
Perbandingan Kinerja Model IndoBERT Base Uncased dan IndoBERTweet untuk Analisis Emosi Komentar YouTube terkait RUU Perampasan Aset

Bima Kaka Bani Adam

Sistem Informasi, UPN “Veteran” Jawa Timur

22082010007@student.upnjatim.co.id

Abstract

Perkembangan media sosial, khususnya YouTube, telah menjadikan komentar pengguna sebagai sumber data yang kaya untuk menganalisis opini dan emosi publik terhadap isu-isu sosial dan kebijakan pemerintah. Salah satu isu yang menarik perhatian adalah Rancangan Undang-Undang (RUU) Perampasan Aset yang memunculkan berbagai respons emosional dari masyarakat. Penelitian ini bertujuan untuk membandingkan kinerja dua model berbasis transformer, yaitu IndoBERT Base Uncased dan IndoBERTweet, dalam melakukan klasifikasi emosi pada komentar YouTube terkait isu tersebut. Data diperoleh melalui scraping menggunakan YouTube Data API v3 dari beberapa kanal berita, kemudian dilakukan tahap preprocessing yang meliputi cleaning, normalisasi teks, dan case folding. Pelabelan emosi dilakukan secara otomatis menggunakan beberapa pretrained model dari Hugging Face dengan kategori emosi yaitu senang, sedih, marah, takut, dan cinta. Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20. Hasil eksperimen menunjukkan bahwa IndoBERTweet memberikan performa yang lebih baik dibandingkan IndoBERT Base Uncased. IndoBERTweet memperoleh accuracy sebesar 0,80 dengan F1-score macro 0,65, sedangkan IndoBERT Base Uncased memperoleh accuracy sebesar 0,78 dengan F1-score macro 0,61. Hasil ini menunjukkan bahwa IndoBERTweet lebih adaptif dalam menangani teks informal seperti komentar media sosial.

Keywords—Analisis Emosi, IndoBERT-base uncased, IndoBERTweet, RUU Perampasan Aset.

I. PENDAHULUAN

Media sosial, khususnya platform berbagi video seperti YouTube, telah berkembang menjadi platform digital yang dapat digunakan masyarakat untuk menyampaikan pendapat dan perasaan mereka tentang berbagai masalah publik. Komentar pengguna pada video YouTube tidak hanya merepresentasikan tanggapan individu, tetapi juga mencerminkan dinamika persepsi publik terhadap suatu kebijakan atau fenomena sosial. Hal ini memberikan keuntungan bagi masyarakat, untuk lebih aktif dalam memahami dan menanggapi isu-isu politik yang sedang berkembang[1]. Salah satu isu yang menarik perhatian masyarakat adalah Rancangan Undang-Undang (RUU) Perampasan Aset, yang menimbulkan beragam reaksi emosional seperti senang, marah, takut, dan sedih.

Analisis emosi menjadi aspek penting dalam memahami respons publik secara lebih mendalam dibandingkan analisis sentimen, karena mampu mengidentifikasi jenis emosi yang lebih spesifik dalam teks [2]. Dalam konteks ini, perkembangan teknologi Natural Language Processing (NLP),

khususnya berbasis deep learning, telah memberikan solusi efektif dalam melakukan klasifikasi emosi secara otomatis. Model berbasis Bidirectional Encoder Representations from Transformers (BERT) telah berhasil melakukan sejumlah tugas pemrosesan bahasa alami, seperti mengklasifikasikan teks berbahasa Indonesia. [3][4][5].

IndoBERT merupakan salah satu model prelatih yang dikembangkan khusus untuk bahasa Indonesia dan telah digunakan secara luas dalam berbagai tugas NLP, termasuk analisis sentimen dan emosi. Meskipun demikian, IndoBERT dilatih menggunakan korpus formal sehingga memiliki keterbatasan dalam menangani teks informal yang banyak ditemukan di media sosial [6]. Hal ini menjadi tantangan dalam analisis komentar YouTube yang umumnya mengandung bahasa tidak baku, singkatan, serta slang.

Sebagai alternatif, IndoBERTweet dikembangkan dengan memanfaatkan data dari media sosial seperti Twitter, sehingga lebih adaptif terhadap karakteristik bahasa informal dan

kontekstual [7]. Model ini menunjukkan performa yang lebih baik dalam memahami teks media sosial karena mampu menangkap variasi bahasa yang lebih kompleks dan dinamis.

Penelitian ini bertujuan untuk membandingkan kinerja model IndoBERT Base Uncased dan IndoBERTtweet dalam analisis emosi komentar YouTube terkait RUU Perampasan Aset. Penelitian ini diharapkan dapat memberikan insight mengenai model yang paling efektif dalam menangani teks berbahasa Indonesia pada konteks media sosial, serta berkontribusi dalam pengembangan metode analisis emosi berbasis NLP.

II. TINJAUAN PUSTAKA

A. Emosi

Emosi merupakan respons psikologis individu terhadap suatu peristiwa yang dapat mempengaruhi perilaku dan cara berpikir. Menurut [8], emosi adalah reaksi alami manusia terhadap pengalaman yang dirasakan, yang dapat diekspresikan melalui kata-kata, intonasi, maupun perilaku nonverbal. Setiap emosi memiliki fungsi dalam membantu individu merespons situasi, seperti marah terhadap ketidakadilan, takut untuk menghindari ancaman, serta senang sebagai bentuk kepuasan. Dalam konteks sosial, emosi juga mencerminkan sikap masyarakat terhadap suatu isu.

B. Text Mining

Text mining adalah proses menggali informasi dan pola tersembunyi dari data teks yang tidak terstruktur menggunakan metode statistik dan machine learning. Menurut [9], text mining memiliki tujuan untuk mengekstraksi makna dan hubungan antar kata dalam teks agar dapat digunakan untuk analisis lebih lanjut. Text mining merupakan tahap penting dalam analisis teks, khususnya pada Natural Language Processing (NLP), karena mengubah data tidak terstruktur menjadi lebih bersih dan terstandarisasi sehingga meningkatkan kualitas fitur dan akurasi model.

C. Deep Learning

Deep learning merupakan cabang dari machine learning yang meniru cara kerja otak manusia dalam memproses informasi dengan menggunakan arsitektur jaringan saraf tiruan (deep neural network).

Menurut [10], pembelajaran deep learning memiliki kemampuan untuk secara otomatis

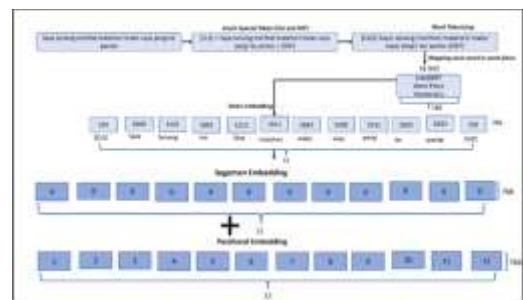
mengekstraksi fitur dari data tanpa memerlukan intervensi manual. Ini menghasilkan sistem yang lebih fleksibel, efektif dan mampu menangani jumlah data yang sangat kompleks. Hal ini menjadikan deep learning sangat efektif digunakan dalam berbagai tugas analisis data, termasuk klasifikasi teks dan analisis emosi pada data tidak terstruktur.

D. IndoBERT

IndoBERT merupakan model bahasa berbasis arsitektur *Bidirectional Encoder Representations for Transformer* (BERT) yang dikembangkan khusus untuk Bahasa Indonesia, sehingga lebih optimal dalam menangani teks berbahasa Indonesia dibandingkan BERT standar yang dilatih menggunakan bahasa Inggris.

Model ini diperkenalkan pertama kali oleh [6], pada pelatihan indobert ini dilatih menggunakan lebih dari 220 juta kata yang dikumpulkan dari sumber Wikipedia, artikel berita, dan korpus web Indonesia. Sementara itu, varian lainnya [11], dilatih pada 26 juta cuitan berbahasa Indonesia yang kaya singkatan, dan banyak kata tidak baku.

Secara arsitektur, IndoBERT menggunakan mekanisme transformer encoder yang bersifat dua arah (bidirectional) dalam memahami konteks kata. Model ini memproses informasi dari kiri ke kanan dan sebaliknya, sehingga mampu menghasilkan representasi konteks yang lebih akurat. Ilustrasi proses tokenisasi pada IndoBERT dapat dilihat pada Gambar 1.



Gambar 1. IndoBERT tokenizer

E. Confusion Matrix

Metode evaluasi Confusion Matrix digunakan untuk mengevaluasi kinerja model klasifikasi sentimen. menurut [12] confusion matrix menampilkan data dalam bentuk tabel dua dimensi yang berisi jumlah prediksi benar dan salah untuk setiap kelas.

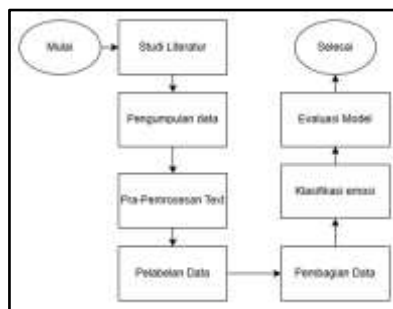
True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) adalah empat komponen utama dari confusion matrix. Keempat nilai ini digunakan untuk menghitung metrik penting seperti akurasi, presisi, recall, dan skor F1, yang digunakan untuk mengevaluasi kinerja model klasifikasi secara keseluruhan.

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad recall = \frac{TP}{TP + FN}$$

$$precision = \frac{TP}{TP + FP} \quad f1 - score = 2 \frac{precision \times recall}{precision + recall}$$

III. METODELOGI

Bagian metodologi penelitian ini menjelaskan tahapan-tahapan yang digunakan agar proses penelitian dapat berjalan secara sistematis dan terstruktur. Alur penelitian tersebut ditampilkan pada Gambar 2 berikut ini.



Gambar 2 Alur Penelitian

Pada setiap proses pada gambar 2 akan dijelaskan dalam poin poin sebagai berikut.

A. Studi Literatur

Studi literatur adalah proses sistematis untuk menelusuri, menelaah, dan menganalisis berbagai sumber ilmiah yang relevan dengan topik penelitian. Kegiatan ini dilakukan untuk memperoleh pemahaman mendalam terhadap teori, konsep, serta hasil penelitian sebelumnya yang menjadi dasar dalam penyusunan kerangka konseptual yang menjadi dasar perancangan metode dan interpretasi hasil penelitian [13].

B. Pengumpulan Data

Data yang dikumpulkan berasal dari komentar video Youtube yang membahas topik Rancangan Undang-Undang (RUU) Perampasan Aset. Pengumpulan data dilakukan melalui scrapping dengan bantuan YouTube Data API v3, yang memungkinkan pengambilan data secara otomatis berdasarkan tautan video. Hasil data yang diperoleh

kemudian disimpan dalam format excel (.xlsx) agar dapat dengan mudah diolah dan dianalisis pada tahap selanjutnya.

C. Pra-Pemrosesan Text

Setelah dilakukan pengumpulan data selanjutnya adalah melakukan pra-pemrosesan text. Menurut [14], tahap pra-pemrosesan berfungsi untuk menghilangkan elemen-elemen yang tidak relevan dan meminimalkan noise pada data, sehingga meningkatkan akurasi dalam proses analisis.

D. Pelabelan

Pelabelan pada penelitian ini bertujuan untuk memberikan kategori emosi pada setiap komentar yang telah dikumpulkan. Proses pelabelan dilakukan secara otomatis, yaitu dengan memanfaatkan pre-trained model. Pada penelitian ini, pelabelan emosi menggunakan beberapa model yang diakses melalui library Transformers dari Python untuk memanggil model dari platform Hugging Face. Pendekatan ini memungkinkan model berperan sebagai annotator otomatis tanpa keterlibatan manusia secara langsung.

Kategori emosi yang digunakan dalam penelitian ini terdiri dari lima kelas, yaitu senang, sedih, marah, takut, dan cinta. Pendekatan ini juga lebih efisien waktu dibandingkan pelabelan manual yang membutuhkan waktu dan tenaga lebih besar.

E. Pembagian Data

Selanjutnya dilakukan pembagian data menjadi dua bagian utama, yaitu data latih (training data) untuk pelatihan model dan data uji (testing data) untuk mengevaluasi kinerja model yang sudah terlatih. Proses pembagian data menggunakan metode Hold-Out, Menurut [16] Hold-Out memberikan keseimbangan yang baik karena proses pembagian hanya dilakukan sekali. Pada skripsi ini dilakukan dua skenario Rasio pembagian data latih dan data testing dengan proporsi 80:20.

F. Klasifikasi Emosi

Pada tahap ini dilakukan klasifikasi emosi dengan menggunakan data train. Proses ini dilakukan menggunakan model Pretrained indobERT, yaitu model berbasis arsitektur Transformer yang telah dilatih secara khusus pada korpus Bahasa Indonesia. Dalam penelitian ini digunakan tiga varian pretrained indobERT, yaitu

Indolem/indoBERT-base-uncased, dan IndoBERTweet, untuk membandingkan performa masing-masing model dalam mengenali dan mengklasifikasikan emosi pada teks Indonesia.

G. Evaluasi

Evaluasi dilakukan menggunakan confusion matrix untuk menghitung nilai accuracy, precision, recall, dan f1-score. Metrik tersebut digunakan untuk mengukur kemampuan model dalam memprediksi label secara tepat dan seimbang. Hasil evaluasi memberikan gambaran mengenai seberapa besar akurasi setiap model pretrained dalam membantu model dalam analisis sentimen dan klasifikasi emosi.

IV. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada tahap ini, dilakukan pengambilan data komentar YouTube yang berasal dari beberapa channel media, yaitu Metro, Kompas.com, dan Official iNews. Data yang diperoleh meliputi tanggal publikasi komentar, username pengguna yang memberikan komentar, serta isi komentar yang digunakan sebagai bahan utama analisis. Dataset yang telah dikumpulkan kemudian disimpan dan dipersiapkan untuk tahap pra-pemrosesan. Hasil dari scarping dapat dilihat pada gambar 3 berikut ini.

	Date	User	Comment
0	2025-11-24 00:34:54	@RuhdaRuhda	ternyata anggota DPR tidak berani membela UU ya...
1	2025-11-24 09:31:12	@RuhdaRuhda	udah gak coak banyak alasan coak dijamin UU ya...
2	2025-11-21 04:02:48	@maemo-wbnc	Ya belum jadi lagi nih
3	2025-11-18 05:25:32	@Suandhara-pfu	akhir masih harus suu tempo kerubahan wujud...
4	2025-11-14 22:18:35	@PambyWaya-71	Kata Henggah akhir Desember mau Paripolason Asat...
...
10185	2025-05-08 11:48:38	@yusufwanto975	Maafkan, Zuhdi, Orlan Anangga betali, E...
10186	2025-05-08 18:19:48	@KawalaAmara	yg penting langsung dibales aja itu dihapus...
10187	2025-05-08 17:37:47	@SaputraArihan-43r	BNATANG RAKUS BERDAS. PENYUSAH RAKYAT. SEMOO...
10188	2025-05-08 17:31:31	@RozChar23	Hanya akan hanya 200rb, itu sudah sangat benci...
10189	2025-05-08 17:18:33	@RohmanGub-06	Das terus #GASURADIKORPRI

Gambar 3. hasil pengumpulan dataset

3.2 Pra-Pemrosesan teks

Pada tahap ini, data mentah yang telah dikumpulkan diproses menjadi data yang lebih bersih dan siap digunakan dalam pemodelan. Proses pra-pemrosesan meliputi penghapusan data duplikat serta elemen yang tidak relevan seperti simbol, URL, dan tanda baca. Selain itu, normalisasi dan case folding diterapkan sehingga teks menjadi lebih konsisten.

Hasil dari tahap ini menunjukkan bahwa data menjadi lebih terstruktur dan seragam, sehingga dapat meningkatkan kualitas input pada model.

Contoh hasil pra-pemrosesan data dapat dilihat pada Gambar 4 berikut.

Teks	Pra-pemrosesan
👍👍👍 lucu kocak para koruptor takut hartanya dirampas... tapi kalau merampas harta rakyat para koruptor ini ugah'n	lucu kocak para koruptor takut hartanya dirampas tapi kalau merampas harta rakyat para koruptor ini ugah'n
Paksa DPR untuk menajektorian UU perampasan aset koruptor Tanpa pandang bulu UU hukuman mati juga menyusul untuk koruptor	paksa DPR untuk menajektorian uu perampasan aset koruptor tanpa pandang bulu uu hukuman mati juga menyusul untuk koruptor

Gambar 4. Hasil Pra-pemrosesan teks

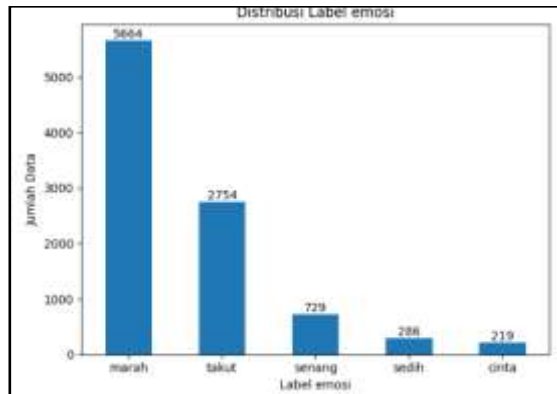
3.3 Pelabelan

Data yang sudah di bersihkan sebelumnya, kemudian dilakukan pelabelan otomatis menggunakan lima pre-trained model yang diperoleh dari platform Hugging Face dengan menggunakan library transformer. Pelabel pre-trained model tersebut yaitu galennolan, ali, stevenlimcorn, AptaArkana, akahana, dan tabularisasi. Hasil dari pelabel tersebut bisa dilihat pada gambar 5 dibawah ini.

Komentar bersih galennolan	ali	stevenlimcorn	AptaArkana	akahana
ya belum jadi lagi	sedih	sedih	senang	sedih
akhir masih harus suu tempo kerubahan wujud...	takut	takut	sedih	sedih
kata Henggah akhir Desember mau Paripolason Asat...	takut	takut	takut	marah
paksa DPR untuk menajektorian UU perampasan aset...	marah	takut	marah	marah
Maafkan, Zuhdi, Orlan Anangga betali, E...	marah	takut	marah	sedih
yg penting langsung dibales aja itu dihapus...	takut	takut	sedih	takut
BNATANG RAKUS BERDAS. PENYUSAH RAKYAT. SEMOO...	sedih	takut	takut	takut
Hanya akan hanya 200rb, itu sudah sangat benci...	marah	takut	marah	takut
Das terus #GASURADIKORPRI	marah	takut	takut	takut

Gambar 5. hasil pelabelan

Setiap pelabel menghasilkan lima kategori emosi yaitu senang, sedih, marah, takut, dan cinta. Setelah dilakukan pelabelan menggunakan lima pretrained, selanjutnya adalah dihitung menggunakan Krippendorff's Alpha untuk menilai kesepakatan antar pelabel. Hasil perhitungan Krippendorff's Alpha sebesar 0.73 yang menunjukkan hasil lima pelabel masih bisa diterima. Langkah akhir pelabelan yaitu menentukan label akhir dari setiap komentar yang telah melalui beberapa pelabelan, dengan pendekatan nilai modus atau label mayoritas. Nilai modus ini akan digunakan sebagai label akhir yang menjadi dasar untuk proses perancangan model yang dapat dilihat pada gambar 6 dibawah ini.



Gambar 6. distribusi label emosi.

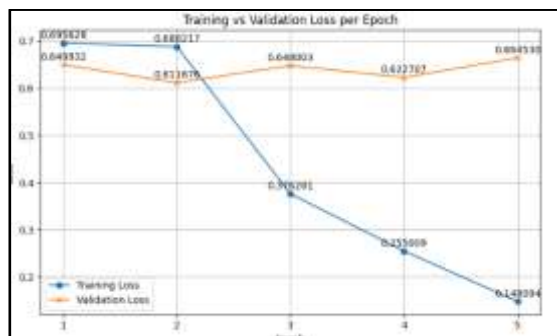
3.4 Pembagian data

Data dibagi menjadi dua bagian utama. Data latih, atau data pelatihan, digunakan untuk melakukan pelatihan model, dan data uji, atau data pengujian, digunakan untuk menilai kinerja model yang telah dilatih. Berdasarkan hasil pembagian data, diperoleh sebanyak 7.721 data (80,0%) sebagai data latih dan 1.931 data (20,0%) sebagai data uji.

3.5 Klasifikasi emosi

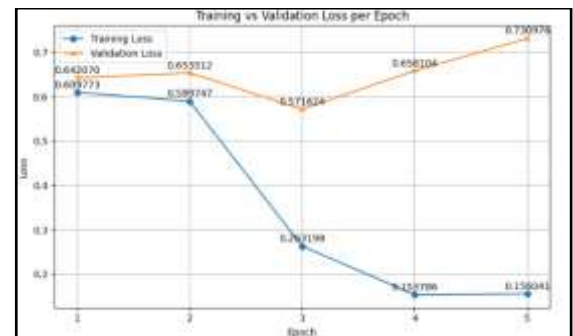
Proses pelatihan dilakukan menggunakan *Trainer* dengan parameter yang sama pada kedua model, yaitu 5 epoch, *learning rate* $2e-5$, dan *batch size* 16. Pengaturan ini memastikan bahwa perbandingan kinerja antar model dilakukan secara konsisten sehingga bersifat adil.

Berdasarkan hasil pelatihan, diperoleh nilai training loss dan validation loss IndoBERT-Base Uncased dan IndoBERTweet ditampilkan pada gambar 7 berikut.



Gambar 7. Train loss dan validation loss IndoBERT-Base Uncased.

Berdasarkan grafik training dan validation loss, model IndoBERT-Base Uncased menunjukkan penurunan training loss yang signifikan dari epoch 1 hingga 5, yang menandakan model mampu mempelajari data latih dengan baik. Namun, validation loss mencapai nilai terendah pada epoch ke-2 (0,611) dan kemudian berfluktuasi hingga meningkat kembali, yang mengindikasikan terjadinya overfitting setelah epoch tersebut. Dengan demikian, performa optimal model IndoBERT-Base Uncased berada pada epoch ke-2.



Gambar 8 train loss dan validation loss IndoBERTtweet

Berdasarkan gambar 8, model IndoBERTtweet menunjukkan penurunan training loss yang signifikan dari epoch 1 hingga 4, menandakan model mampu mempelajari data latih dengan baik. Namun, validation loss mencapai nilai terendah pada epoch ke-3 (0,571) dan kemudian meningkat hingga epoch ke-5, yang mengindikasikan terjadinya overfitting. Dengan demikian, performa optimal model IndoBERTtweet berada pada epoch ke-3. Selanjutnya hasil dari kinerja masing masing model ditunjukkan pada tabel 3.1 dibawah ini.

Tabel 1. HASIL KINERJA MODEL

Model	Accuracy	F1 Macro	F1 Weights
IndoBERT-Base Uncased	0.78	0.61	0.78
IndoBERTweet	0.80	0.65	0.80

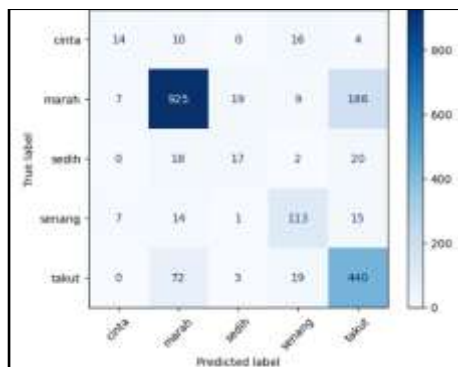
Berdasarkan hasil pada Tabel 1, IndoBERTtweet menunjukkan kinerja yang lebih baik dibandingkan IndoBERT-Base Uncased. Hal ini terlihat dari nilai accuracy yang lebih tinggi (0,80), yang menunjukkan kemampuan model dalam memprediksi data dengan lebih tepat. Selain itu, Nilai F1 Macro yang lebih tinggi (0,65)

mengindikasikan bahwa IndoBERTweet lebih mampu mengenali berbagai kategori emosi secara seimbang, termasuk pada kelas minoritas.

Pada F1 Weighted, IndoBERTweet juga unggul dengan nilai 0,80 dibandingkan 0,78, yang menunjukkan performa yang lebih stabil pada data yang tidak seimbang. Secara keseluruhan, IndoBERTweet lebih efektif dalam klasifikasi emosi.

3.6 Evaluasi

Visualisasi confusion matrix untuk kedua model ditampilkan pada Gambar 9 (IndoBERT-Base Uncased) dan Gambar 10 (IndoBERTweet). Visualisasi ini digunakan untuk melihat distribusi prediksi serta pola kesalahan pada masing-masing kelas emosi. Dengan demikian, confusion matrix menjadi alat yang penting untuk mengevaluasi performa model secara lebih mendalam dibandingkan hanya menggunakan metrik agregat seperti accuracy dan F1-score. Berikut adalah confusion Matrix IndoBERT-Base Uncased pada gambar 9.

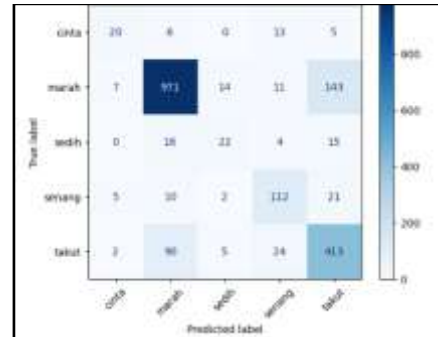


Gambar 9 Confusion Matrix IndoBERT-Base Uncased

Berdasarkan gambar 9, model IndoBERT-Base Uncased menunjukkan performa terbaik pada kelas marah dan takut, dengan jumlah prediksi benar masing-masing sebanyak 925 dan 440. Hal ini menunjukkan bahwa model mampu mengenali kedua emosi tersebut dengan cukup baik. Namun, pada kelas lain seperti cinta, sedih, dan senang, performa model masih relatif rendah.

Hal ini terlihat dari jumlah prediksi yang benar yang lebih sedikit serta adanya kesalahan klasifikasi ke kelas lain. Beberapa data dengan label marah masih salah diprediksi sebagai takut, dan data takut juga cukup sering diprediksi sebagai marah. Selain itu, kelas sedih memiliki jumlah prediksi benar yang

paling sedikit, yang menunjukkan bahwa model masih kesulitan dalam membedakan emosi tersebut.



Gambar 10 Confusion Matrix IndoBERTtweet

Berdasarkan gambar 10, model IndoBERTtweet menunjukkan performa yang baik terutama pada kelas marah (971) dan takut (413) dengan jumlah prediksi benar yang tinggi. Hal ini menandakan model cukup akurat dalam mengenali kedua emosi tersebut. Selain itu, kelas senang juga memiliki hasil yang cukup baik dengan 112 prediksi benar. Namun, pada kelas cinta dan sedih, performa masih lebih rendah dibandingkan kelas lainnya. Kesalahan klasifikasi masih terjadi, terutama antara kelas marah dan takut, yang sering saling tertukar.

Secara keseluruhan, IndoBERTtweet menunjukkan performa yang lebih merata dibandingkan model sebelumnya, meskipun masih terdapat kendala pada kelas dengan jumlah data yang lebih sedikit.

V. KESIMPULAN

Menurut penelitian yang telah dilakukan, dapat disimpulkan bahwa implementasi model berbasis transformer untuk klasifikasi emosi pada komentar YouTube menghasilkan performa yang cukup baik. Secara umum, kedua model yang digunakan menunjukkan kemampuan dalam mempelajari pola data, meskipun terdapat indikasi overfitting pada beberapa epoch. Hasil evaluasi menunjukkan bahwa model IndoBERTweet memiliki performa yang lebih unggul dibandingkan IndoBERT-Base Uncased, dengan nilai akurasi sebesar 0,80 dan F1-score yang lebih tinggi. Hal ini menunjukkan bahwa IndoBERTweet lebih efektif dalam mengklasifikasikan emosi, terutama pada data yang bersifat informal seperti komentar media sosial.

Selain itu, hasil analisis menunjukkan bahwa model cenderung lebih baik dalam mengenali emosi dengan jumlah data yang besar, seperti marah dan takut, dibandingkan emosi lainnya seperti sedih dan

cinta. Hal ini mengindikasikan bahwa distribusi data mempengaruhi performa model dalam klasifikasi emosi. Adapun saran untuk penelitian selanjutnya adalah dapat dilakukan pengembangan dengan menambahkan teknik penanganan data tidak seimbang (*imbalanced data*), seperti oversampling atau data augmentation. Penelitian selanjutnya juga dapat mengeksplorasi model lain atau melakukan optimasi parameter untuk memperoleh performa yang lebih optimal.

Physics: Conference Series, Institute of Physics Publishing, Mar. 2017. doi: 10.1088/1742-6596/801/1/012072.

- [15] A. Rahman, M. Zulkifle, M. K. Alam Ansari, S. -, and M. Aslam, "Literature Review: A Significant Feature of Writing a Research Report," *World J. Res. Rev.*, vol. 11, no. 1, Jul. 2020, doi: 10.31871/wjrr.11.1.13.
- [16] V. R. Joseph, "Optimal ratio for data splitting," *Stat. Anal. Data Min.*, vol. 15, no. 4, pp. 531–538, Aug. 2022, doi: 10.1002/sam.11583.

REFERENCES

- [1] Ayuni, M. S., & Sari, D. (n.d.). PERAN MEDIA SOSIAL DALAM MENINGKATKAN PARTISIPASI POLITIK GENERASI MILENIAL DI INDONESIA. <https://jurnalsentral.com/index.php/jdss>
- [2] Peng, S., et al. (2022). A survey on deep learning for textual emotion analysis in social networks. *Digital Communications and Networks*. <https://doi.org/10.1016/j.dcan.2021.10.003>
- [3] Apriansyah, F. M., Ramadhan, T. I., Hidayat, C. R., & Wijaya, A. K. (2025). Perbandingan IndoBERT dan IndoRoBERTa untuk analisis sentimen pada film dokumenter Dirty Vote. *Jurnal Informatika: Jurnal Pengembangan IT*, 10(3), 593–605. <https://doi.org/10.30591/jpit.v10i3.8607>
- [4] Christian, W., Adamlu, D., Yu, A., & Suhartono, D. (2025). Leveraging IndoBERT and DistilBERT for Indonesian emotion classification in e-commerce reviews. <http://arxiv.org/abs/2509.14611>
- [5] Nabiihah, G. Z., Alam, I. N., Purwanto, E. S., & Hidayat, M. F. (2024). Indonesian multilabel classification using IndoBERT embedding and MBERT classification. *International Journal of Electrical and Computer Engineering*, 14(1), 1071–1078. <https://doi.org/10.11591/ijece.v14i1.pp1071-1078>
- [6] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," Online. [Online]. Available: <https://huggingface.co/>
- [7] F. Indriani, R. A. Nugroho, M. R. Faisal, and D. Kartini, "Comparative Evaluation of IndoBERT, IndoBERTweet, and mBERT for Multilabel Student Feedback Classification," *Jurnal RESTI*, vol. 8, no. 6, pp. 748–757, Dec. 2024, doi: 10.29207/resti.v8i6.6100.
- [8] Ekman, U., et al. (2020). Evaluation of a novel psychological intervention tailored for patients with early cognitive impairment (PIPCL): Study protocol of a randomized controlled trial. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.600841>
- [9] Text mining: predictive methods for analyzing unstructured information. *Scholars Portal*, 2019.
- [10] H. Xie, Z. Tan, and X. Lv, "Application of Artificial Intelligence in Financial Risk Management in a Company," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2024, pp. 349–354. doi: 10.1145/1122445.1122456.
- [11] F. Koto, J. H. Lau, and T. Baldwin, "IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," Sep. 2021, [Online]. Available: <http://arxiv.org/abs/2109.04607>
- [12] Zeng, G. (2025). Invariance properties and evaluation metrics derived from the confusion matrix in multiclass classification. *Mathematics*, 13(16). <https://doi.org/10.3390/math13162609>
- [13] A. Rahman, M. Zulkifle, M. K. Alam Ansari, S. -, and M. Aslam, "Literature Review: A Significant Feature of Writing a Research Report," *World J. Res. Rev.*, vol. 11, no. 1, Jul. 2020, doi: 10.31871/wjrr.11.1.13.
- [14] A. F. Hidayatullah and M. R. Ma'Arif, "Pre-processing Tasks in Indonesian Twitter Messages," in *Journal of*