Proses *fine-tuning* dilakukan terhadap model *W11wo/Indonesian-RoBERTa-base-sentiment-classifier* menggunakan dataset kumpulan *tweet* yang dikumpulkan dari platform twiter atau X pada masa PPKM dengan total data mencapai 23.645 *tweet* yang telah diberi label (positif, netral, negatif). Tahapan *preprocessing* meliputi pembersihan teks, tokenisasi, serta penerapan *class weighting* untuk mengatasi ketidakseimbangan kelas pada data latih. Evaluasi model dilakukan menggunakan metrik *precision*, *recall*, dan *F1-score*, serta divisualisasikan dalam bentuk *confusion matrix*. Hasil pengujian menunjukkan bahwa model mampu mengenali sentimen dengan performa yang baik pada kelas positif dan netral, meskipun performa pada kelas negatif masih perlu ditingkatkan. Sistem *dashboard* yang dikembangkan berhasil mengintegrasikan proses scraping, prediksi sentimen, dan visualisasi data, sehingga menjadi solusi efektif untuk memahami opini pengguna secara otomatis. Penelitian ini berkontribusi pada pengembangan teknologi NLP berbasis transformer dalam konteks bahasa Indonesia, sekaligus membuktikan potensi penerapannya dalam sistem pengambilan keputusan berbasis data.

Kata kunci: Analisis Sentimen, *Fine-tuning*, *Google Play Store*, *Indonesian-RoBERTa*, Instagram.

ABSTRACT

The rapid growth of User reviews on the Instagram application in the Google Play Store presents a significant challenge in processing and understanding sentiment quickly and accurately. Manual processing becomes inefficient as the volume of data increases. This study aims to develop a sentiment analysis dashboard based on the Indonesian-RoBERTa model, capable of automatically collecting, analyzing, and visualizing User sentiment data. The process begins by retrieving reviews from the Google Play Store using the google_play_scraper library, which are then processed within a dashboard and analyzed using the Indonesian-RoBERTa model that has been previously fine-tuned. The fine-tuning was performed on the w11wo/Indonesian-RoBERTa-base-sentiment-classifier model using a dataset consisting of 23,645 labeled tweets (positive, neutral, negative) collected from the Twitter or X platform during the PPKM period. The preprocessing steps included text cleaning, Tokenization, and the application of class weighting to address class imbalance in the training data. The model was evaluated using precision, recall, and F1-score metrics, along with a confusion matrix for performance visualization. The evaluation results show that the model performs well in recognizing positive and neutral sentiments, although the performance on the negative class still requires improvement. The developed dashboard successfully integrates scraping, sentiment prediction, and data visualization processes, offering an effective solution for automatically understanding User opinions. This research contributes to the advancement of transformer-based NLP technology in the Indonesian language context and demonstrates its potential for application in data-driven decision support systems.

Keywords: *Fine-tuning, Google Play Store, Indonesian-RoBERTa, Instagram Sentiment Analysis.*

KATA PENGANTAR

Puji dan syukur saya panjatkan ke hadapan Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam, Ibu Dr. Nur Elfi Husda, S.Kom., M.SI;
2. Bapak Welly Sugianto, S.T., M.M, selaku Dekan Fakultas Teknik Dan Komputer Universitas Putera Batam;
3. Bapak Andi Maslan, S.T., M.SI., Ph.D. selaku ketua Program Studi Teknik Informatika Universitas Putera Batam;
4. Pak Koko Handoko, S.Kom., M.Kom. selaku pembimbing skripsi saya yang telah membimbing dan memberikan arahan kepada penulis dalam proses pengerjaan skripsi ini;
5. Dosen dan Staff Universitas Putera Batam yang sudah membimbing dan memberikan ilmu kepada penulis selama penulis menempuh Pendidikan di Universitas Putera Batam;
6. Kepada Kedua orang tua penulis yang selalu memberi dukungan, semangat motivasi dan mendoakan penulis untuk menyelesaikan skripsi ini;
7. Saudara perempuan dan saudra pria penulis yang selalu mendoakan, meberikan semangat, dan dukungan penuh kepada penulis untuk menyelesaikan skripsi ini;

8. Teman teman yang ada di Braincore yang selalu mendukung saya untuk menyelesaikan skripsi ini;
9. Eric Julianto selaku Founder dari Braincore yang sudah membimbing saya dan memberikan ilmu kepada penulis selama menjadi bagian dari Braincore;
10. Semua pihak yang tidak bisa penulis sebutkan satu persatu yang telah banyak membantu dan memberikan semangat dalam proses pengerjaan skripsi ini;

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin

Batam, 18 Juli 2025



Brian Natanael Nainggolan

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS.....	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR RUMUS.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Identifikasi masalah	4
1.3. Batasan masalah.....	5
1.4. Rumusan Masalah.....	5
1.5. Tujuan Penelitian	6
1.6. Manfaat penelitian	6
1.6.1. Tujuan Teoritis:.....	6
1.6.2. Tujuan Praktis	7
BAB II TINJAUAN PUSTAKA.....	8
2.1. Teori Dasar.....	8
2.1.1. <i>Deep learning</i>	9
2.1.2. <i>Arsitektur dasar Deep learning</i>	10
2.1.3. <i>Training Pipeline dalam Deep Learning</i>	12
2.1.4. <i>Machine learning</i>	18
2.1.5. <i>Supervised learning</i>	20
2.1.6. <i>Unsupervised learning</i>	22
2.1.7. <i>Reinforcement learning</i>	23
2.1.8. <i>Natural Language Processing, Transformers, BERT, dan RoBERTa</i>	24
2.1.9. <i>Natural Language Processing (NLP)</i>	25
2.1.10. <i>Transformers</i>	29
2.1.11. <i>BERT (Bidirectional Encoder Representations from Transformers)</i>	32
2.1.12. <i>Indonesian-RoBERTa</i>	36
2.1.13. <i>Software pendukung</i>	39
2.1.14. <i>Visual Studio Code</i>	39
2.1.15. <i>Google Colab</i>	41
2.1.16. <i>Postman</i>	42
2.2. Penelitian terdahulu	44
2.3. Objek Penelitian.....	52
2.3.1. Instagram	53
2.3.2. <i>Google Play Store</i>	54
2.4. Kerangka pemikiran.....	54

BAB III METODOLOGI PENELITIAN.....	56
3.1. Desain Penelitian	56
3.2. Metode Pengumpulan data	59
3.2.1. Dataset <i>Finetuning</i>	59
3.2.2. Dataset model <i>Indonesian-RoBERTa</i>	65
3.3. Operasional variabel	66
3.4. Metode Perancangan.....	68
3.4.1. <i>Preprocessing</i> data.....	69
3.4.2. Perhitungan bobot kelas (<i>class weights</i>).....	70
3.4.3. Pembagian dataset	72
3.4.4. <i>Tokenization</i>	73
3.4.5. Proses <i>Training</i> (<i>Fine-tuning Process</i>).....	74
3.4.6. Evaluasi Model	76
3.4.7. Analisis Prediksi Model.....	80
3.4.8. <i>Confusion matrix</i>	83
3.4.9. Perancangan Website (UML).....	85
3.5. Desain Antarmuka	99
3.6. Algoritma.....	106
3.6.1. Proses <i>RoBERTa</i> dalam analisis sentimen	106
3.6.2. Implementasi model <i>Indonesian-RoBERTa</i>	113
3.7. Waktu dan tempat penelitian.....	116
3.7.1. Tempat penelitian.....	117
3.7.2. Waktu penelitian	117
BAB IV HASIL DAN PEMBAHASAN.....	119
4.1. Hasil Penelitian.....	119
4.1.1. <i>Preprocessing</i> data.....	119
4.1.2. Bobot Kelas	125
4.1.3. Hasil Pembagian Dataset	127
4.1.4. <i>Tokenization</i>	128
4.1.5. Hasil <i>Training</i>	130
4.1.6. Evaluasi Model	132
4.1.7. Hasil analisis prediksi model	138
4.1.8. <i>Confusion matrix</i>	140
4.1.9. Hasil Perancangan Web <i>Dashboard</i>	143
4.2. Pembahasan	149
4.2.1. Analisis Hasil.....	149
4.2.2. Hasil Pengujian.....	151
BAB V KESIMPULAN DAN SARAN	165
5.1. Kesimpulan	165
5.2. Saran	167
DAFTAR PUSTAKA.....	114
LAMPIRAN.....	118
Lampiran 1. Pendukung penelitian	118
Lampiran 2. Daftar riwayat hidup.....	137
Lampiran 3. Surat Keterangan Penelitian	138
Lampiran 4. LoA Jurnal	139

DAFTAR GAMBAR

Gambar 2.1 Struktur <i>Neural Network Deep learning</i>	11
Gambar 2.2 <i>Learning process neural network</i>	14
Gambar 2.3 Berbagai jenis teknik <i>machine learning</i>	20
Gambar 2.4 <i>Supervied Learning Workflow</i>	21
Gambar 2.5 Arsitektur Transformers	30
Gambar 2.6 Arsitektur BERT	33
Gambar 2.7 Kerangka Pemikiran	55
Gambar 3.1 Alur Penelitian.....	56
Gambar 3.2 Metode Perancangan.....	69
Gambar 3.3 <i>Confusion matrix</i>	84
Gambar 3.4 <i>Use Case Diagram</i>	86
Gambar 3.5 <i>Activity diagram Login</i>	90
Gambar 3.6 <i>Activity diagram register</i>	91
Gambar 3.7 <i>Activity diagram</i> klasifikasi teks <i>sentiment</i>	92
Gambar 3.8 <i>Activity diagram</i> hasil <i>sentiment</i> dan visualisasi data.....	93
Gambar 3.9 <i>Sequence Diagram Login dan Logout</i>	94
Gambar 3.10 <i>Sequence Diagram Register dan Login</i>	95
Gambar 3.11 <i>Sequence Diagram</i> klasifikasi teks sentimen.....	96
Gambar 3.12 <i>Sequence Diagram</i> Hasil Klasifikasi	97
Gambar 3.13 <i>Class Diagram</i> Penelitian	99
Gambar 3.14 Halaman <i>Register</i>	100
Gambar 3.15 Halaman <i>Login</i>	101
Gambar 3.16 Halaman <i>Home</i>	102
Gambar 3.17 Halaman Klasifikasi.....	103
Gambar 3.18 Ulasan yang berhasil diambil dari <i>Google Play Store</i>	104
Gambar 3.19 Hasil Sentimen dan visualisasi data.....	105
Gambar 3.20 Proses <i>RoBERTa</i> untuk klasifikasi teks	106
Gambar 4.2 Halaman <i>Login</i>	143
Gambar 4.3 Halaman <i>Register</i>	144
Gambar 4.4 Halaman <i>Home</i>	145
Gambar 4.5 Halaman Klasifikasi.....	146
Gambar 4.6 Hasil data ulasan yang berhasil di ambil	147
Gambar 4.7 Hasil Sentimen dan visualisasi data.....	148

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	45
Tabel 3.1 Contoh Data Mentah <i>Tweet</i> PPKM.....	60
Tabel 3.2 Informasi Model <i>Indonesian-RoBERTa</i>	65
Tabel 3.3 Tabel operasional <i>variable</i>	67
Tabel 3.4 Perhitungan Bobot Kelas	71
Tabel 3.5 Implementasi Model <i>Indonesian-RoBERTa</i>	113
Tabel 3.6 Tabel kegiatan selama penelitian	118
Tabel 4.1 Rincian awal dataset	119
Tabel 4.2 Distribusi akhir dataset setelah <i>preprocessing</i>	120
Tabel 4.3 Dataset yang sudah melewati tahap <i>processing</i> data	121
Tabel 4.4 Perhitungan Bobot Kelas	125
Tabel 4.5 interpretasi dari hasil perhitungan bobot kelas	126
Tabel 4.6 Pembagian Dataset ke dalam <i>Training, Validation, dan Test Set</i>	127
Tabel 4.7 Jumlah dataset persubset.....	128
Tabel 4.8 Contoh teks yang sudah di tokenisasi.....	129
Tabel 4.9 Hasil Pelatihan Model <i>per-epoch</i>	130
Tabel 4.10 Hasil perhitungan <i>precision, recall</i> dan <i>F1-score</i> untuk setiap kelas	133
Tabel 4.11 Hasil <i>Macro Averaging, Micro Averaging, dan Weighted Averaging</i>	136
Tabel 4.12 Ringkasan Hasil Prediksi	139
Tabel 4.13 Interpretasi Setiap Elemen <i>Confusion matrix</i>	140
Tabel 4.14 Ringkasan performa model untuk setiap kelas	141
Tabel 4.15 Hasil pengujian model <i>Indonesian-RoBERTa</i> pada <i>Dashboard Sentiment Analysis</i>	151

DAFTAR RUMUS

Rumus 3.1 <i>Class Weight</i>	71
Rumus 3.2 <i>Precision</i>	77
Rumus 3.3 <i>Recall</i>	77
Rumus 3.4 <i>F1-Score</i>	77
Rumus 3.5 <i>Precision Macro</i>	79
Rumus 3.6 <i>Recall Macro</i>	79
Rumus 3.7 <i>F1-Score Macro</i>	79
Rumus 3.8 <i>Softmax</i>	81