

**PERBANDINGAN PERFORMA TRANSFORMER-BASED
LANGUAGE MODEL UNTUK SENTIMENT ANALYSIS
TOPIK VIDEO GAME PADA TWITTER (X)**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

MOHAMMAD CALVIN WIRANATA

21.83.0667

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**PERBANDINGAN PERFORMA TRANSFORMER-BASED
LANGUAGE MODEL UNTUK SENTIMENT ANALYSIS
TOPIK VIDEO GAME PADA TWITTER (X)**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

MOHAMMAD CALVIN WIRANATA

21.83.0667

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

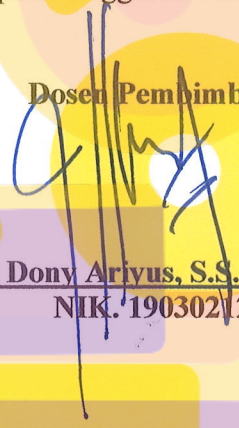
SKRIPSI

**PERBANDINGAN PERFORMA TRANSFORMER-BASED
LANGUAGE MODEL UNTUK SENTIMENT ANALYSIS
TOPIK VIDEO GAME PADA TWITTER (X)**

yang disusun dan diajukan oleh
Mohammad Calvin Wiranata
21.83.0667

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Juli 2025.

Dosen Pembimbing,


Dr. Dony Ariyus, S.S., M.Kom.
NIK. 190302128

HALAMAN PENGESAHAN

SKRIPSI

**PERBANDINGAN PERFORMA TRANSFORMER-BASED
LANGUAGE MODEL UNTUK SENTIMENT ANALYSIS
TOPIK VIDEO GAME PADA TWITTER (X)**

yang disusun dan diajukan oleh

Mohammad Calvin Wiranata

21.83.0667

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Juli 2025.

Susunan Dewan Penguji

Nama Penguji

Jeki Kuswanto, S.Kom., M.Kom.

NIK. 190302456

Wahid Miftahul Ashari, S.Kom., M.T.

NIK. 190302452

Dr. Dony Ariyus, S.S., M.Kom.

NIK. 190302128

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Mohammad Calvin Wiranata
NIM : 21.83.0667

Menyatakan bahwa Skripsi dengan judul berikut:

**PERBANDINGAN PERFORMA TRANSFORMER-BASED LANGUAGE
MODEL UNTUK SENTIMENT ANALYSIS TOPIK VIDEO GAME PADA
TWITTER (X)**

Dosen Pembimbing : Dr. Dony Ariyus, S.S., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 Juli 2025.

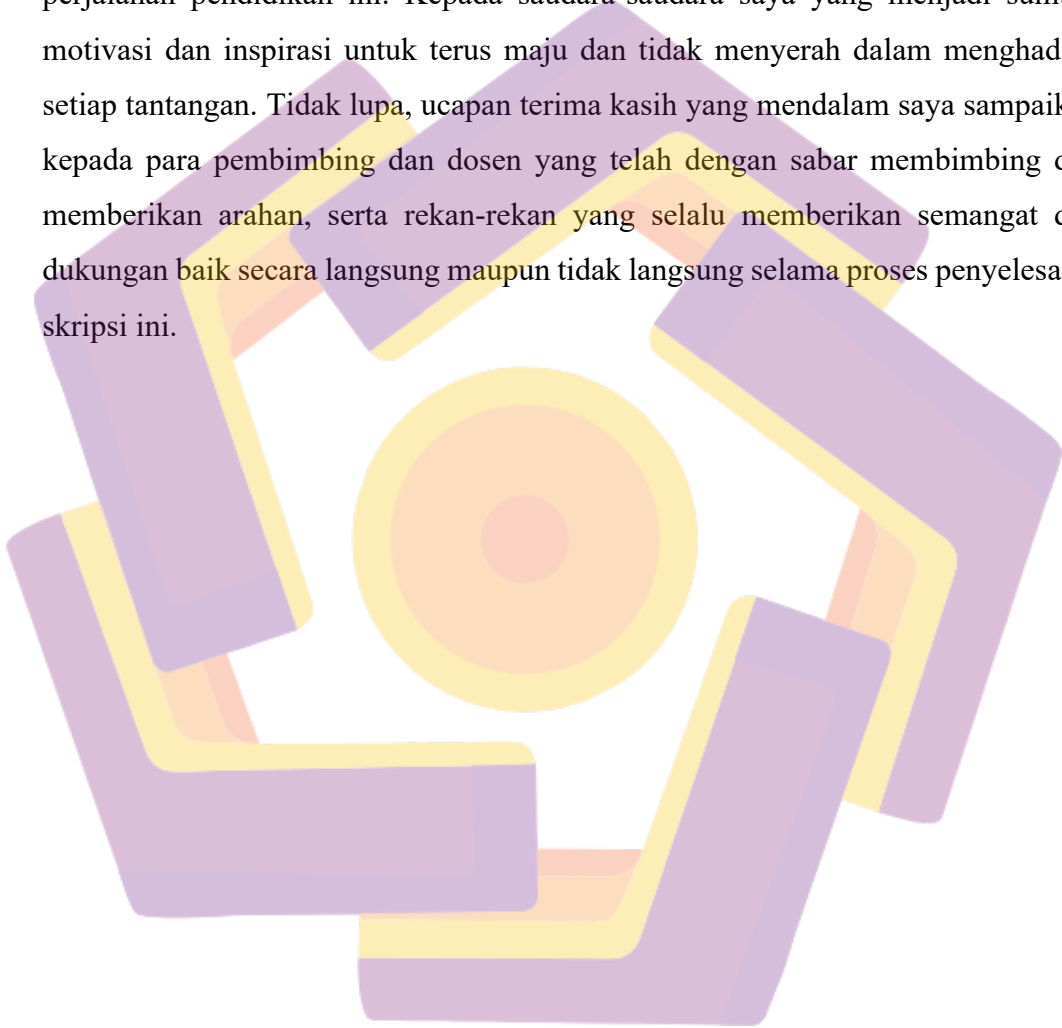
Yang Menyatakan,



Mohammad Calvin Wiranata

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Tuhan Yang Maha Esa, karya ini saya persembahkan kepada keluarga tercinta, khususnya orang tua saya yang selalu memberikan cinta, doa, dukungan, dan pengorbanan yang tak terhingga sepanjang perjalanan pendidikan ini. Kepada saudara-saudara saya yang menjadi sumber motivasi dan inspirasi untuk terus maju dan tidak menyerah dalam menghadapi setiap tantangan. Tidak lupa, ucapan terima kasih yang mendalam saya sampaikan kepada para pembimbing dan dosen yang telah dengan sabar membimbing dan memberikan arahan, serta rekan-rekan yang selalu memberikan semangat dan dukungan baik secara langsung maupun tidak langsung selama proses penyelesaian skripsi ini.



KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa, atas rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “**Perbandingan Performa Transformer-based Language Model Untuk Sentiment Analysis Topik Video Game pada Twitter (X)**” sebagai salah satu syarat untuk menyelesaikan program Sarjana (S1) Jurusan Teknik Komputer Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Dalam kesempatan ini, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah memberikan dukungan, bimbingan, dan bantuan selama proses penyusunan skripsi ini:

- 1 Dr. Dony Ariyus, S.S., M.Kom., selaku dosen pembimbing yang dengan sabar memberikan arahan, masukan, dan dukungan selama proses penyusunan skripsi ini.
- 2 Staff pengajar Fakultas Ilmu Komputer yang telah membagi ilmu yang sangat berharga selama penulis menempuh pendidikan di Universitas AMIKOM Yogyakarta.
- 3 Tim dosen penguji yang telah memberikan kritik dan saran yang membangun demi kesempurnaan karya ini.
- 4 Orang tua dan keluarga tercinta yang senantiasa memberikan doa, semangat, dan dukungan dalam langkah pendidikan penulis.
- 5 Rekan-rekan mahasiswa Teknik Komputer, yang selalu memberikan semangat dan bantuan selama proses penyusunan skripsi.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Oleh karena itu, penulis terbuka terhadap kritik dan saran yang membangun untuk perbaikan di masa depan. Semoga karya ini dapat memberikan manfaat bagi pembaca dan menjadi kontribusi positif dalam perkembangan ilmu pengetahuan.

Yogyakarta, 21 Juli 2025

Penulis

DAFTAR ISI

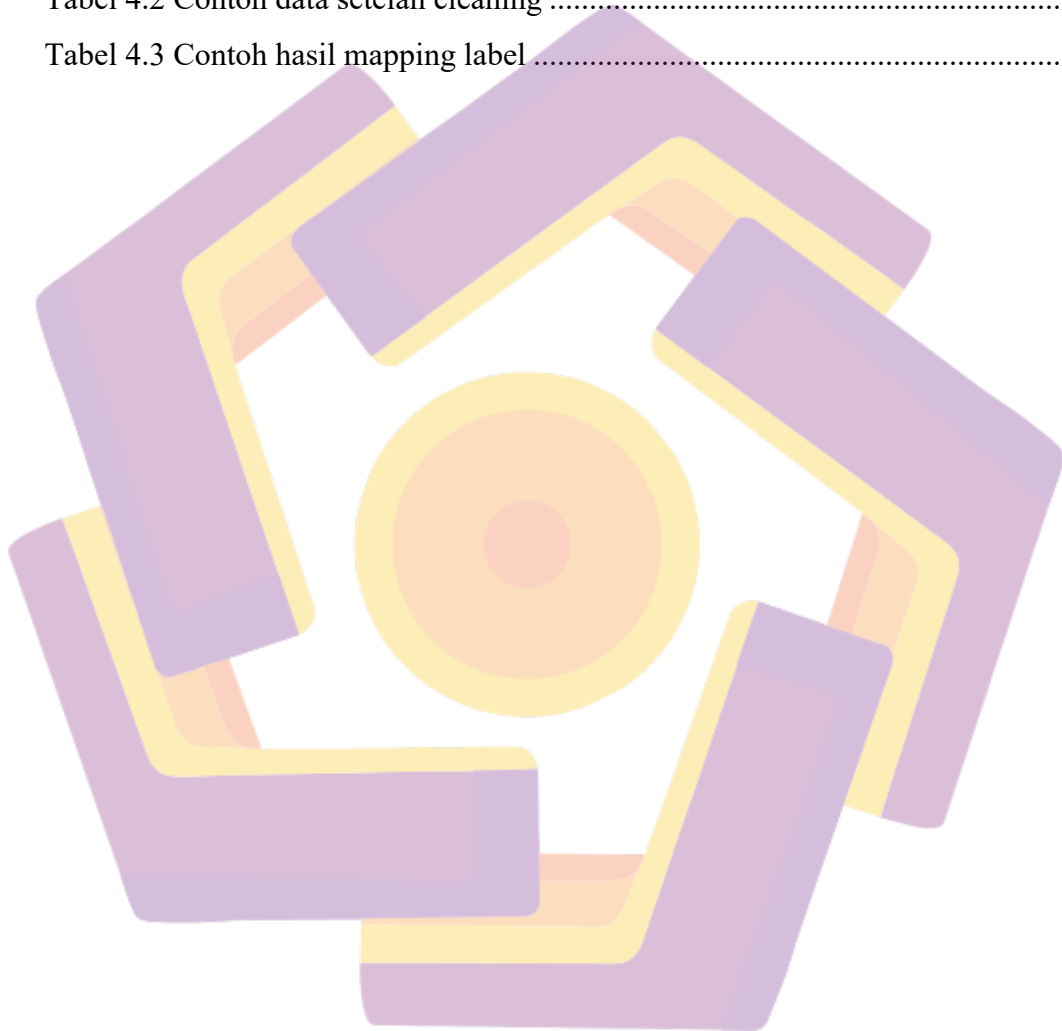
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xiii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH.....	xv
INTISARI	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5

2.2	Dasar Teori.....	17
2.2.1	Natural Language Processing (NLP)	17
2.2.2	Twitter (X)	17
2.2.3	Tweet Harvest	18
2.2.4	Embedding	18
2.2.5	HuggingFace	19
2.2.6	Transformer.....	20
2.2.7	VADER (<i>Valence Aware Dictionary and sEntiment Reasoner</i>) ...	22
2.2.8	BERT	24
2.2.9	DistilBERT	28
2.2.10	RoBERTa (<i>Robustly Optimized BERT</i>)	31
2.2.11	ALBERT (A Lite BERT).....	34
2.2.12	XLNet	35
2.2.13	Konfigurasi Pelatihan.....	37
2.2.14	Metrik Evaluasi.....	38
BAB III METODE PENELITIAN		42
3.1	Alur Penelitian	42
3.1.1	Pengumpulan Data	43
3.1.2	<i>Preprocessing Data</i>	44
3.1.3	Visualisasi Data	46
3.1.4	Persiapan Data	47
3.1.5	Training Data	50
3.1.6	Evaluasi Model	54
3.1.7	Penarikan Kesimpulan	56
3.2	Alat dan Bahan.....	57

BAB IV HASIL DAN PEMBAHASAN	59
4.1 Pengumpulan Data	59
4.1.1 <i>Data Scraping</i>	59
4.1.2 <i>Keyword Filtering</i>	60
4.1.3 Pelabelan Sentimen	62
4.2 <i>Preprocessing</i>	63
4.3 Visualisasi Data	65
4.4 Persiapan Data	66
4.4.1 Label Mapping	66
4.4.2 Splitting Data	68
4.4.3 Konversi Format Dataset	69
4.5 Training Model	69
4.5.1 Tokenisasi	70
4.5.2 Definisi Parameter Pelatihan	71
4.5.1 Training Model	72
4.6 Evaluasi Model	73
4.7 Perbandingan Hasil Evaluasi	77
BAB V PENUTUP	83
5.1 Kesimpulan	83
5.2 Saran	83
REFERENSI	85
LAMPIRAN	89

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 3.1 Tabel spesifikasi perangkat.....	58
Tabel 4.1 Contoh data sebelum cleaning	63
Tabel 4.2 Contoh data setelah cleaning	64
Tabel 4.3 Contoh hasil mapping label	67



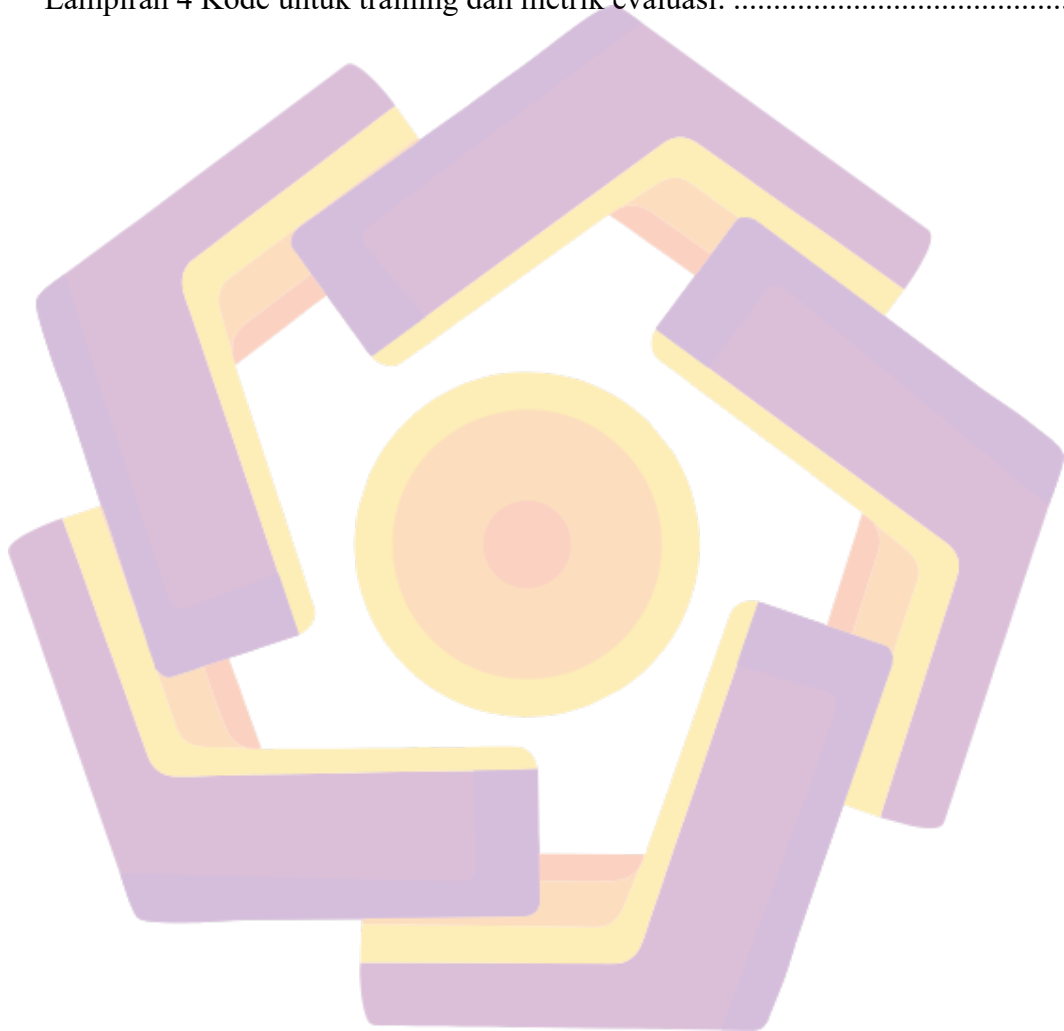
DAFTAR GAMBAR

Gambar 2.1 Visualisasi word embedding	18
Gambar 2.2 Alur kerja model transformer	20
Gambar 2.3 Diagram alur penilaian sentimen VADER.....	23
Gambar 2.4 Ilustrasi arsitektur transformer BERT	25
Gambar 2.5 Masked Language Model (MLM) BERT	26
Gambar 2.6 Next Sentence Prediction (NSP) BERT	27
Gambar 2.7 Model knowledge distillation teacher-student	29
Gambar 2.8 Diagram model arsitektur BERT base dan DistilBERT	31
Gambar 2.9 Arsitektur model RoBERTa	33
Gambar 2.10 Arsitektur model ALBERT	35
Gambar 2.11 Permutation language modelling	37
Gambar 3.1 Alur Penelitian	42
Gambar 3.2 Tampilan Tweet Harvest pada Google Colab	43
Gambar 3.3 Isi hasil scraping data menggunakan Tweet Harvest	43
Gambar 3.4 Laman dokumentasi nltk.sentiment.vader.....	44
Gambar 3.5 Contoh format tweet berpola sama	45
Gambar 3.6 Contoh visualisasi data berbentuk wordcloud.....	46
Gambar 3.7 Contoh visualisasi persentase klasifikasi sentimen.....	47
Gambar 3.8 Contoh mapping label sentimen.....	48
Gambar 3.9 Contoh perbedaan hasil tokenizer BERT dan RoBERTa	50
Gambar 3. 10 Contoh pemanggilan model	51
Gambar 3.11 Perbandingan performa optimizer untuk training neural network ...	52
Gambar 3.12 Konfigurasi TrainingArguments yang akan diterapkan pada semua model yang akan dicoba.....	53
Gambar 3.13 Contoh beberapa metrik sebagai perbandingan performa model.....	54
Gambar 3.14 Kode untuk menampilkan hasil evaluasi model.....	55
Gambar 4.1 Proses scraping tweets berdasarkan keyword dan jangka tanggal	59
Gambar 4.2 Pengumpulan hasil scraping menjadi satu file baru	60
Gambar 4.3 Kode untuk drop kolom yang tidak diperlukan	61

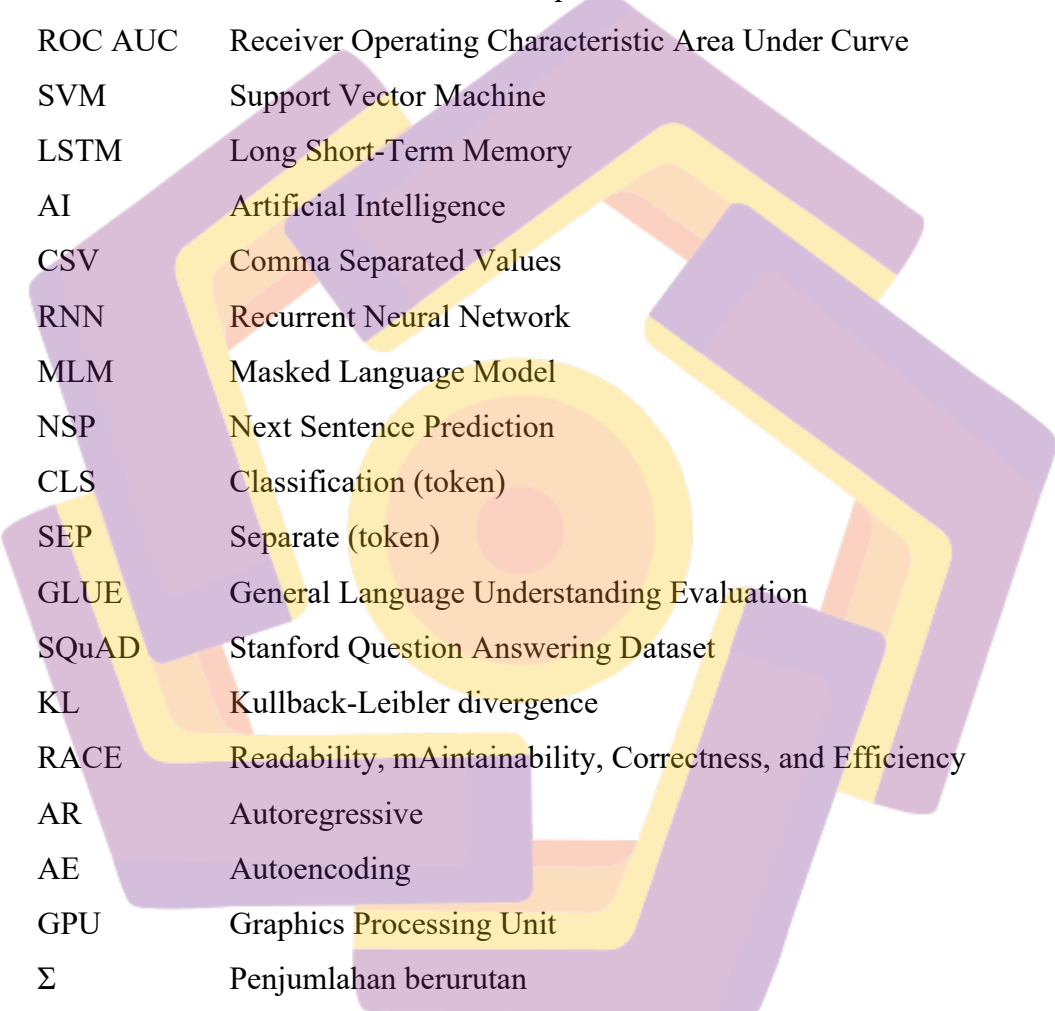
Gambar 4.4 Kode untuk melakukan proses keyword filtering kolom “full_text”	62
Gambar 4.5 Kode untuk proses labeling menggunakan VADER	63
Gambar 4.6 Hasil preprocessing data	65
Gambar 4.7 Visualisasi WordCloud	65
Gambar 4.8 Visualisasi persentase distribusi kelas sentimen	66
Gambar 4.9 Kode yang digunakan untuk mapping label	67
Gambar 4.10 Kode yang digunakan untuk splitting data	68
Gambar 4.11 Kode yang digunakan untuk splitting data	69
Gambar 4.12 Kode untuk mengubah model yang digunakan	70
Gambar 4.13 Kode yang digunakan untuk tokenisasi	70
Gambar 4.14 Contoh hasil tokenisasi (BERT)	71
Gambar 4.15 Argumen pelatihan yang digunakan dalam penelitian	72
Gambar 4.16 Kode yang digunakan untuk menetapkan fungsi metrik evaluasi	72
Gambar 4.17 Output metrik	72
Gambar 4.18 Fungsi untuk melakukan training dan mengukur lamanya proses	73
Gambar 4.19 Kode yang digunakan untuk menampilkan metrik evaluasi	74
Gambar 4.20 Output classification report (BERT)	74
Gambar 4.21 Output weighted & macro averages (BERT)	74
Gambar 4.22 Kode yang digunakan untuk menampilkan skor ROC AUC (BERT)	75
Gambar 4.23 Kode yang digunakan untuk memaparkan hasil dan plotting confusion matrix	75
Gambar 4.24 Kode yang digunakan untuk memaparkan hasil dan plotting confusion matrix	76
Gambar 4.25 Metrik evaluasi model-model yang digunakan	79
Gambar 4.26 Perbandingan kinerja model per-epoch	79
Gambar 4.27 Perbandingan evaluasi antar model	80
Gambar 4.28 Perbandingan confusion matrix	82

DAFTAR LAMPIRAN

Lampiran 1 Kode yang digunakan untuk menyatukan data:.....	89
Lampiran 2 Kode untuk preprocessing:	92
Lampiran 3 Kode untuk visualisasi data:	97
Lampiran 4 Kode untuk training dan metrik evaluasi:	99

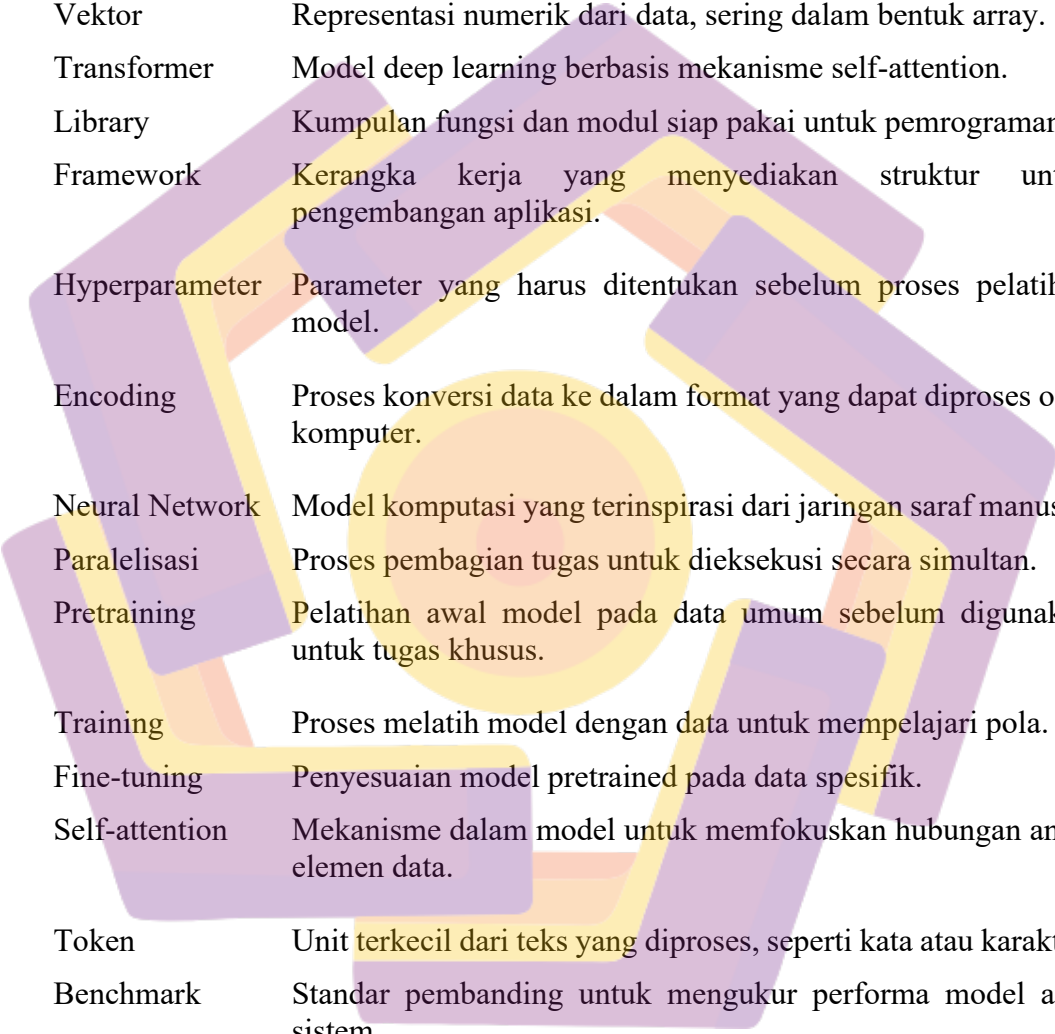


DAFTAR LAMBANG DAN SINGKATAN

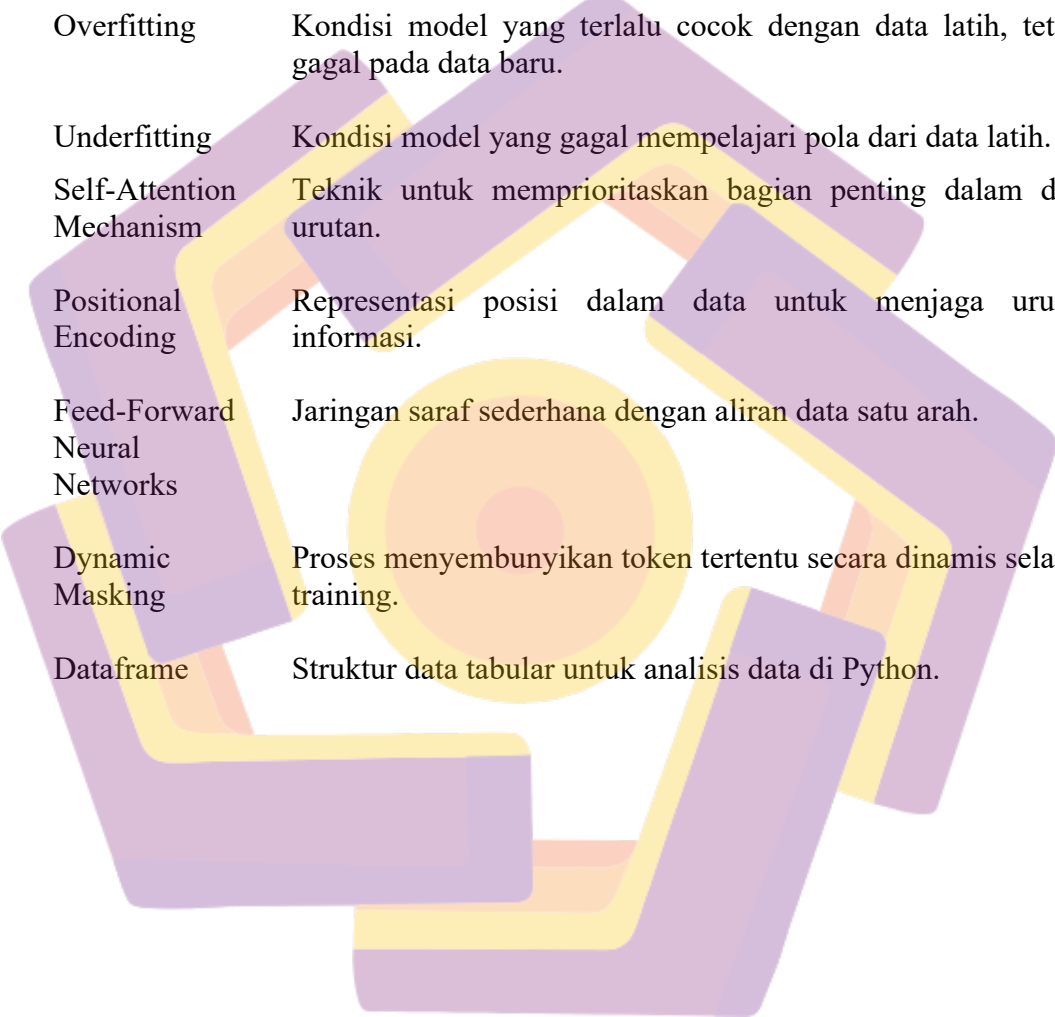


NLP	Natural Language Processing
VADER	Valence Aware Dictionary and sEntiment Reasoner
NLTK	Natural Language Toolkit
BERT	Bidirectional Encoder Representations from Transformers
ROC AUC	Receiver Operating Characteristic Area Under Curve
SVM	Support Vector Machine
LSTM	Long Short-Term Memory
AI	Artificial Intelligence
CSV	Comma Separated Values
RNN	Recurrent Neural Network
MLM	Masked Language Model
NSP	Next Sentence Prediction
CLS	Classification (token)
SEP	Separate (token)
GLUE	General Language Understanding Evaluation
SQuAD	Stanford Question Answering Dataset
KL	Kullback-Leibler divergence
RACE	Readability, mAintainability, Correctness, and Efficiency
AR	Autoregressive
AE	Autoencoding
GPU	Graphics Processing Unit
Σ	Penjumlahan berurutan
TP	True Positive
FP	False Positive
FN	False Negative
OvR	One-vs-Rest
OvO	One-vs-One
URL	Uniform Resource Locator
BBPE	Byte-Pair Encoding

DAFTAR ISTILAH



Tensor	Struktur data multidimensi untuk merepresentasikan data.
Scrape	Proses pengambilan data secara otomatis dari situs web.
Vektor	Representasi numerik dari data, sering dalam bentuk array.
Transformer	Model deep learning berbasis mekanisme self-attention.
Library	Kumpulan fungsi dan modul siap pakai untuk pemrograman.
Framework	Kerangka kerja yang menyediakan struktur untuk pengembangan aplikasi.
Hyperparameter	Parameter yang harus ditentukan sebelum proses pelatihan model.
Encoding	Proses konversi data ke dalam format yang dapat diproses oleh komputer.
Neural Network	Model komputasi yang terinspirasi dari jaringan saraf manusia.
Paralelisasi	Proses pembagian tugas untuk dieksekusi secara simultan.
Pretraining	Pelatihan awal model pada data umum sebelum digunakan untuk tugas khusus.
Training	Proses melatih model dengan data untuk mempelajari pola.
Fine-tuning	Penyesuaian model pretrained pada data spesifik.
Self-attention	Mekanisme dalam model untuk memfokuskan hubungan antar elemen data.
Token	Unit terkecil dari teks yang diproses, seperti kata atau karakter.
Benchmark	Standar pembandingan untuk mengukur performa model atau sistem.
Cross-entropy	Fungsi loss yang mengukur perbedaan antara prediksi dan label.
Arsitektur	Desain struktur model atau sistem.
Embedding	Representasi data dalam ruang vektor untuk mempermudah pemrosesan.
Permutasi	Penyusunan ulang elemen dalam berbagai urutan.



<i>Epoch</i>	Satu siklus penuh pelatihan model pada seluruh data.
Metrik	Ukuran yang digunakan untuk mengevaluasi performa model.
Matriks	Susunan data dalam bentuk dua dimensi atau lebih.
Preprocessing	Proses awal untuk mempersiapkan data sebelum digunakan.
Case folding	Konversi teks menjadi huruf kecil untuk konsistensi.
Pipeline	Alur kerja terintegrasi untuk mengotomatisasi proses.
Overfitting	Kondisi model yang terlalu cocok dengan data latih, tetapi gagal pada data baru.
Underfitting	Kondisi model yang gagal mempelajari pola dari data latih.
Self-Attention Mechanism	Teknik untuk memprioritaskan bagian penting dalam data urutan.
Positional Encoding	Representasi posisi dalam data untuk menjaga urutan informasi.
Feed-Forward Neural Networks	Jaringan saraf sederhana dengan aliran data satu arah.
Dynamic Masking	Proses menyembunyikan token tertentu secara dinamis selama training.
Dataframe	Struktur data tabular untuk analisis data di Python.

INTISARI

Dalam era digital, media sosial seperti Twitter sering digunakan untuk menyampaikan opini publik yang kaya akan informasi sentimen. Namun, analisis sentimen dari data Twitter menghadapi tantangan terkait volume data yang besar, keberagaman gaya bahasa, serta ketidakseimbangan kelas sentimen. Penelitian ini bertujuan untuk menguji dan membandingkan performa lima model berbasis deep learning, yaitu BERT, DistilBERT, RoBERTa, ALBERT, dan XLNet, dalam analisis sentimen data Twitter.

Metode yang digunakan melibatkan pengumpulan data Twitter melalui scraping menggunakan tool Tweet Harvest. Data kemudian melalui beberapa tahapan seperti pembersihan, labeling dengan VADER, serta splitting ke dalam format Hugging Face. Model dilatih menggunakan algoritma transformer dengan parameter yang seragam untuk memastikan validitas perbandingan. Evaluasi dilakukan berdasarkan metrik akurasi, precision, recall, F1-score, serta area under ROC curve (ROC AUC). Hasil evaluasi juga dianalisis melalui confusion matrix.

Hasil penelitian menunjukkan bahwa model ALBERT memiliki performa terbaik dengan akurasi sebesar 90.59% dan nilai ROC AUC lebih tinggi di antara model lainnya. Hasil ini menunjukkan bahwa model lightweight seperti ALBERT mampu mengungguli model lain meskipun menggunakan arsitektur yang lebih sederhana. Penelitian ini memberikan kontribusi dalam memilih model optimal untuk analisis sentimen Twitter, yang dapat dimanfaatkan oleh peneliti di bidang NLP, perusahaan analitik data, dan pengembang aplikasi. Penelitian lanjutan dapat mengeksplorasi performa model dengan data dari berbagai domain atau menambahkan teknik augmentasi data untuk mengatasi ketidakseimbangan kelas.

Kata kunci: Analisis Sentimen, Twitter, NLP, Deep Learning, ALBERT.

ABSTRACT

In the digital era, social media platforms like Twitter are widely used to express public opinions, rich in sentiment information. However, sentiment analysis of Twitter data faces challenges such as large data volumes, diverse language styles, and class imbalance. This study aims to evaluate and compare the performance of five deep learning-based models BERT, DistilBERT, RoBERTa, ALBERT, and XLNet in analyzing sentiment from Twitter data.

The methodology involved collecting Twitter data through scraping using the Tweet Harvest tool. The data underwent several preprocessing stages, including cleaning, sentiment labeling with VADER, and splitting into Hugging Face-compatible formats. The models were trained using transformer-based algorithms with standardized parameters to ensure valid comparisons. Evaluation metrics included accuracy, precision, recall, F1-score, and area under the ROC curve (ROC AUC). The results were further analyzed using confusion matrices.

The findings reveal that the ALBERT model achieved the best performance, with an accuracy of 90.59% and the highest ROC AUC among all models. This indicates that lightweight models like ALBERT can outperform others despite having simpler architectures. This study provides contributions to identifying optimal models for Twitter sentiment analysis, benefiting NLP researchers, data analytics companies, and application developers. Future research can explore model performance across different domains or incorporate data augmentation techniques to address class imbalances.

Keyword: Sentiment Analysis, Twitter, NLP, Deep Learning, ALBERT.