

**ANALISIS SENTIMEN PADA PLATFROM TWITTER
TERHADAP KETERLIBATAN INEOS SEBAGAI PEMILIK
BARU MANCHESTER UNITED MENGGUNAKAN MODEL
ROBERTA DAN VADER
SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
FARHAN FATHURROZIQ
21.61.0213

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**ANALISIS SENTIMEN PADA PLATFROM TWITTER
TERHADAP KETERLIBATAN INEOS SEBAGAI PEMILIK
BARU MANCHESTER UNITED MENGGUNAKAN MODEL
ROBERTA DAN VADER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

FARHAN FATHURROZIQ

21.61.0213

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PADA PLATFROM TWITTER TERHADAP
KETERLIBATAN INEOS SEBAGAI PEMILIK BARU MANCHESTER
UNITED MENGGUNAKAN MODEL ROBERTA DAN VADER**

yang disusun dan diajukan oleh

FARHAN FATHURROZIQ

21.61.0213

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 20 Agustus 2025

Dosen Pembimbing,



Hanafi, S.Kom, M.Eng, Ph.D
NIK. 190302024

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PADA PLATFROM TWITTER
TERHADAP KETERLIBATAN INEOS SEBAGAI PEMILIK
BARU MANCHESTER UNITED MENGGUNAKAN MODEL**

ROBERTADAN VADER

yang disusun dan diajukan oleh

FARHAN FATHURROZIQ

21.61.0213

Telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Ferian Fauzi Abdullah, S.Kom., M.Kom.
NIK. 190302276

Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 20 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Farhan Fathurroziq
NIM : 21.61.0213

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN PADA PLATFROM TWITTER TERHADAP KETERLIBATAN INEOS SEBAGAI PEMILIK BARU MANCHESTER UNITED MENGGUNAKAN MODEL ROBERTA DAN VADER

Dosen Pembimbing : Hanafi, S.Kom, M.Eng, Ph.D

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, 20 Agustus 2025

Yang Menyatakan,



Farhan Fathurroziq

HALAMAN PERSEMBAHAN

Dengan rasa syukur kepada Allah SWT, skripsi ini penulis persembahkan kepada:

- 1) Allah SWT – sebagai penolong, pemberi kekuatan, dan sumber harapan dalam seluruh proses kehidupan dan penelitian ini. Tanpa pertolongan dan kasih karunia-Nya, penulis tidak akan mampu menyelesaikan skripsi ini.
- 2) Orang tua dan keluarga tercinta – atas motivasi, cinta, doa yang tak pernah putus, serta dukungan finansial yang diberikan. Kehadiran mereka memberikan inspirasi untuk terus bertumbuh dan memegang teguh prinsip hidup “Kerja keras, kerja tuntas.”
- 3) Diri penulis sendiri – atas semangat dan keteguhan hati untuk menuntaskan kewajiban ini, meskipun menghadapi keraguan, lelah, dan tantangan. Penulis bangga atas pencapaian yang berhasil diraih hingga dapat menuntaskan skripsi ini.
- 4) Teman-teman kelas 21BCI, teman-teman di Pontianak, sepupu dan teman seperjuangan – terima kasih atas tawa, kebahagiaan, nasihat, motivasi, serta dukungan yang diberikan. Kehadiran kalian telah menemani dan membantu penulis dalam proses penyelesaian skripsi.
- 5) Dosen Pembimbing – atas bimbingan, arahan, dan kesabaran yang diberikan selama proses penelitian ini, sehingga penulis dapat menyelesaikan skripsi dengan baik.
- 6) Guru dan Dosen – mulai dari tingkat SD, SMP, SMA, hingga perguruan tinggi, atas ilmu, moral, dan adab yang diajarkan sehingga penulis dapat menempuh pendidikan hingga tahap ini.
- 7) Penulis buku dan pembuat konten bertema pengembangan diri – yang telah menciptakan karya-karya inspiratif dan memotivasi penulis untuk terus berproses menjadi pribadi yang lebih baik.
- 8) Bagi teman yang belum menyelesaikan studi kuliahnya. Skripsi ini saya persembahkan agar dapat memotivasi segera menyelesaikannya studinya.

Semua pihak di atas telah menjadi bagian penting dari proses yang Penulis lalui. Tanpa kalian, penelitian ini tidak akan mencapai titik akhir seperti saat ini.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan Skripsi dengan judul:

ANALISIS SENTIMEN PADA PLATFROM TWITTER TERHADAP KETERLIBATAN INEOS SEBAGAI PEMILIK BARU MANCHESTER UNITED MENGGUNAKAN MODEL ROBERTA DAN VADER

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Sarjana di Universitas Amikom Yogyakarta. Dengan selesainya skripsi ini, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Windha Mega Pradnya D, M.Kom. selaku Ketua Program Studi S1 Informatika.
4. Bapak Hanafi, S.Kom, M.Eng, Ph.D selaku dosen pembimbing yang telah memberikan bimbingan, arahan, serta semangat selama proses penelitian hingga penulisan skripsi ini.

Penulis juga menyampaikan terima kasih kepada seluruh pihak yang telah membantu secara langsung maupun tidak langsung dalam penyelesaian skripsi ini. Penulis menyadari bahwa penyusunan skripsi ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kesempurnaan skripsi ini. Penulis berharap karya ini dapat memberikan manfaat bagi semua pihak yang membaca dan membutuhkannya.

Yogyakarta, 22 September 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 LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.5.1 Manfaat Teoretis	6
1.5.2 Manfaat Praktis	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	9

2.1	Studi Literatur	9
2.2	Dasar Teori	15
2.3	Natural Language Processing (NLP)	15
2.3.1	Jenis-Jenis Natural Language Processing (NLP)	16
2.3.2	Faktor-Faktor Yang Mempengaruhi NLP	18
2.4	Analisis Sentimen	19
2.5	Lexicon-Based Approach	20
2.6	Machine Learning Approach	20
2.7	VADER & Lexicon-Based Approach Model	21
2.8	RoBERTa & Machine Learning Approach Model	22
2.8.1	Penghitungan Sentiment Score dengan Model RoBERTa	24
2.9	Logistic Regression sebagai Meta-Classfier pada Ensemble Stacking	26
BAB III METODE PENELITIAN		28
3.1	Objek Penelitian	28
3.2	Alur Penelitian	28
3.2.1	Pengumpulan Data	29
3.2.2	Preprocessing, Splitting and Balancing Data	30
3.2.3	Model Preparation	32
3.2.4	Fine-tuning RoBERTa model	33
3.2.5	Optimizing VADER threshold	34
3.2.6	Meta-classification model	35
3.2.7	Evaluate model	36
3.2.8	Analyze Manchester United fan sentiment	39
3.3	Alat dan Bahan	39
3.3.1	Alat	39

3.3.2	Bahan	40
BAB IV HASIL DAN PEMBAHASAN		42
4.1	Gambaran Umum	42
4.2	Pengumpulan data	42
4.2	Eksplorasi Data	44
4.2.1	Eksplorasi Data Pelatihan	44
4.2.2	Eksplorasi Data Tweet	48
4.3	Hasil Evaluasi Model RoBERTa	50
4.4	Hasil Evaluasi Model VADER	51
4.5	Hasil Evaluasi Model Meta-Classification	51
4.6	Perbandingan ketiga model	52
4.7	Analisis Sentimen Manchester United terhadap Keterlibatan INEOS	54
4.7.1	Analisis sentimen positif	55
4.7.2	Analisis sentimen netral	56
4.7.3	Analisis sentimen negatif	58
BAB V PENUTUP		60
5.1	Kesimpulan	60
5.2	Saran	60
REFERENSI		62

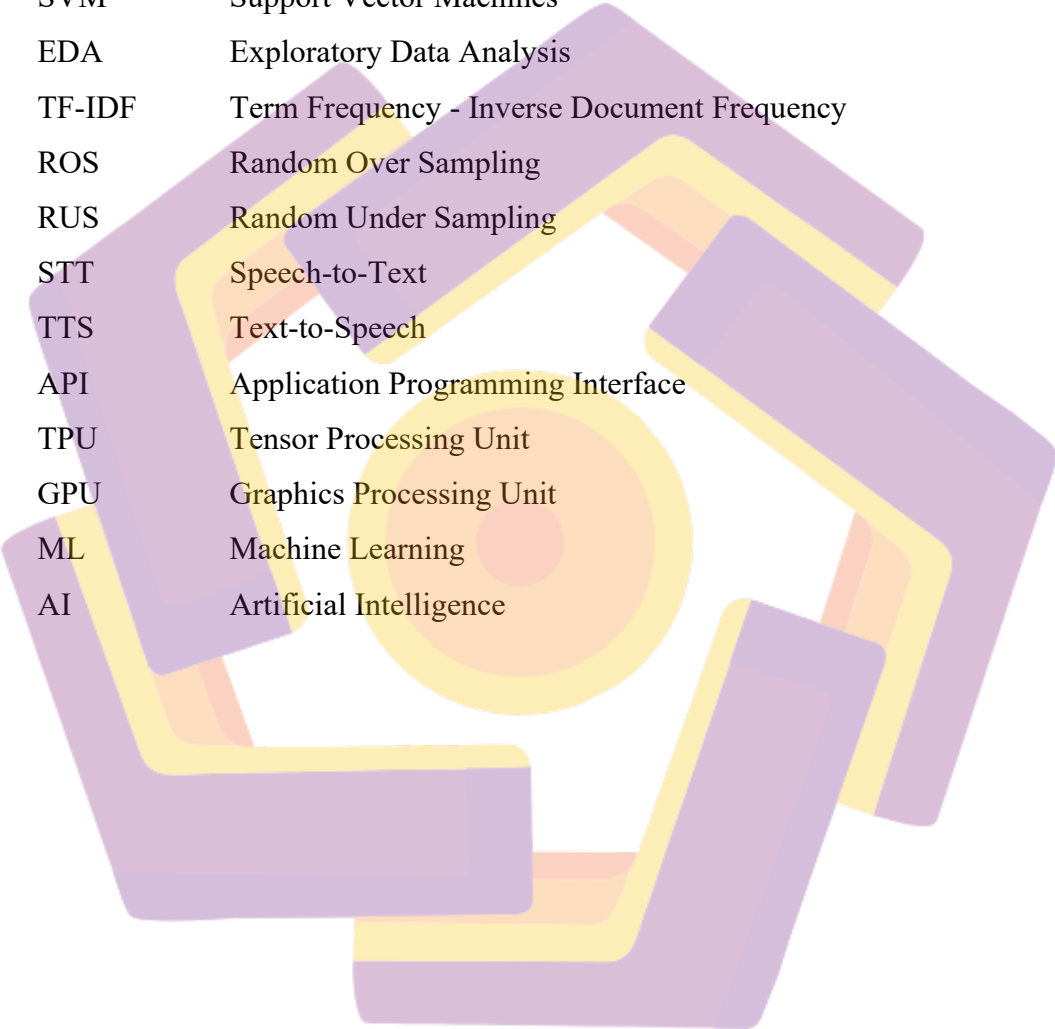
DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	13
Tabel 3. 1 Perbandingan data training awal dan setelah pembagian	31
Tabel 3. 2 Pembagian data balancing	32
Tabel 3. 3 Jumlah data train atau input meta-classfier model	36
Tabel 3. 4 Jumlah data uji meta-classifier model	36
Tabel 4. 1 Evaluasi fine-tuning RoBERTa pada 3 versi dataset	50
Tabel 4. 2 Perbandingan threshold default dan dioptimalkan VADER	51
Tabel 4. 3 Hasil evaluasi meta-classifier model	51
Tabel 4. 4 Metrik ketiga model	52
Tabel 4. 5 Perbandingan prediksi model terhadap sentimen Manchester United	54

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur transformers	23
Gambar 3. 1 Alur Penelitian	29
Gambar 3. 2 Potongan kode optimasi threshold VADER	35
Gambar 4. 1 Hasil Jumlah data training yang didapatkan	44
Gambar 4. 2 Distribusi Label Sentimen Data Pelatihan	45
Gambar 4. 3 Distribusi panjang tweet per kelas sentimen	46
Gambar 4. 4 WordCloud data pelatihan	47
Gambar 4. 5 Persebaran distribusi data tweet	49
Gambar 4. 6 WordCloud data tweet	50
Gambar 4. 7 Confusion Matrix ketiga model	53
Gambar 4. 8 Visualisasi perbandingan prediksi model terhadap sentimen Manchester United	55
Gambar 4. 9 Visualisasi Wordcloud Positive sentimen data tweet	56
Gambar 4. 10 Visualisasi Wordcloud Neutral sentimen data tweet	57
Gambar 4. 11 Visualisasi Wordcloud Negative sentimen data tweet	58

DAFTAR LAMBANG DAN SINGKATAN

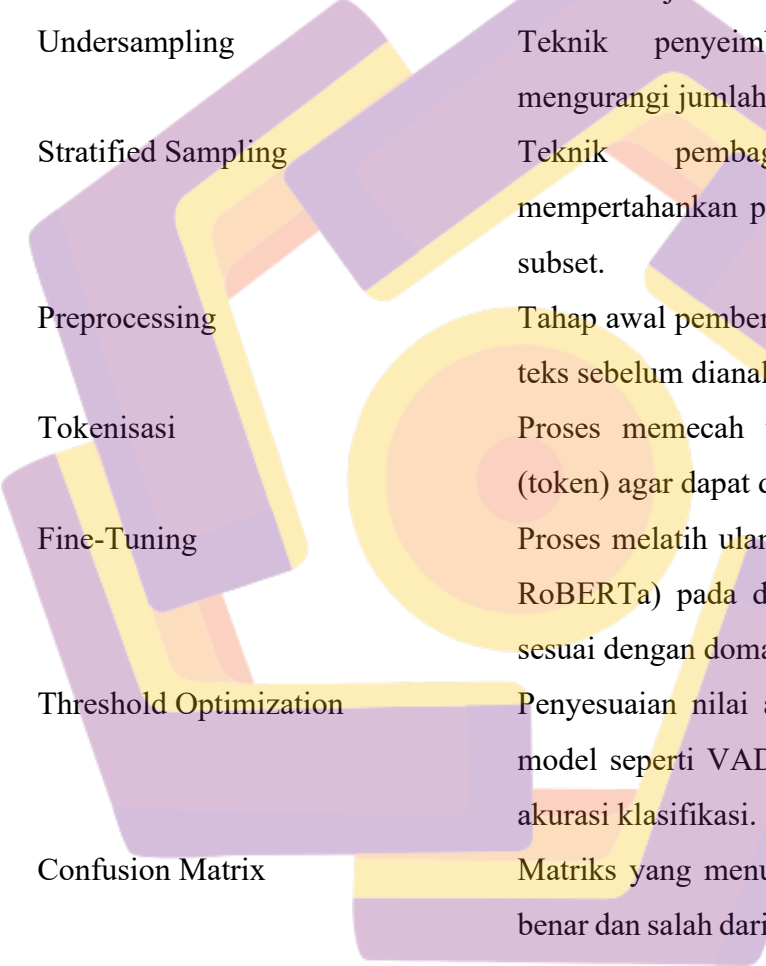


NLP	Natural Language Processing
RoBERTa	Robustly Optimized BERT Pretraining Approach
VADER	Valence Aware Dictionary and Sentiment Reasoner
SVM	Support Vector Machines
EDA	Exploratory Data Analysis
TF-IDF	Term Frequency - Inverse Document Frequency
ROS	Random Over Sampling
RUS	Random Under Sampling
STT	Speech-to-Text
TTS	Text-to-Speech
API	Application Programming Interface
TPU	Tensor Processing Unit
GPU	Graphics Processing Unit
ML	Machine Learning
AI	Artificial Intelligence

DAFTAR ISTILAH



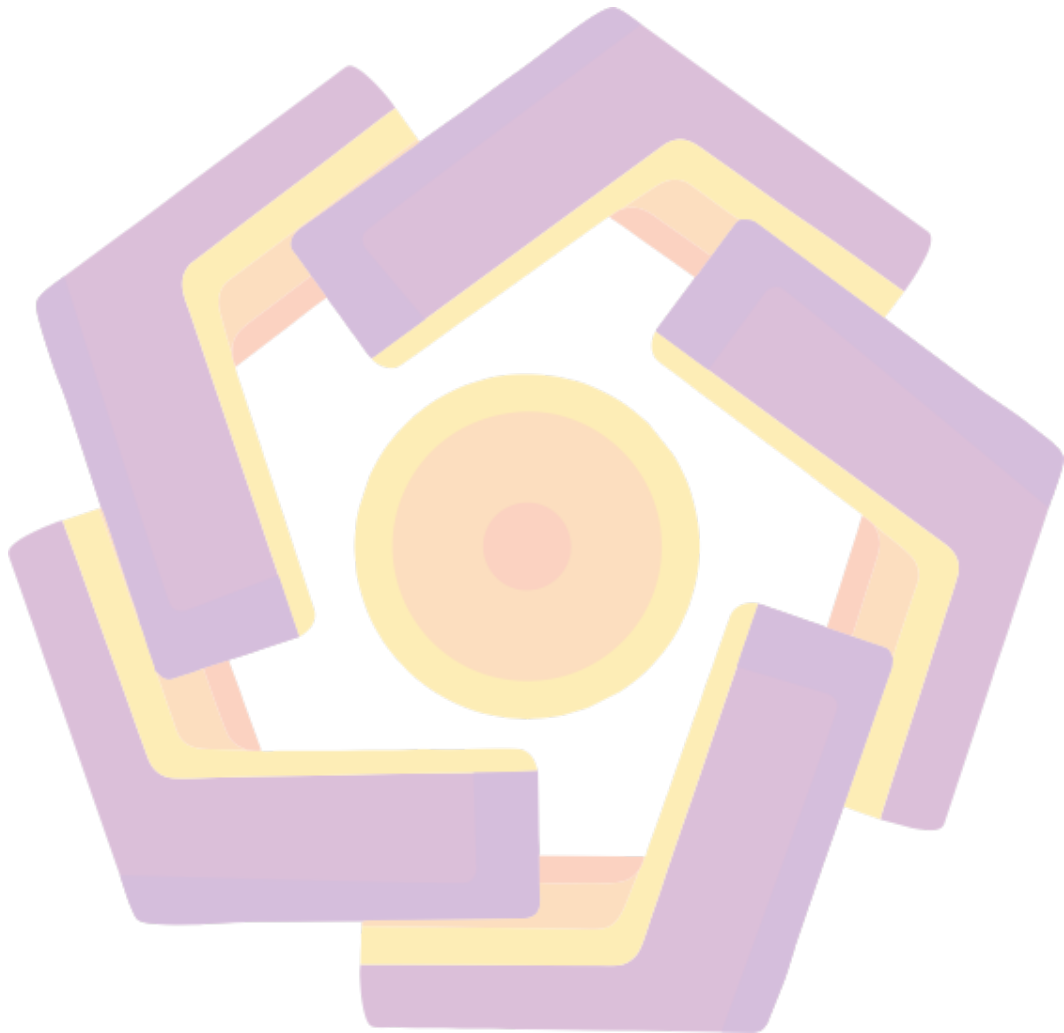
Natural Language Processing	Bidang kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia.
Analisis Sentimen	Proses mengklasifikasikan teks menjadi kategori opini (positif, negatif, netral).
RoBERTa	Model NLP berbasis transformer yang dioptimalkan dari BERT untuk pemahaman konteks bahasa yang lebih baik.
VADER	Model berbasis leksikon yang dirancang untuk analisis sentimen pada teks pendek seperti tweet.
Lexicon-Based Approach	Pendekatan analisis sentimen menggunakan kamus kata yang memiliki nilai polaritas.
Machine Learning Approach	Pendekatan analisis sentimen menggunakan algoritma pembelajaran mesin untuk mengenali pola dari data latih.
Meta-Classifer	Model yang menggabungkan hasil prediksi dari beberapa model (base models) untuk menghasilkan keputusan akhir yang lebih baik.
Logistic Regression	Algoritma klasifikasi yang digunakan sebagai meta-classifier untuk ensemble stacking dalam penelitian ini.
Precision	Rasio prediksi positif yang benar dibandingkan dengan seluruh prediksi positif.
Recall	Rasio prediksi positif yang benar dibandingkan dengan seluruh data positif sebenarnya.



F1-Score	Rata-rata harmonis antara precision dan recall, digunakan untuk mengukur keseimbangan keduanya.
Accuracy	Proporsi prediksi yang benar dibandingkan seluruh data.
Oversampling	Teknik penyeimbangan data dengan menambah jumlah sampel kelas minoritas.
Undersampling	Teknik penyeimbangan data dengan mengurangi jumlah sampel kelas mayoritas.
Stratified Sampling	Teknik pembagian data dengan mempertahankan proporsi kelas pada setiap subset.
Preprocessing	Tahap awal pembersihan dan penyiapan data teks sebelum dianalisis atau diproses model.
Tokenisasi	Proses memecah teks menjadi unit kecil (token) agar dapat diproses oleh model.
Fine-Tuning	Proses melatih ulang model pralatih (seperti RoBERTa) pada dataset khusus agar lebih sesuai dengan domain data.
Threshold Optimization	Penyesuaian nilai ambang (threshold) pada model seperti VADER untuk meningkatkan akurasi klasifikasi.
Confusion Matrix	Matriks yang menunjukkan jumlah prediksi benar dan salah dari model untuk setiap kelas.
WordCloud	Representasi visual kata yang sering muncul dalam dataset, dengan ukuran kata sesuai frekuensinya.
Tweet-Harvest	Alat yang digunakan untuk melakukan crawling data tweet dari Twitter.

Open-Access

Sistem publikasi yang memungkinkan artikel/jurnal dapat diakses secara gratis oleh siapa pun tanpa hambatan berbayar.



INTISARI

Keterlibatan INEOS dalam kepemilikan Manchester United menjadi topik yang banyak diperbincangkan oleh publik, khususnya di media sosial seperti Twitter. Dinamika opini publik yang muncul menimbulkan kebutuhan untuk menganalisis sentimen penggemar terhadap perubahan tersebut. Permasalahan yang diangkat dalam penelitian ini adalah bagaimana sentimen terhadap keterlibatan INEOS sebagai pemilik baru Manchester United berdasarkan data yang diungkap melalui platform Twitter, serta bagaimana model-model Natural Language Processing (NLP) dapat digunakan untuk mengevaluasi sentimen tersebut secara akurat.

Penelitian ini menggunakan pendekatan NLP dengan dua model utama. RoBERTa, sebagai model berbasis pembelajaran mesin, dan VADER, sebagai model berbasis leksikon. Data diambil melalui proses crawling menggunakan *Tweet-Harvest* dan dilakukan preprocessing, splitting, serta balancing. RoBERTa dilakukan fine-tuning terhadap data latih berlabel, sedangkan VADER dioptimalkan melalui penyesuaian threshold. Hasil prediksi keduanya kemudian dikombinasikan ke dalam model meta-classifier berbasis Logistic Regression untuk meningkatkan akurasi klasifikasi.

Hasil evaluasi menunjukkan bahwa model gabungan RoBERTa dan VADER memiliki performa terbaik dengan akurasi 71% dan F1-score 71%, mengungguli performa masing-masing model secara terpisah. Model gabungan juga menunjukkan kemampuan yang lebih seimbang dalam mengklasifikasikan sentimen positif, netral, dan negatif. Penelitian ini berkontribusi dalam pengembangan metode hybrid untuk analisis sentimen berbasis media sosial, dan dapat dimanfaatkan oleh praktisi komunikasi, analis media, maupun stakeholder klub untuk memahami persepsi publik secara lebih mendalam. Rekomendasi selanjutnya adalah penggunaan dataset multimodal dan eksplorasi teknik ensemble lanjutan.

Kata kunci: Analisis Sentimen, RoBERTa, VADER, Meta-classifier, Logistic Regression.

ABSTRACT

INEOS' involvement in the ownership of Manchester United has become a hot topic of discussion among the public, especially on social media platforms such as Twitter. The dynamics of public opinion that have emerged necessitate an analysis of fan sentiment toward this change. The issue addressed in this study is how sentiment toward INEOS involvement as the new owner of Manchester United is reflected in data revealed through the Twitter platform, as well as how Natural Language Processing (NLP) models can be used to accurately evaluate such sentiment.

This study employs an NLP approach with two main models: RoBERTa, a machine learning-based model, and VADER, a lexicon-based model. Data was collected through a crawling process using Tweet-Harvest and underwent preprocessing, splitting, and balancing. RoBERTa was fine-tuned on labeled training data, while VADER was optimized through threshold adjustment. The prediction results of both models were then combined into a Logistic Regression-based meta-classifier model to improve classification accuracy.

The evaluation results showed that the combined RoBERTa and VADER model had the best performance with an accuracy of 71% and an F1-score of 71%, outperforming the performance of each model separately. The combined model also demonstrated more balanced capabilities in classifying positive, neutral, and negative sentiments. This research contributes to the development of hybrid methods for social media-based sentiment analysis and can be utilized by communication practitioners, media analysts, and club stakeholders to gain a deeper understanding of public perception. Further recommendations include the use of multimodal datasets and the exploration of advanced ensemble techniques.

Keyword: Sentiment analysis, RoBERTa, VADER, Meta-classifier, Logistic Regression.