Sentiment Analysis of the Indonesian Megathrust Earthquake and Tsunami Issue Using BERT and Roberta Methods

Hendra Rahman [1)*], Taswanda Taryo [2)], Sudarno Wiharjo [3)]

[1)2)3)]Teknik Informatika S-2, Program Pascasarjana, Universitas Pamulang
[*)]**Correspondence author**: hendra.rahman@bmkg.go.id, Tangerang Selatan, Indonesia
DOI: https://doi.org/10.37012/jtik.v12i1.3563

Abstract

The megathrust earthquake in Indonesia is a major potential natural disaster capable of triggering high-magnitude earthquakes and tsunamis, thereby influencing public perception. This study aims to analyze public opinion and identify the main topics related to the megathrust earthquake issue using Bidirectional Encoder Representations from Transformers (BERT) and Robustly Optimized BERT Pretraining Approach (RoBERTa) models. The dataset consists of 16,592 comments collected from the X social media platform during the period 2012–2025, which were classified into three sentiment categories, positive, negative, and neutral. The research methodology included exploratory data analysis, text preprocessing, model training, and evaluation using four experimental scenarios. The results indicate that the best performance was achieved using an 80:10:10 train–validation test split with ten training epochs. The BERT model outperformed RoBERTa, achieving an accuracy of 92,4350%, precision of 92,4291%, recall of 92,4350%, and F1-score of 92,4292%. These findings demonstrate that BERT is more effective in capturing the linguistic context of the Indonesian language. Furthermore, this study contributes to the advancement of artificial intelligence-based sentiment analysis for monitoring public opinion on disaster-related issues and provides a valuable foundation for developing more effective risk communication strategies, disaster mitigation education, and evidence-based policymaking that is more responsive to public perception.

Keywords: *Sentiment Analysis, Megathrust, BERT, RoBERTa, Indonesia Earthquake*

Abstrak

Gempa megathrust di Indonesia merupakan potensi bencana besar yang dapat memicu gempa bumi berkekuatan tinggi dan tsunami sehingga memengaruhi persepsi masyarakat. Penelitian ini bertujuan menganalisis opini publik serta mengidentifikasi topik utama terkait isu gempa megathrust menggunakan model BERT dan RoBERTa. Dataset terdiri atas 16.592 komentar dari media sosial X periode 2012–2025 yang diklasifikasikan ke dalam sentimen positif, negatif, dan netral. Tahapan penelitian meliputi eksplorasi data, prapemrosesan teks, pelatihan model, dan evaluasi dengan empat skenario. Hasil menunjukkan bahwa konfigurasi terbaik diperoleh pada pembagian data 80:10:10 dengan epoch 10. Model BERT memberikan kinerja lebih baik dibandingkan RoBERTa dengan akurasi 92,4350%, presisi 92,4291%, recall 92,4350%, dan F1-score 92,4292%. Temuan ini menunjukkan bahwa BERT lebih efektif dalam memahami konteks linguistik bahasa Indonesia. Hasil penelitian ini memberikan kontribusi terhadap pengembangan analisis sentimen berbasis kecerdasan buatan untuk pemantauan opini publik terkait kebencanaan serta dapat dimanfaatkan sebagai dasar penyusunan strategi komunikasi risiko, edukasi mitigasi bencana, dan pengambilan kebijakan yang lebih responsif terhadap persepsi masyarakat.

Kata Kunci: Analisis Sentimen, Megathrust, BERT, RoBERTa, Gempa Indonesia

PENDAHULUAN

Indonesia adalah negara yang terdiri dari banyak pulau dan terletak di garis khatulistiwa, di antara dua benua serta dua lautan. Keadaan ini membuat Indonesia memiliki ciri-ciri meteorologi dan klimatologi yang sangat rumit, termasuk kemungkinan muncul-nya cuaca serta iklim yang ekstrim. Dari sudut pandang geologi, Indonesia ter-letak di pertemuan empat lempeng tektonik utama di dunia, yaitu Lempeng Indo-Australia, Lempeng Eurasia, Lempeng Pasifik, dan Lempeng Philippine.(Hutchings & Mooney, 2021).

Posisi geografis negara ini menjadikan Indonesia terletak di ring of fire dan salah satu area dengan frekuensi gempa yang paling sering terjadi di seluruh dunia. Megathrust adalah zona di mana lempeng tektonik bertemu, dan ini memiliki kemampuan untuk memicu gempa yang sangat besar serta berpotensi mendatangkan tsunami. Indonesia memiliki sejumlah zona megathrust aktif yang tersebar di Sumatra, Jawa, Bali, Nusa Tenggara, Sulawesi, dan Maluku, sehingga menjadikannya sebagai negara dengan resiko gempa dan tsunami yang cukup tinggi. Ancaman dari megathrust menjadi perhatian yang serius karena dapat mengakibatkan kerusakan pada infrastruktur, menimbulkan korban jiwa, dan memberikan dampak sosial serta ekonomi yang signifikan. (Supendi et al., 2023).

Perkembangan teknologi komunikasi dan jejaring sosial membuat masyarakat lebih giat dalam menyampaikan pendapat, tanggapan, serta pandangan mengenai masalah bencana melalui saluran digital, terutama di platform media sosial X (sebelumnya *Twitter*)(Naufal et al., 2023)(Contreras et al., 2022). Informasi terkait gempa, tsunami, dan megathrust sering kali menjadi bahan diskusi hangat di kalangan masyarakat, khususnya saat peringatan awal dikeluarkan, isu menjadi viral, atau saat terjadi gempa besar. Jejaring sosial berfungsi sebagai sumber informasi(Seneviratne et al., 2024) sekaligus tempat untuk menyebarkan pandangan masyarakat dengan sangat cepat dan luas(Naufal et al., 2023)(Fakhruddin et al., 2024).

Berdasarkan pengumpulan informasi melalui metode web *scraping* di media sosial X antara tahun 2012 sampai Oktober 2025, ditemukan sebanyak 16.592 data komentar yang berkaitan dengan topik gempa bumi dan tsunami megathrust di Indonesia. Hasil tersebut mengindikasikan bahwa kepedulian masyarakat mengenai isu megathrust mengalami variasi setiap tahunnya. Pada tahun 2024, terjadi lonjakan jumlah diskusi yang sangat besar hingga

mencapai 5.352 komentar, yang mencerminkan tingginya minat publik terhadap topik ini. Sementara itu, pada tahun-tahun sebelumnya, jumlah komentar cenderung lebih rendah dan tidak konsisten(Tan & Lee, 2023).

Banyaknya informasi pendapat masyarakat yang ada di platform sosial media dapat digunakan untuk memahami perasaan umum mengenai isu megathrust, baik itu berupa perasaan positif, negatif, atau netral. Namun, data dari media sosial memiliki sifat yang tidak teratur, mengandung akronim, penggunaan bahasa yang tidak formal, simbol, dan gangguan yang memerlukan teknik analisis yang dapat mengerti konteks bahasa dengan lebih baik(Firdaus et al., 2026).

Salah satu teknik yang sering diterapkan dalam analisis sentimen adalah pendekatan *Natural Language Processing* (NLP). Dalam studi ini, digunakan model yang berlandaskan transformer, yaitu BERT (*Bidirectional Encoder Representations from Transformers*) dan RoBERTa (*Robustly Optimized BERT Pre-training Approach*). Kedua metode ini dipilih karena mereka mampu menangkap konteks kalimat dengan mendalam dan telah terbukti menghasilkan kinerja yang baik dalam berbagai tugas pengklasifikasian teks, termasuk analisis sentimen(Srivastava & Soni, 2022; Tan et al., 2023).

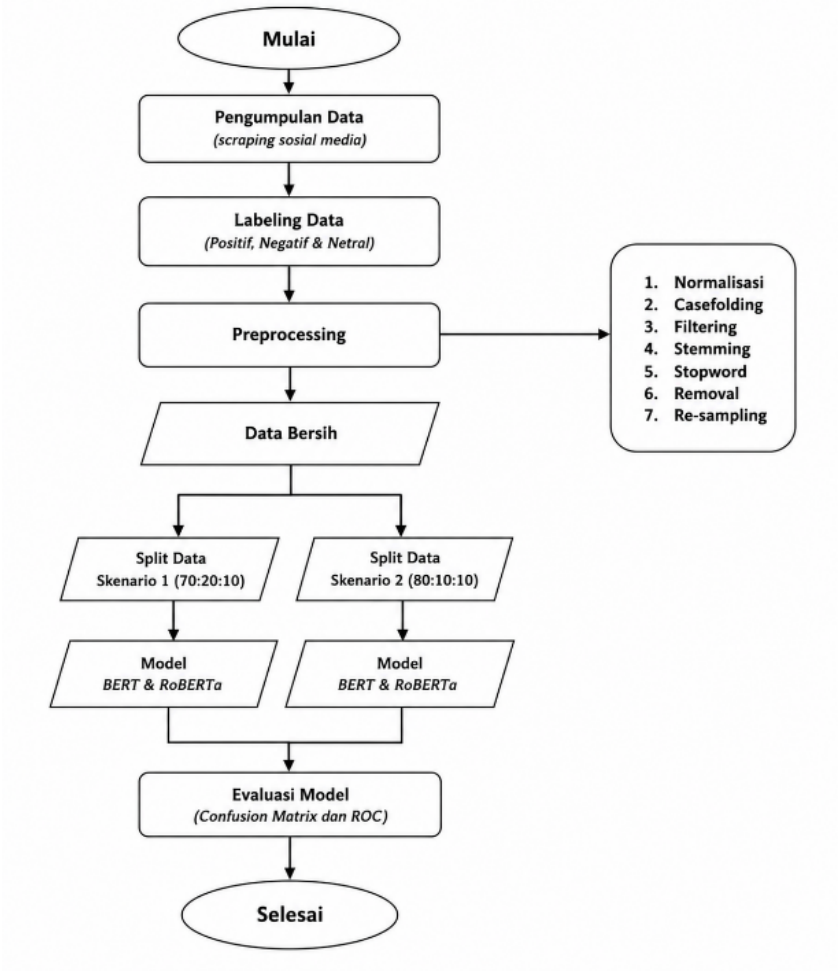
Metode BERT berfungsi dengan cara memahami keterkaitan antar kata secara dua arah, yang membuatnya lebih efektif dalam menangkap konteks kalimat dibandingkan dengan teknik yang lebih konvensional(Circuit & Circuit, 2023). Di sisi lain, RoBERTa adalah perkembangan dari BERT yang memiliki penyempurnaan dalam proses pelatihannya, sehingga dapat memberikan hasil yang lebih baik dan konsisten. Dengan menganalisis kedua metode tersebut, penelitian ini bertujuan untuk menentukan model yang paling optimal dalam mengklasifikasikan sentimen terkait isu gempa bumi dan tsunami megathrust di Indonesia.

Berdasarkan isu yang ada, studi ini mengusung judul “Analisis Sentimen mengenai Gempabumi dan Tsunami Megathrust di Indonesia dengan Pemanfaatan Metode BERT dan RoBERTa”. Melalui penelitian ini, diharapkan bisa memberikan wawasan tentang pandangan masyarakat terkait isu megathrust serta berfungsi sebagai acuan dalam pengembangan sistem analisis sentimen menggunakan deep learning di sektor kebencanaan(Sufi, 2024).

METODE PENELITIAN

A. Tahapan Penelitian

Tahapan penelitian dirancang untuk memastikan analisis sentimen mengenai isu gempabumi dan tsunami megatrust di Indonesia dilakukan secara terstruktur dan sistematis.



Gambar 1. Flowchart Alur Penelitian

B. Dataset

Metode pemilihan data dalam penelitian ini dilakukan dengan pendekatan purposive sampling, dimana pemilihan data berdasarkan kriteria tertentu yang relevan sesuai dengan tujuan penelitian(Campbell et al., 2020). Dalam hal ini, data yang dipilih untuk penelitian adalah komentar-komentar yang ada di platform sosial media X (sebelumnya *Twitter)*. Kriteria pemilihan sampel yang digunakan meliputi antara lain, (1) Data diambil

dari komentar-komentar sosial media X (sebelumnya *Twitter)*, menggunakan teknik *web scraping*, (2) Periode data dari 2012 sampai dengan 2025 selama 14 tahun rentang waktu periode tahun 2012 hingga Oktober 2025, (3) Relevansi teks komentar, hanya mengandung teks deskriptif minimal terdiri dari tiga kata; (4) Bahasa yang digunakan dalam ulasan adalah bahasa Indonesia(Campbell et al., 2020). Dataset mencakup 16.592 baris data ulasan dengan fitur *content* yang merupakan ulasan pengguna. Fitur *score* digunakan untuk proses klasifikasi ulasan pengguna ke dalam tiga kelas yaitu positif, netral, dan negatif.

C. Pra-pemrosesan Data

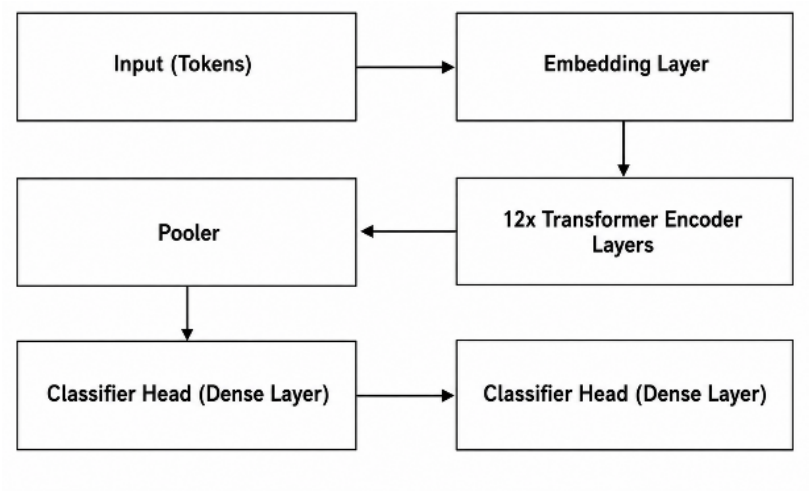
Pada tahap pra-pemrosesan data dilakukan untuk memastikan bahwa data dalam kondisi yang siap digunakan dalam proses pelatihan model. Beberapa langkah yang dilakukan dalam pra-pemrosesan data meliputi.

1. Normalisasi merupakan proses dalam mengoreksi ejaan Bahasa Indonesia dengan memanfaatkan fitur Google AI di Google Translate, dilakukan penerjemahan dua kali pada teks ulasan, yaitu dari Bahasa Indonesia ke Bahasa Inggris, lalu diterjemahkan lagi dari Bahasa Inggris ke Bahasa Indonesia.
2. *Case folding* merupakan proses menjadi semua karakter teks ulasan menjadi huruf kecil.
3. *Filtering* merupakan proses menghilangkan karakter teks yang mengandung angka, tanda baca, URL, kode ASCII, dan emoji.
4. *Stemming* merupakan proses mencari kata dasar dengan menghapus awalan dan akhiran dari suatu kata tertentu.
5. Penghapusan *Stopwords* merupakan proses penghapusan kata umum yang tidak memiliki makna dalam teks ulasan.
6. *Re-sampling imbalanced data* teknik yang digunakan untuk menyetarakan jumlah kelas minoritas mengikuti jumlah kelas mayoritas.

D. *Model Transformer*

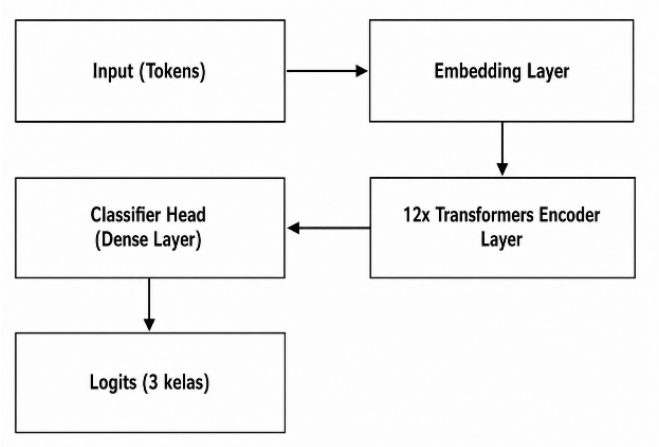
Model BERT yang digunakan dalam penelitian ini adalah *Bert For Sequence Classification* dari ekstensi *indobenchmark/indobert-base-p1*. Model *Bert For Sequence Classification* diadaptasi untuk tugas klasifikasi dengan menambahkan *classification*

head di atas representasi [CLS]. Pretrained model tersebut telah dilatih sebelumnya menggunakan *masked language modeling (MLM) objective* dan *next sentence prediction (NSP) objective*(Devlin et al., 2019)(Bashiri & Naderi, 2024).



Gambar 2. Arsitektur Model BERT

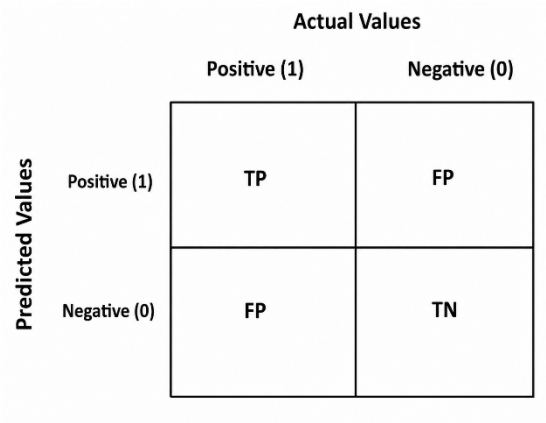
Model RoBERTa yang digunakan dalam penelitian adalah *Roberta For Sequence Classification* dari ekstensi model RoBERTa *flax-community/indonesian-roberta-base*(Wilie et al., 2019). RoBERTa base sentiment Clasifier adalah model klasifikasi teks sentimen yang didasarkan model RoBERTa. Model RoBERTaBase Indonesia yang telah dilatih sebelumnya, kemudian disempurnakan pada kumpulan data SmSA idonlu yang terdiri dari komentar dan ulasan(Liu et al., 2019)(Bashiri & Naderi, 2024)



Gambar 3. Arsitektur Model RoBERTa

E. Evaluasi Model

Evaluasi model dilakukan dengan mengukur sejauh mana model mampu mengklasifikasikan data secara benar. *Confusion Matrix* merupakan suatu alat yang digunakan untuk mengevaluasi kinerja model algoritma dengan membandingkan prediksi model dengan label yang sebenarnya. Ini memberikan pandangan yang lebih detail tentang bagaimana model melakukan klasifikasi, terutama dalam konteks klasifikasi dengan lebih dari dua kelas. *Confusion matrix* terdiri dari empat elemen utama yang membantu dalam menilai akurasi model deep learning dalam analisis sentimen antara lain, akurasi *(acuration),* presisi (*precision*), Sensitifitas (*recall*) dan *F1-score*(Chicco & Jurman, 2023).

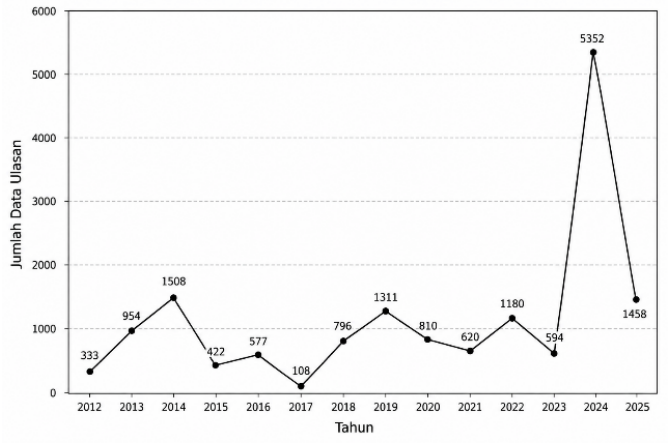


Gambar 4. Confusion Matrix

HASIL DAN PEMBAHASAN

Eksplorasi Data Awal

Eksplorasi data awal dilakukan untuk memahami karakteristik dataset yang digunakan penelitian.



Gambar 5. Tren Ulasan Pengguna Media Sosial X (Sebelumnya *Twitter*)

Gambar 5 menunjukkan tren ulasan pengguna media sosial X, Dimana data dikumpulkan dari ulasan pada platform media sosial X periode 2012 sampai dengan 2025 selama 14 tahun dengan total 16.592 data

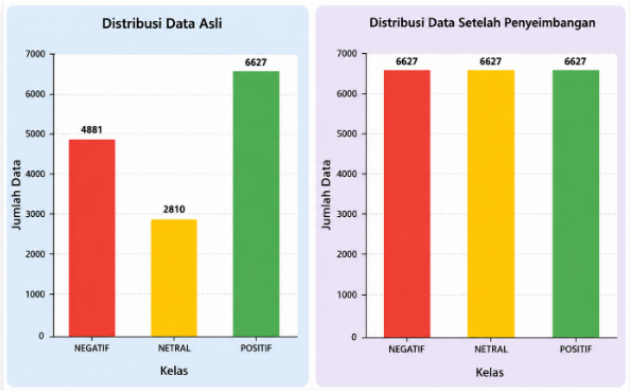
Analisis Hasil Pra-pemrosesan Data

Pada penelitian ini dilakukan pre-processing untuk memperbaiki struktur data sehingga menjadi lebih rapi dan terorganisir. Contoh hasil preprocessing.

Tabel 1. Contoh Hasil Keluaran Tahap Prapemrosesan Data

No	Masukan	Keluaran
1	[MAFIA PEDULI: MITIGASI ABOUT MEGATHRUST] Sempat terjadi peristiwa yang menggempaikan Indonesia karena berawal dari sebuah postingan di Facebook yang mengisyaratkan bahwa akan ada bencana hebat di Pulau Jawa hal ini membuat warga banyak yang khawatir maka BMKG pun angkat - https://t.co/pHm6nocXX8	[MAFIA PEDULI: MITIGASI MEGATHRUST] Ada kejadian yang menggempakan Indonesia karena bermula dari postingan di Facebook yang menandakan akan terjadi bencana besar di Pulau Jawa. Hal ini membuat banyak pihak khawatir sehingga BMKG mengambil tindakan - https://t.co/pHm6nocXX8
2	[2] .. siap siaga dan tidak gagap dalam menghadapi ancaman gempa dan tsunami. Kota Padang adalah kota yang memiliki potensi gempabumi dan tsunami dikarenakan letak pantainya di bagian barat berhadapan dengan zona sumber gempabumi Megathrust yang menurut para pakar memiliki..	[2] .. bersiaplah dan jangan ragu menghadapi ancaman gempa bumi dan tsunami. Kota Padang merupakan kota yang berpotensi terjadinya gempa bumi dan tsunami karena letak pantainya di sebelah barat menghadap zona sumber gempa Megathrust yang menurut para ahli...
3	[3] Kondisi Kalimantan sendiri merupakan pulau yang memiliki faktor kegempaan relatif lebih kecil dibandingkan pulau lainnya di Indonesia. Kalimantan kata dia tidak berada di Zona Megathrust atau zona tumbukan antara lempeng Indo-Australia dan Eurasia.	[3] Kalimantan sendiri merupakan pulau yang memiliki faktor kegempaan yang relatif lebih kecil dibandingkan pulau-pulau lain di Indonesia. Kalimantan, kata dia, tidak berada di Zona Megathrust atau zona tumbukan lempeng Indo-Australia dan Eurasia.

Tabel 1 menunjukkan contoh hasil prapemrosesan data ulasan pengguna berhasil mentransformasikan teks mentah menjadi dataset terstruktur melalui normalisasi, *case folding, filtering, stemming, re-sampling*, dan *label encoding*. Proses ini menghilangkan noise, menstandarisasi bentuk kata, mengurangi redundansi, serta memastikan keseimbangan dan konsistensi data.



Gambar 6. Grafik Distribusi Data Sebelum dan Sesudah Proses Penyeimbangan (*Balancing*)

Gambar 6 menunjukkan grafik distribusi data sebelum dan sesudah proses penyeimbangan *(balancing)*.Gambar ini merupakan perbandingan distribusi data sebelum dan sesudah proses penyeimbangan *(balancing)*.Dimana grafik distribusi data asli (kiri) menunjukkan bahwa dataset awal tidak seimbang *(imbalanced*). Kelas positif memiliki jumlah data terbanyak (6.627), diikuti negatif (4.881), sedangkan netral merupakan kelas dengan jumlah data paling sedikit (2.810). Ketidakseimbangan ini berpotensi menyebabkan model lebih cenderung memprediksi kelas mayoritas. Pada grafik distribusi data setelah penyeimbangan (kanan) menunjukkan bahwa seluruh kelas telah diseimbangkan menjadi 6.627 data untuk masing-masing kelas (negatif, netral, dan positif). Proses penyeimbangan ini bertujuan mengurangi bias model terhadap kelas mayoritas sehingga proses pelatihan menjadi lebih adil dan diharapkan menghasilkan performa klasifikasi yang lebih baik, terutama pada kelas yang sebelumnya memiliki jumlah data lebih sedikit.

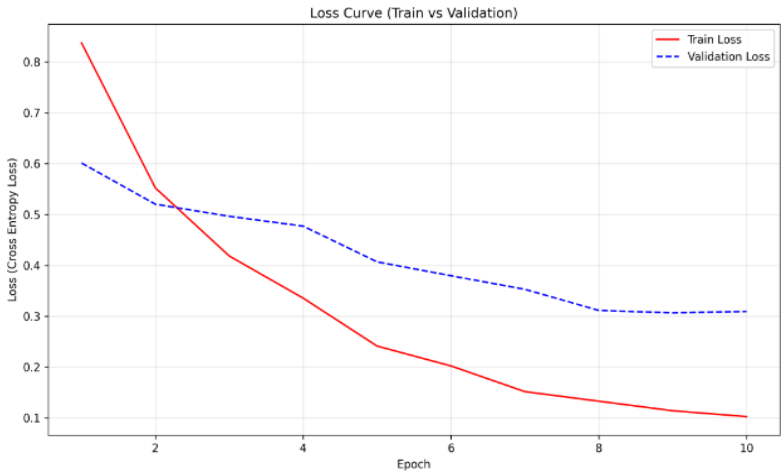
Pemodelan BERT

Pada model BERT tahap pemodelan dilakukan dengan 2 skenario percobaan (70:15:15) dan (80:10:10) dengan percobaan 10 epoch.

Tabel 2. Detail Loss Model BERT Skenario Percobaan 10 Epoch

Epoch	Skenario pembagian dataset			
	70:15:15		80:10:10	
	Loss Train	*Loss* Validasi	*Loss* Train	*Loss* Validasi
1.	0.8376	0.6013	0.8423	0.5917
2.	0.5522	0.5203	0.5456	0.5119
3.	0.4184	0.4965	0.3973	0.4540
4.	0.3355	0.4771	0.2989	0.4068
5.	0.2413	0.4069	0.2377	0.3039
6.	0.2022	0.3796	0.1849	0.2532
7.	0.1517	0.3529	0.1417	0.2008
8.	0.1329	0.3114	0.1252	0.1792
9.	0.1141	0.3066	0.1129	0.1730
10.	0.1026	0.3094	0.1000	0.1692

Berdasarkan Tabel 2 detail loss model BERT, dapat diamati pola perubahan nilai *loss* model BERT pada kedua skenario pembagian dataset (70:15:15 dan 80:10:10) selama proses pelatihan berlangsung sebanyak 10 *epoch*.



Gambar 7. Kurva Loss Skenario 70:15:15 Epoch 10 Pelatihan Model BERT

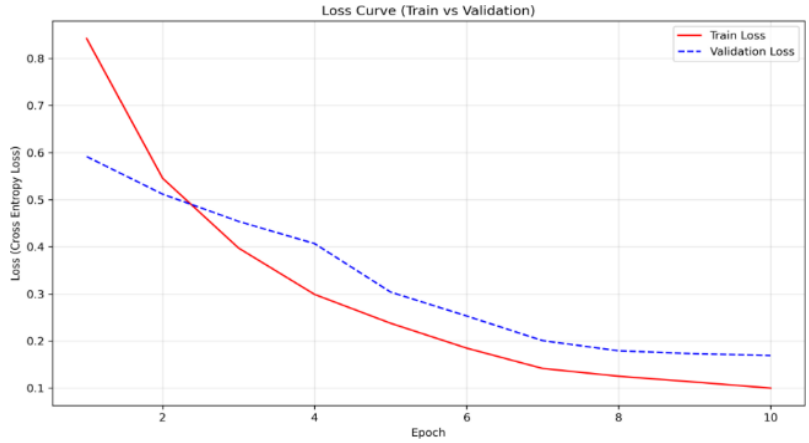
Gambar 7 menjelaskan kurva loss skenario 70:15:15 model BERT dengan 10 *epoch*, nilai *loss* train menurun secara konsisten dari 0,8376 (*epoch* 1) menjadi 0,1026 (*epoch* 10), atau berkurang sebesar 0,7350 (sekitar 87,8%). Pola serupa terjadi pada skenario 80:10:10, dengan *loss* train turun dari 0,8423 menjadi 0,1000 (penurunan sekitar 88,1%).

Penurunan yang konsisten dan cukup besar pada kedua skenario ini menunjukkan

bahwa model semakin baik dalam mempelajari pola pada data latih seiring bertambahnya jumlah *epoch*.

Berbeda dengan *loss* train, perilaku *loss* validasi pada kedua skenario menunjukkan perbedaan yang cukup mencolok. Pada skenario 70:15:15, *loss* validasi menurun secara konsisten dari 0,6013 (*epoch* 1) hingga mencapai titik terendah 0,3066 pada *epoch* 9, kemudian sedikit meningkat menjadi 0,3094 pada *epoch* 10.

Kenaikan ini relatif kecil (sekitar 0,0028), namun cukup untuk menandakan bahwa *epoch* ke-9 merupakan titik performa terbaik model pada skenario ini, sementara *epoch* ke-10 mulai menunjukkan kecenderungan *overfitting*.



Gambar 8. Kurva Loss Skenario 80:15:15 Epoch 10 Pelatihan Model BERT

Gambar 8 kurva loss skenario 80:10:10 model BERT dengan 10 *epoch* menunjukkan l*oss* validasi terus menurun secara konsisten tanpa mengalami kenaikan sama sekali sepanjang 10 *epoch*, dari 0,5917 (*epoch* 1) hingga mencapai titik terendah 0,1692 pada *epoch* 10 yaitu *epoch* terakhir itu sendiri. Penurunan ini sangat signifikan, yaitu sekitar 0,4225 atau 71,4% dari nilai awal. Hal ini mengindikasikan bahwa hingga *epoch* ke-10, model pada skenario 80:10:10 belum menunjukkan tanda-tanda *overfitting* yang berarti dan kemampuan generalisasinya terhadap data validasi masih terus meningkat.

Pemodelan RoBERTa

Pada model RoBERTa tahap pemodelan dilakukan dengan 2 skenario percobaan (70:15:15) dan (80:10:10) dengan percobaan 10 epoch.

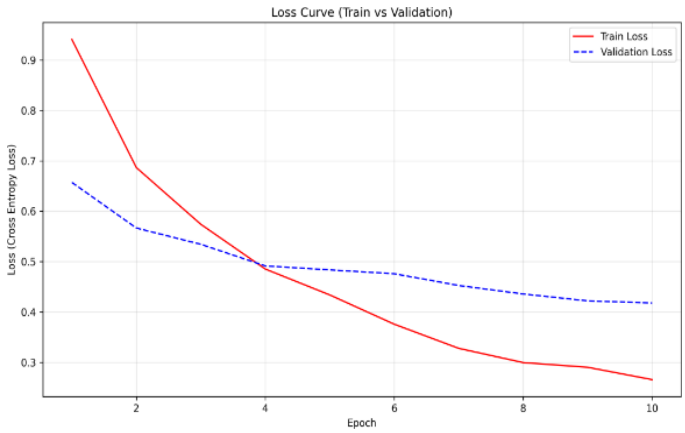
Tabel 3. Detail Loss Model RoBERTa Skenario Percobaan 10 Epoch

Epoch	Skenario pembagian dataset			
	70:15:15		80:10:10	
	Loss Train	*Loss* Validasi	*Loss* Train	*Loss* Validasi
1.	0.9404	0.7527	0.9409	0.6572
2.	0.6848	0.6162	0.6866	0.5667
3.	0.5773	0.5838	0.5741	0.5344
4.	0.5220	0.5183	0.4853	0.4915
5.	0.4561	0.5110	0.4339	0.4836
6.	0.4064	0.4713	0.3758	0.4760
7.	0.3536	0.4922	0.3279	0.4526
8.	0.3297	0.4827	0.2996	0.4358
9.	0.3042	0.4866	0.2903	0.4221
10.	0.2941	0.4990	0.2658	0.4180

Berdasarkan Tabel 3 menunjukkan pola perubahan nilai *loss* model RoBERTa pada kedua skenario pembagian dataset (70:15:15 dan 80:10:10) selama proses pelatihan berlangsung sebanyak 10 *epoch*.

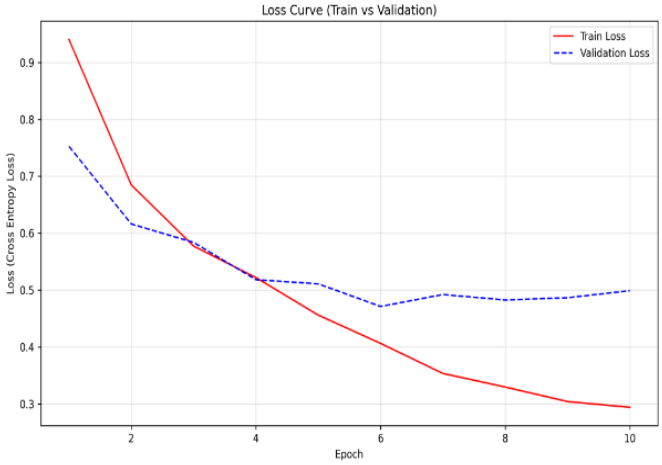
Pada skenario 70:15:15, *loss* train menurun secara konsisten dari 0,9404 (*epoch* 1) menjadi 0,2941 (*epoch* 10), atau berkurang sebesar 0,6463 (sekitar 68,7%). Pada skenario 80:10:10, *loss* train turun dari 0,9409 menjadi 0,2658 (penurunan sekitar 71,7%).

Pola penurunan yang konsisten pada kedua skenario menunjukkan bahwa model semakin baik dalam mempelajari pola pada data latih seiring bertambahnya *epoch*.



Gambar 9. Kurva Loss Skenario 70:15:15 Epoch 10 Pelatihan Model RoBERTa

Gambar 9 menjelaskan pada skenario 70:15:15, *loss* validasi menurun sampai titik terendah 0,4713 pada *epoch* ke-6, kemudian berfluktuasi naik-turun pada *epoch* berikutnya (0,4922 pada *epoch* 7, turun sedikit ke 0,4827 pada *epoch* 8, lalu kembali naik ke 0,4866 pada *epoch* 9, dan 0,4990 pada *epoch* 10) tanpa pernah kembali ke titik terendahnya. Pola ini menunjukkan bahwa *epoch* ke-6 merupakan titik performa terbaik model pada skenario ini, sementara *epoch-epoch* setelahnya mengindikasikan munculnya *overfitting*. Selain itu, *loss* train dan *loss* validasi berada pada level yang hampir sama (*crossover* sementara) di sekitar *epoch* ke-3 dan ke-4, sebelum *loss* validasi secara konsisten melampaui *loss* train mulai *epoch* ke-5 dan seterusnya. *Gap* antara keduanya kemudian melebar cukup signifikan, dari 0,0549 pada *epoch* 5 menjadi 0,2049 pada *epoch* 10.



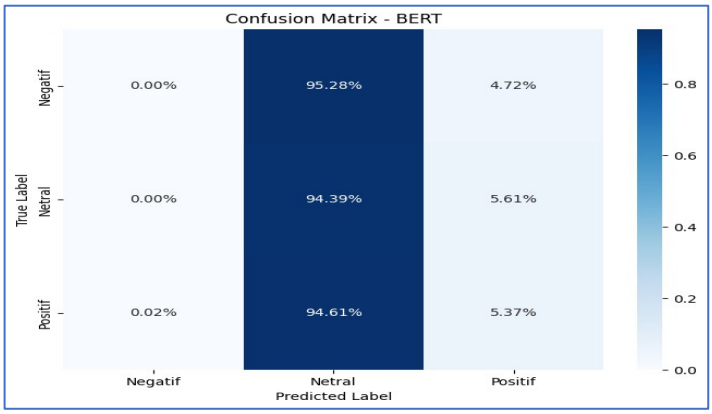
Gambar 10. Kurva Loss Skenario 80:15:15 Epoch 10 Pelatihan Model RoBERTa

Gambar 10 kurva loss model RoBERTa pada skenario 80:10:10 10 *epoch*, menunjukkan *loss* validasi menurun secara konsisten tanpa mengalami kenaikan sama sekali sepanjang 10 *epoch*, dari 0,6572 (*epoch* 1) hingga mencapai titik terendah 0,4180 pada *epoch* ke-10 yaitu *epoch* terakhir itu sendiri. Hal ini mengindikasikan bahwa model pada skenario ini belum menunjukkan tanda *overfitting* yang berarti dan kemampuan generalisasinya masih terus meningkat hingga akhir pelatihan.

Evaluasi Model BERT

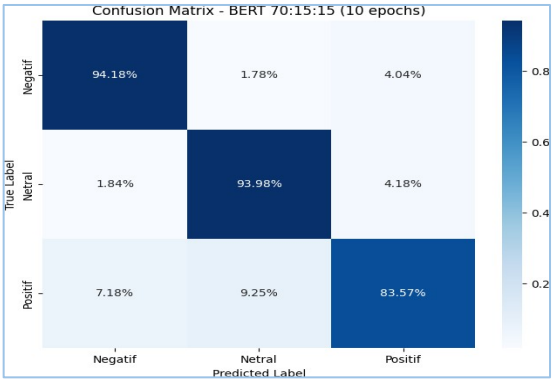
Dalam proses pembangunan sistem klasifikasi sentimen berbasis teks, evaluasi performa model menjadi tahap penting untuk memahami sejauh mana model mampu mengenali dan membedakan berbagai ekspresi sentimen dalam data. Pada penelitian ini,

dilakukan dua pendekatan utama, yaitu menggunakan model *pretrained indobenchmark/indobert-base-p1* secara langsung tanpa pelatihan lanjutan, serta pendekatan *fine-tuning* di mana model yang sama disesuaikan lebih lanjut menggunakan dataset ulasan pengguna X/*Twitter* dengan dua skenario pembagian dataset. Seperti yang ditampilkan pada Gambar 11,



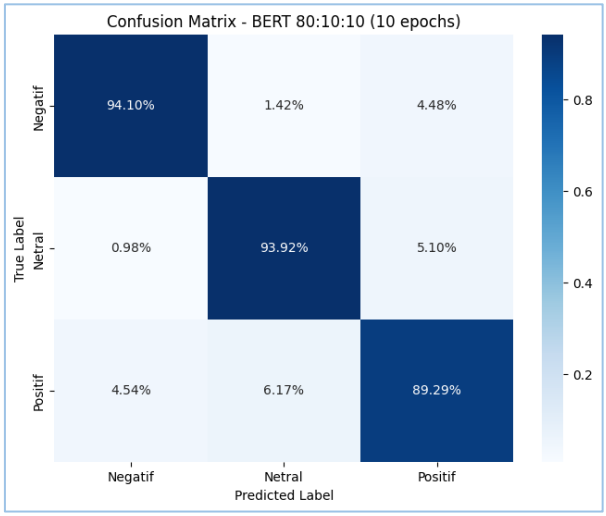
Gambar 11. Conffusion Matrix Pre-Trained model BERT

Gambar 11 Confusion matrix pre-trained model BERT menunjukkan model BERT dalam kondisi *pre-trained* tanpa *fine-tuning* menunjukkan performa yang sangat rendah dan belum dapat diandalkan. Sebagian besar prediksi model cenderung mengarah ke kelas netral.Hal ini seperti pada gambar 11 terlihat dari jumlah prediksi yang salah di mana sebanyak 95.28% data berlabel negatif dan 94.61% data berlabel positif diklasifikasikan sebagai netral. Prediksi yang benar untuk kelas negatif 0% dan untuk kelas positif hanya 0.02%. Ini menunjukkan adanya ketidakseimbangan model terhadap distribusi kelas yang menyebabkan model bias terhadap satu kelas dominan yaitu netral.



Gambar 12. Conffusion Matrix Fine-Tunning Model BERT Epoch 10 (Skenario 70:15:15)

Gambar 12 menunjukkan *confusion matrix fine-tunning* model BERT skenario 70:15:15 dengan epoch 10, dimana kelas Negatif berhasil diklasifikasikan dengan benar sebesar 94,18%, dengan sebesar 1,78% salah diprediksi sebagai Netral dan 4,04% diprediksi sebagai Positif. Kelas Netral memiliki performa yang hampir setara, dengan tingkat klasifikasi benar sebesar 93,98%, sementara 1,84% salah diprediksi sebagai Negatif dan 4,18% sebagai Positif. Kelas Positif menunjukkan performa yang paling rendah di antara ketiga kelas, dengan tingkat klasifikasi benar sebesar 83,57%, sedangkan sisanya tersebar cukup signifikan, 7,18% salah diprediksi sebagai Negatif dan 9,25% salah diprediksi sebagai Netral.



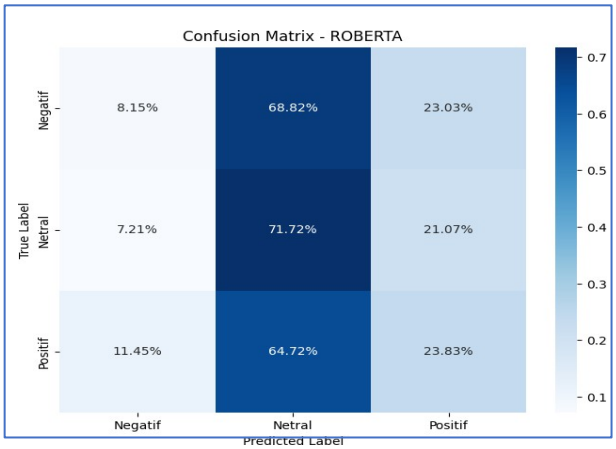
Gambar 13. Conffusion Matrix Fine-Tunning Model BERT Epoch 10 (Skenario 80:10:10)

Gambar 13 Confusion matrix fine-tunning model BERT skenario 80:10:10 epoch 10 menunjukkan kelas Negatif diklasifikasikan dengan benar sebesar 94,10%, dengan kesalahan 1,42% ke Netral dan 4,48% ke Positif. Kelas Netral memiliki tingkat klasifikasi benar sebesar 93,92%, dengan 0,98% salah diprediksi sebagai Negatif dan 5,10% sebagai Positif. Kelas Positif pada skenario ini mencapai tingkat klasifikasi benar sebesar 89,29%, lebih tinggi dibandingkan skenario 70:15:15 dengan kesalahan sebesar 4,54% ke Negatif dan 6,17% ke Netral.

Evaluasi Model RoBERTa

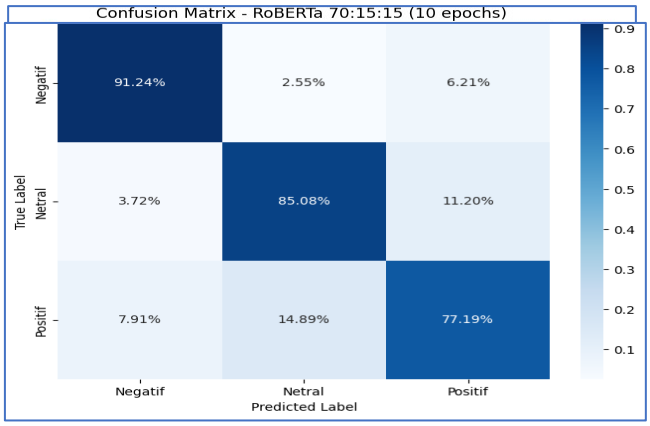
Pada penelitian ini, dilakukan dua pendekatan utama, yaitu menggunakan model *pre-trained flax-community/indonesian-RoBERTa-base* secara langsung tanpa pelatihan lanjutan, serta pendekatan *fine-tunning* di mana model yang sama disesuaikan lebih lanjut

menggunakan dataset ulasan pengguna X/*Twitter* dengan dua skenario pembagian dataset.



Gambar 14. Conffusion Matrix Pre-Trained Model RoBERTa

Berdasarkan gambar 14 Confusion matrix pre-trained model RoBERTa, menunjukkan model RoBERTa dalam kondisi *pre-trained* belum mampu melakukan klasifikasi sentimen secara optimal terhadap data ulasan pengguna *X/Twitter.* Hal ini terlihat dari dominasi prediksi pada kelas netral, di mana jumlah prediksi tertinggi berada pada kolom netral untuk seluruh label sebenarnya, baik positif, netral maupun negatif.Sebagai contoh, sebagian besar data yang seharusnya berlabel positif dan negatif justru diklasifikasikan sebagai netral, yang mencerminkan bahwa model masih cenderung memetakan berbagai ekspresi sentimen ke kelas yang netral.



Gambar 15. Conffusion Matrix Scenario 70:15:15 Epoch Fine-Tunning Model RoBERTa

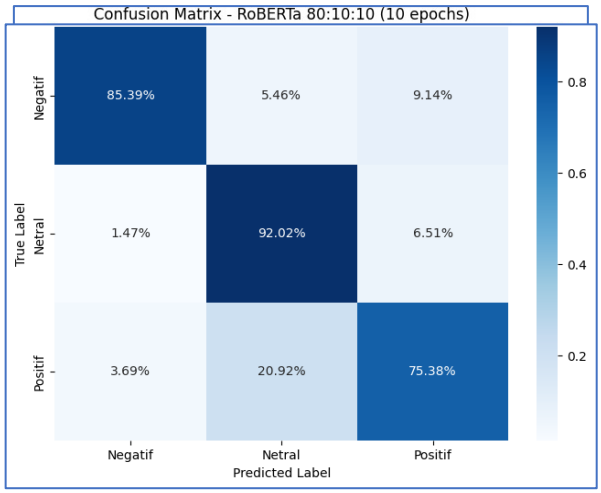
Gambar 15 Confusion matrix scenario 70:15:15 epoch 10 fine-tunning model RoBERTa, menunjukan kelas negative, model kinerja terbaik dengan tingkat klasifikasi

benar sebesar 91,24%. Ini menunjukkan bahwa model dapat mengidentifikasi sebagian besar data bersentimen negatif dengan sukses.

Pada kelas ini, kesalahan klasifikasi sangat kecil, 6,21% data negatif yang keliru diprediksi sebagai positif, dan 2,55% diprediksi sebagai netral. Ini menunjukkan bahwa sentimen negatif memiliki karakteristik linguistik yang cukup unik sehingga mudah dibedakan oleh model.

Pada kelas netral model memiliki akurasi sebesar 85,08% dalam mengklasifikasikan data, dimana kesalahan klasifikasi terbesar adalah 11,20% pada prediksi data netral positif, dan hanya 3,72% pada prediksi data netral negatif. Menurut pola ini, ada kemiripan dalam ekspresi sentimen netral dan positif. Ini membuat batas keputusan, atau batas keputusan, antara kedua kelas tersebut lebih sulit dipisahkan daripada batas kelas negatif.

Pada kelas positif, model mendapatkan nilai terendah dengan klasifikasi benar 77,19%. Kesalahan terbesar terjadi pada saat prediksi data netral sebesar 14,89% dan negatif sebesar 7,91%. Kesalahan prediksi yang tinggi menjadi netral ini menunjukkan bahwa ada tumpang tindih, atau overlap, fitur antara sentimen netral dan positif. Akibatnya, beberapa ekspresi positif yang cenderung halus atau tidak jelas dianggap netral.



Gambar 16. Conffusion Matrix Scenario 80:10:10 Epoch 10 Fine-Tunning Model RoBERTa

Gambar 16 Confusion matrix skenario 80:10:10 model RoBERTa epoch 10 fine-tunning, menunjukkan Model memperoleh recall sebesar 85,39% pada kelas Negatif. Sebagian besar data berlabel Negatif berhasil diprediksi dengan tepat, dengan kesalahan

klasifikasi sebesar 9,14% terprediksi sebagai Positif dan 5,46% sebagai Netral. Performa ini tergolong baik, meskipun sedikit lebih rendah dibandingkan skema 70:15:15. Kesalahan ke arah Positif yang relatif menonjol dapat dipengaruhi oleh teks negative yang turut memuat ungkapan harapan atau doa, sehingga sebagian terbaca sebagai sentimen positif oleh model.

Kelas Netral menunjukkan performa terbaik pada skema ini dengan recall sebesar 92,02%. Hampir seluruh data netral berhasil dikenali dengan benar, dengan kesalahan yang sangat kecil ke arah Positif (6,51%) maupun Negatif (1,47%). Hasil ini mengindikasikan bahwa dengan proporsi data latih yang lebih besar (80%), model semakin mampu mengenali karakteristik teks netral yang umumnya bersifat informatif dan faktual, seperti laporan parameter gempa, lokasi episenter, atau imbauan resmi.

Kelas Positif kembali menjadi kelas dengan recall terendah, yaitu 75,38%. Kesalahan klasifikasi paling dominan terjadi pada arah Positif → Netral yang cukup besar, yaitu 20,92%, sementara kesalahan ke arah Negatif relatif kecil (3,69%). Tingginya kebingungan antara Positif dan Netral menunjukkan bahwa model kesulitan membedakan ungkapan bersentimen positif dari pernyataan netral, terlebih ketika teks positif mengandung unsur informatif. Faktor jumlah data positif yang lebih sedikit (*class imbalance*) juga diduga turut memperburuk performa pada kelas ini.

Komparasi Model BERT dan RoBERTa

Tahap akhir penelitian ini adalah mengevaluasi performa dua model klasifikasi berbasis *transformer*, yakni BERT dan RoBERTa, dalam mengklasifikasikan sentimen data ulasan yang telah melalui keseluruhan tahapan prapemrosesan dan pelabelan.

Pengujian dilakukan pada dua skenario pembagian data, yaitu 70:15:15 dan 80:10:10 (latih:validasi:uji), serta pada dua durasi pelatihan, yakni 5 *epoch* dan 10 *epoch*, guna mengamati pengaruh proporsi data latih sekaligus jumlah iterasi pelatihan terhadap kinerja model.

Performa setiap konfigurasi diukur menggunakan empat parameter evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil komparasi disajikan pada Tabel 4 untuk pelatihan 5 *epoch* dan Tabel 5 untuk pelatihan 10 *epoch*. Secara umum, hasil pengujian memperlihatkan dua kecenderungan konsisten, penambahan proporsi data latih dari 70% menjadi 80% selalu meningkatkan kinerja, dan penambahan jumlah *epoch* dari 5 menjadi 10

turut mendorong peningkatan performa pada seluruh konfigurasi.

Tabel 4. Komparasi Evaluasi Model 5 Epoch

Parameter Evaluasi	BERT		RoBERTa	
	Skenario 70:15:15	Skenario 80:10:10	Skenario 70:15:15	Skenario 80:10:10
Accuracy	88,6022%	90,6494%	82,2544%	84,9505%
Precision	88,5685%	90,6185%	82,2199%	84,8718%
Recall	88,6022%	90,6494%	82,2544%	84,9505%
F1-score	88,5543%	90,6218%	82,1034%	84,8544%

Tabel 4 komparasi model BERT dan RoBERTa menunjukkan kinerja tertinggi pada keseluruhan pengujian. Pada skenario 70:15:15, model ini memperoleh *accuracy* sebesar 88,6022%, *precision* 88,5685%, *recall* 88,6022%, dan *F1-score* 88,5543%. Kinerja tersebut meningkat pada skenario 80:10:10, dengan *accuracy* mencapai 90,6494%, *precision* 90,6185%, *recall* 90,6494%, dan *F1-score* 90,6218%. Peningkatan ini setara dengan selisih sekitar 2,05 poin persen pada *accuracy* dan 2,07 poin persen pada *F1-score* ketika proporsi data latih ditambah dari 70% menjadi 80%.

Model RoBERTa menunjukkan pola peningkatan yang serupa, meskipun pada level kinerja yang lebih rendah. Pada skenario 70:15:15, model ini memperoleh *accuracy* 82,2544%, *precision* 82,2199%, *recall* 82,2544%, dan *F1-score* 82,1034%. Pada skenario 80:10:10, kinerjanya naik menjadi *accuracy* 84,9505%, *precision* 84,8718%, *recall* 84,9505%, dan *F1-score* 84,8544%. Selisih peningkatan antarskenario pada RoBERTa bahkan sedikit lebih besar dibanding BERT, yakni sekitar 2,70 poin persen pada *accuracy* dan 2,75 poin persen pada *F1-score*.

Apabila kedua model dibandingkan secara langsung pada skenario yang sama, BERT secara konsisten lebih unggul. Pada skenario 70:15:15, BERT mengungguli RoBERTa dengan selisih *accuracy* sebesar 6,35 poin persen (88,6022% berbanding 82,2544%) dan selisih *F1-score* sebesar 6,45 poin persen. Pada skenario 80:10:10, keunggulan BERT sedikit menyempit namun tetap signifikan, dengan selisih *accuracy* sebesar 5,70 poin persen (90,6494% berbanding 84,9505%) dan selisih *F1-score* sebesar 5,77 poin persen. Konsistensi keunggulan ini di seluruh metrik dan skenario mengindikasikan bahwa BERT

lebih sesuai dalam menangani karakteristik data ulasan pada penelitian ini dibanding RoBERTa.

Tabel 5. Komparasi Evaluasi Model 10 Epoch

Parameter Evaluasi	BERT		RoBERTa	
	Skenario 70:15:15	Skenario 80:10:10	Skenario 70:15:15	Skenario 80:10:10
Accuracy	90,5739%	92,4350%	88,8587%	91,8374%
Precision	90,5974%	92,4291%	88,8554%	91,8375%
Recall	90,5739%	92,4350%	88,8587%	91,8374%
F1-score	90,5064%	92,4292%	88,8565%	91,8363%

Tabel 5 komparasi evaluasi model BERT dan RoBERTa dengan 10 *epoch* menunjukkan hasil komparasi pada dua skenario pembagian data, yaitu 70:15:15 dan 80:10:10 (latih : validasi : uji), dengan pelatihan sebanyak 10 *epoch*. Kinerja kedua model diukur menggunakan empat parameter evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Secara umum, hasil pada tabel memperlihatkan dua kecenderungan utama, model BERT tetap mengungguli RoBERTa pada seluruh skenario, dan skenario pembagian data 80:10:10 menghasilkan kinerja yang lebih tinggi dibandingkan skenario 70:15:15 pada kedua model. Meskipun demikian, dibandingkan hasil pada 5 *epoch*, selisih kinerja antarkedua model pada 10 *epoch* menyempit secara signifikan.

Model BERT kembali mencatatkan kinerja tertinggi pada keseluruhan pengujian 10 *epoch*. Pada skenario 70:15:15, model ini memperoleh *accuracy* 90,5739%, *precision* 90,5974%, *recall* 90,5739%, dan *F1-score* 90,5064%. Kinerja tersebut meningkat pada skenario 80:10:10, dengan *accuracy* mencapai 92,4350%, *precision* 92,4291%, *recall* 92,4350%, dan *F1-score* 92,4292%. Peningkatan antar skenario ini setara dengan selisih sekitar 1,86 poin persen pada *accuracy* dan 1,92 poin persen pada *F1-score* ketika proporsi data latih ditambah dari 70% menjadi 80%. Konsistensi BERT dalam meraih nilai tertinggi di seluruh metrik menunjukkan kestabilan model ini dalam mengklasifikasikan sentimen data ulasan.

Model RoBERTa menunjukkan pola peningkatan yang serupa namun pada level kinerja yang sedikit lebih rendah. Pada skenario 70:15:15, model ini memperoleh *accuracy*

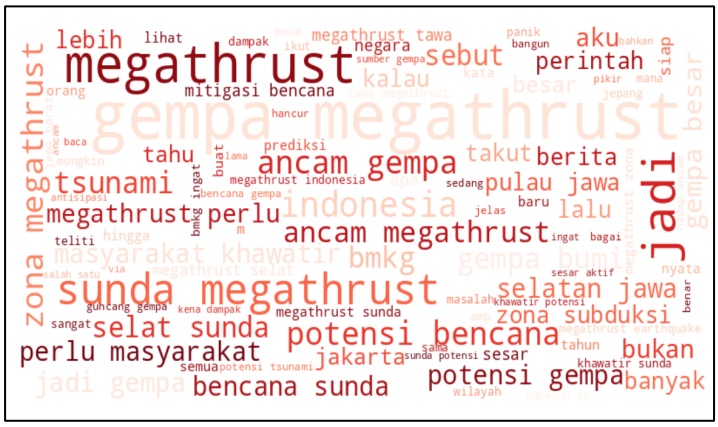
88,8587%, *precision* 88,8554%, *recall* 88,8587%, dan *F1-score* 88,8565%. Pada skenario 80:10:10, kinerjanya naik menjadi *accuracy* 91,8374%, *precision* 91,8375%, *recall* 91,8374%, dan *F1-score* 91,8363%. Selisih peningkatan antarskenario pada RoBERTa bahkan lebih besar dibanding BERT, yakni sekitar 2,98 poin persen pada *accuracy* maupun *F1-score*. Hal ini mengindikasikan bahwa RoBERTa lebih sensitif terhadap penambahan proporsi data latih dibandingkan BERT.

Apabila kedua model dibandingkan secara langsung pada skenario yang sama, BERT tetap unggul, namun dengan margin yang jauh lebih tipis dibanding pengujian 5 *epoch*. Pada skenario 70:15:15, BERT mengungguli RoBERTa dengan selisih *accuracy* sebesar 1,72 poin persen (90,5739% berbanding 88,8587%) dan selisih *F1-score* sebesar 1,65 poin persen. Pada skenario 80:10:10, keunggulan BERT menyusut drastis menjadi hanya sekitar 0,60 poin persen pada *accuracy* (92,4350% berbanding 91,8374%) dan 0,59 poin persen pada *F1-score*. Menyempitnya selisih ini, terutama pada skenario 80:10:10, mengindikasikan bahwa pada durasi pelatihan yang lebih panjang, kinerja RoBERTa mulai mendekati BERT, meskipun BERT masih mempertahankan posisinya sebagai model terbaik.

Secara keseluruhan, dari delapan konfigurasi yang diuji pada dua durasi pelatihan, performa tertinggi dicapai oleh model BERT dengan skenario pembagian data 80:10:10 pada pelatihan 10 *epoch*, dengan nilai *accuracy* sebesar 92,4350% dan *F1-score* sebesar 92,4292%. Konfigurasi inilah yang dipandang paling andal dan dipilih untuk digunakan pada tahap klasifikasi sentimen dalam penelitian ini.

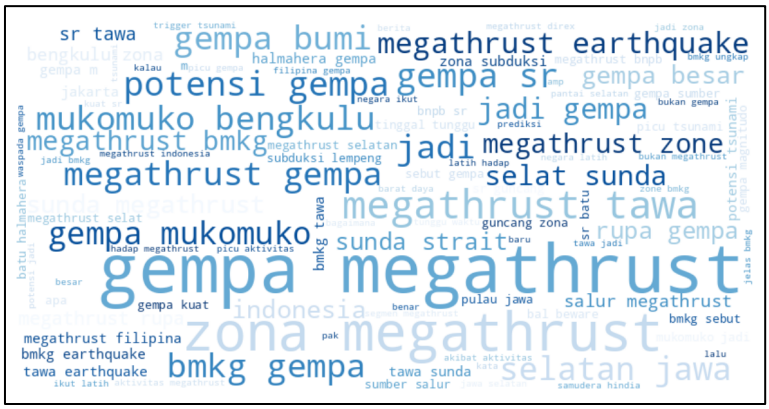
Analisis Sentimen Ulasan Pengguna *X/Twitter*

Setelah model berhasil mengklasifikasikan sentimen, tahap selanjutnya adalah menganalisis konten ulasan untuk mendapatkan wawasan yang komprehensif mengenai persepsi publik terhadap isu yang saat ini viral di masyarakat yaitu *"gempa dan tsunami megathrust"*. Analisis ini dilakukan dengan mengidentifikasi kata kunci dan frasa yang paling sering muncul pada setiap kategori sentimen. Untuk meringkas tema utama secara visual, digunakan metode *workcloud* seperti yang disajikan pada Gambar 17.



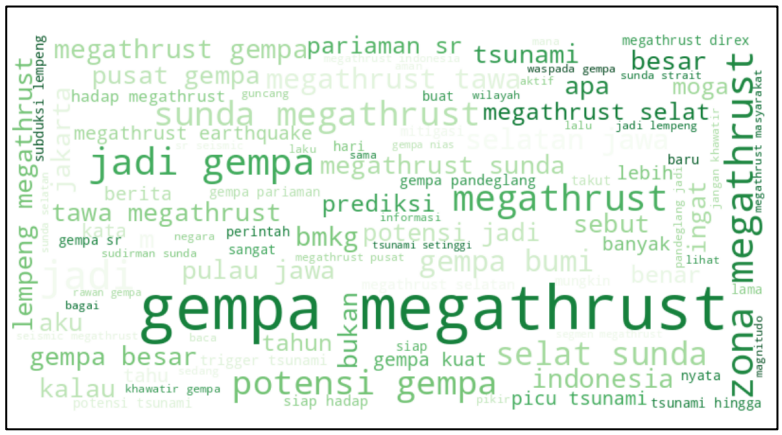
Gambar 17. Visualisasi Wordcloud Kata Kunci Ulasan Pengguna X/Twitter Sentiment Negative

Pada Gambar 17 Visualisasi *wordcloud* ini menampilkan kata-kata yang paling dominan untuk kategori sentimen negatif .Pada ulasan negatif, kata-kata dominan seperti *"ancam gempa"*, *"ancam megathrust"*, *"kena dampak"* dan *"khawatir potensi"* mencerminkan frasa negatif terhadap isu gempabumi dan tsunami *megathrust*.



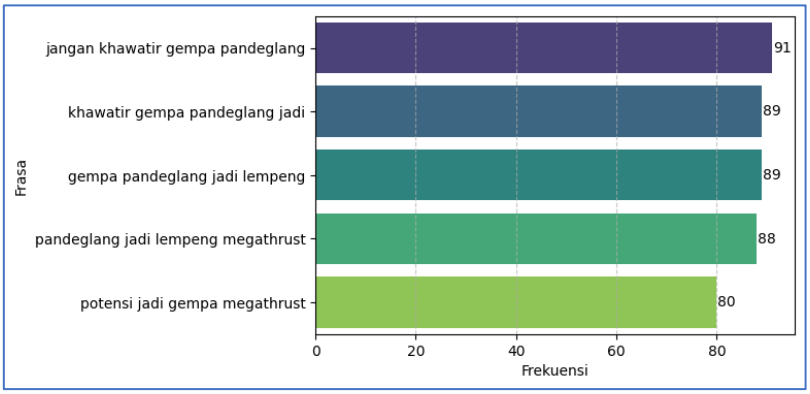
Gambar 18. Visualisasi Wordcloud Kata Kunci Ulasan Pengguna X/Twitter Sentiment Netral

Sementara itu, Pada Gambar 18 Visualisasi *wordcloud* sentimen netral tetap menampilkan kata kunci seperti "gempa megathrust", "potensi gempa" namun dengan tambahan frasa seperti "waspada gempa", "jangan khawatir" masih berfokus pada "gempa megathrust" namun disertai konteks yang menunjukkan kesiapsiagaan.



Gambar 19. Visualisasi Wordcloud Kata Kunci Ulasan Pengguna X/Twitter Sentiment Positif

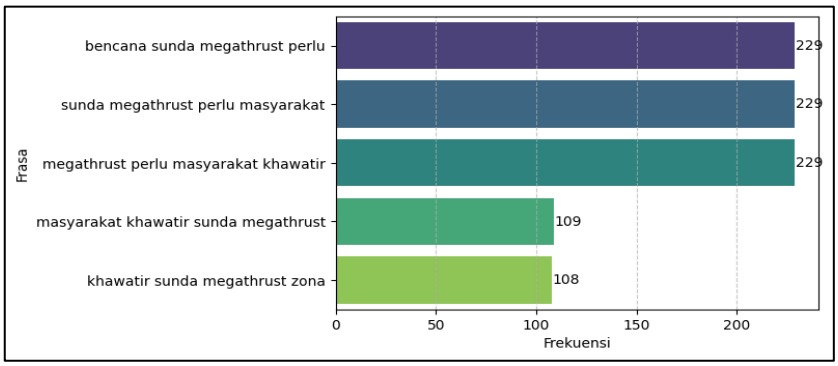
Pada Gambar 19 Visualisasi *wordcloud* kata kunci seperti "gempa megathrust", "potensi gempa" dengan tambahan frasa seperti "waspada gempa", "jangan khawatir" masih berfokus pada "gempa megathrust" namun disertai konteks yang menunjukkan kesiapsiagaan.Analisis dilanjutkan dengan mengidentifikasi frasa yang paling sering muncul di seluruh ulasan pengguna untuk memahami tema-tema utama yang menjadi fokus perhatian.



Gambar 20. Fourgram Kata Kunci Terbanyak Alasan Pengguna X/Twitter pada Sentiment Positif

Gambar 20, menunjukkan visualisasi frekuensi *fourgram* sentimen positif dengan jelas. Dari visualisasi tersebut, frasa yang paling dominan adalah *"jangan khawatir"* yang sering muncul diikuti dengan frasa *"gempa megathrust"*. Kemunculan frasa-frasa ini dengan frekuensi tinggi mengindikasikan kesiapsiagaan masyarakat terhadap isu yang viral saat ini yaitu *"gempa dan tsunami megathrust"*.

Analisis lebih mendalam terhadap ulasan bersentimen negatif, khususnya melalui identifikasi frasa kunci terlihat bahwa frasa yang paling dominan adalah *"khawatir zona megathrust sunda"*. Munculnya frasa ini menandakan bahwa masih kurangnya edukasi terhadap masyarakat sehingga menurunnya kesiapsiagaan ditunjukkan pada gambar 21.



Gambar 21. Fourgram Kata Kunci Terbanyak Alasan Pengguna X/Twitter pada Sentiment Negatif

KESIMPULAN DAN REKOMENDASI

Penelitian ini menunjukkan bahwa tahapan *preprocessing* yang sistematis, meliputi normalisasi teks, pembersihan data, *stemming*, penghapusan *stopwords*, dan *re-sampling*, mampu menghasilkan dataset yang berkualitas untuk analisis sentimen. Dalam evaluasi model dengan empat skenario, skenario 80:10:10 epoch 10 menunjukkan model terbaik pada kedua model BERT dan RoBERTa. Konfigurasi performa model BERT memiliki akurasi 92,4350%, presisi 92,4291%, recall 92,4350%, dan skor F1-nya 92,4292%. Sebaliknya, konfigurasi model RoBERTa memiliki akurasi 91.8374%, presisi 91,8375%, recall 91,8374%,dan skor F1-nya 91,8363%. Hasil ini menunjukkan bahwa BERT lebih cocok dengan linguistik ulasan bahasa Indonesia dalam aplikasi publik yang dominan.

Penelitian selanjutnya disarankan memperluas korpus data dan menerapkan pendekatan klasifikasi yang lebih komprehensif, seperti *multi-label* atau *aspect-based sentiment analysis*, agar mampu menggali opini masyarakat secara lebih spesifik. Selain itu, penggunaan model Transformer lain, seperti IndoBERT, XLNet, atau ALBERT, serta penyempurnaan proses pelabelan data secara manual dapat dilakukan untuk meningkatkan akurasi dan mengurangi bias klasifikasi. Di sisi praktis, hasil penelitian ini dapat menjadi

masukan bagi pemerintah dan instansi terkait untuk memperkuat edukasi serta strategi komunikasi publik guna meningkatkan kesiapsiagaan masyarakat dalam menghadapi potensi gempabumi dan tsunami megathrust.

REFERENSI

Bashiri, H., & Naderi, H. (2024). Comprehensive review and comparative analysis of transformer models in sentiment analysis. In *Knowledge and Information Systems* (Vol. 66, Issue 12). Springer London. https://doi.org/10.1007/s10115-024-02214-3

Campbell, S., Greenwood, M., Prior, S., Walkem, K., Young, S., & Bywaters, D. (2020). *Purposive sampling : complex or simple ? Research case examples*. https://doi.org/10.1177/1744987120927206

Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, 1–23. https://doi.org/10.1186/s13040-023-00322-4

Circuit, M. I., & Circuit, B. I. (2023). *Jurnal media informatika budidarma*. *7*, 2022–2023. https://doi.org/10.30865/mib.v7i1.5380

Contreras, D., Wilkinson, S., & Balan, N. (2022). *Assessing post-disaster recovery using sentiment analysis : The case of L ' Aquila ,*. https://doi.org/10.1177/87552930211036486

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 4171–4186.

Fakhruddin, M. R., Wijaya, R., & Bijaksana, M. A. (2024). *Analyzing Public Sentiments on Disaster Relief Efforts Through Social Media Data*. *11*(1), 44–50.

Firdaus, A. R., Ayudewi, F. M., Firdonsyah, A., Studi, P., Informasi, T., Yogyakarta, U. A., Sleman, K., Yogyakarta, D. I., Embeddings, C. W., Seimbang, D. T., & Data, T. A. (2026). *ANALISIS SENTIMEN MULTI-PLATFORM DIGITAL MENGGUNAKAN BERT*. *10*(1), 1220–1226.

Hutchings, S. J., & Mooney, W. D. (2021). The Seismicity of Indonesia and Tectonic Implications. *Geochemistry, Geophysics, Geosystems*, *22*(9), 1–42. https://doi.org/10.1029/2021GC009812

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. *1*.

Naufal, G., Perdana, R., Muhammad, A., & Rusli, R. (2023). *ANALYSIS OF TWITTER SOCIAL MEDIA FUNCTIONS IN RESPONDING TO THE CIANJUR EARTHQUAKE IN INDONESIA*. *8*(2), 165–182.

Seneviratne, K., Nadeeshani, M., & Senaratne, S. (2024). *Use of Social Media in Disaster Management : Challenges and Strategies*.

Srivastava & Soni, 2022; Tan et al., 2023. (2023). *Analisis Sentimen Terhadap Kemajuan Kecerdasan Buatan di*. *9*(November), 137–146. https://doi.org/10.34128/jsi.v9i2.649

Sufi, F. (2024). (2024). *A Sustainable Way Forward : Systematic Review of Transformer Technology in Social-Media-Based Disaster Analytics*.

Supendi, P., Widiyantoro, S., Rawlinson, N., Yatimantoro, T., Muhari, A., Hanifa, N. R., Gunawan, E., Shiddiqi, H. A., Imran, I., Anugrah, S. D., Daryono, D., Prayitno, B. S., Adi, S. P., Karnawati, D., Faizal, L., & Damanik, R. (2023). On the potential for megathrust earthquakes and tsunamis off the southern coast of West Java and southeast Sumatra, Indonesia. *Natural Hazards*, *116*(1), 1315–1328. https://doi.org/10.1007/s11069-022-05696-y

Tan, K. L., & Lee, C. P. (2023). *applied sciences A Survey of Sentiment Analysis : Approaches , Datasets , and Future Research*.

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2019). *IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*.