

# Klasifikasi Sentimen Multimodal Pada Media Sosial X Terkait Isu Kesehatan Mental Dengan Ekspansi Fitur FastText dan Model CNN-GRU

1st I Gusti Bagus Bagaskara Kerta Yasa  
Informatika  
Universitas Telkom  
Bandung, Indonesia  
bagasskara@student.telkomuniversity.ac.id

2nd Erwin Budi Setiawan  
Informatika  
Universitas Telkom  
Bandung, Indonesia  
erwinbudisetiawan@telkomuniversity.ac.id

**Abstrak**— Media sosial kini berfungsi sebagai platform utama bagi orang untuk membagikan pengalaman dan pandangan, termasuk yang berkaitan dengan kesehatan mental. Namun, analisis data yang ada menghadapi beberapa tantangan, seperti perbedaan jenis seperti teks dan gambar dan pola ekspresi yang rumit. Penelitian ini bertujuan untuk merancang model analisis sentimen multimodal yang dapat mendeteksi masalah kesehatan mental di media sosial X dengan menggabungkan metode *Convolutional Neural Network* (CNN), *Gated Recurrent Units* (GRU), dan ekspansi fitur FastText. Metodologi yang diusulkan mencakup pengumpulan data secara real-time dari media sosial X, meliputi teks dan gambar. Data teks diolah dengan teknik praproses standar dan representasi fitur FastText, sedangkan data visual diambil menggunakan VGG-16 untuk mengenali pola visual yang relevan. 24.742 pasangan tweet gambar dikumpulkan dari platform Twitter dan dianalisis melalui sistem pemungutan suara mayoritas. Untuk membangun korpus kemiripan FastText, 63.512 data dari portal berita digital CNN dan Twitter digabungkan menggunakan modalitas dilakukan melalui lapisan integrasi untuk menghasilkan klasifikasi sentimen akhir (positif dan negatif). Hasil evaluasi pada dataset uji menunjukkan bahwa metode ini mampu meningkatkan akurasi deteksi sentimen hingga 0,12% dibandingkan metode yang hanya berbasis teks. Secara keseluruhan, akurasi yang diperoleh mencapai 87,89%. Dengan capaian ini, diharapkan penelitian ini dapat menjadi referensi dalam pemantauan isu kesehatan mental di media sosial X secara lebih efektif.

**Keywords**—Analisis Sentimen Multimodal, VGG-16, Hybrid CNN-GRU, TF-IDF, FastText

## LATAR BELAKANG

Salah satu aspek kesehatan yang penting adalah kondisi di mana seseorang tidak menunjukkan tanda-tanda masalah psikologis apa pun. Individunya bisa menjalani kehidupan sehari-hari dengan baik meski menghadapi berbagai tantangan yang datang secara alami. Mereka memiliki kemampuan untuk mengelola tekanan dan stres dalam kehidupan mereka dengan baik. Kesehatan mental sama pentingnya dengan kesehatan fisik karena keduanya saling terkait dan bisa saling memengaruhi satu sama lain secara signifikan dalam kehidupan seseorang. Stres yang berkepanjangan harus diatasi dengan baik untuk mencegah memperburuk kondisi kesehatan mental dan mencegah timbulnya masalah yang lebih serius pada waktu yang akan datang [1].

Penggunaan media sosial di kalangan remaja saat ini marajalela dengan cara-cara yang tidak baik sehingga mereka dapat mengalami masalah mental yang mengarah ke dampak

yang parah sehingga perlu menjaga kesehatan mental mereka. Gangguan yang paling umum terjadi di kalangan remaja antara lain adalah kecemasan, gangguan emosi, perilaku negatif, menyakiti diri sendiri, dan bunuh diri. Jika seorang remaja menunjukkan gejala-gejala yang berkaitan dengan gangguan mental ini, ada kebutuhan khusus untuk memastikan bahwa seseorang harus mendapatkan bantuan dari orang yang profesional dalam bidang ini. Kesehatan mental yang optimal tercapai ketika seseorang berada dalam kondisi damai dan tenteram, sehingga mampu menikmati aktivitas sehari-hari serta menjalin hubungan yang harmonis dengan orang di sekitarnya. Individu dengan kondisi mental yang sehat dapat mengembangkan potensi diri secara maksimal, menghadapi berbagai tantangan hidup, dan membangun relasi positif. Sebaliknya, gangguan kesehatan mental dapat memicu perubahan suasana hati, menurunnya kemampuan berpikir, serta ketidakstabilan emosi yang berpotensi memengaruhi perilaku secara negatif. Salah satu akibat negatifnya adalah berkembangnya gangguan mental yang dapat mengganggu kehidupan sehari-hari, dengan penyebab utama tidak hanya menghalangi interaksi sosial atau hubungan dengan orang lain, tetapi juga menurunkan prestasi di sekolah dan produktivitas dalam pekerjaan. Oleh karena itu, kita harus menerapkan gaya hidup sehat yang dapat kita manfaatkan [2].

Twitter yang sekarang lebih dikenal sebagai X adalah sebuah platform media sosial yang sangat populer saat ini. Penelitian sebelumnya menunjukkan bahwa banyak remaja menghabiskan lebih banyak waktu di media sosial X setiap harinya. Mereka bahkan akan l mencuri waktu untuk membuka aplikasi X walaupun sedang menyelesaikan tugas atau melakukan aktivitas lainnya [3]. Beberapa remaja kini lebih suka mengungkapkan emosinya melalui cuitan di aplikasi. Mereka juga cenderung menggunakan akun pribadi atau akun terkunci karena dianggap aman karena hanya pengikutnya yang bisa melihatnya. Sehingga tidak mudah bagi pihak ketiga untuk mengakses informasi tweet dan pemilik akun. Dengan aplikasi X, mereka bisa mengungkapkan keluh kesahnya atas apa yang dirasakannya tanpa harus khawatir dengan tanggapan orang lain terhadap postingannya. Hal ini juga sangat berkontribusi penting dengan kesehatan mental mereka.

Pada penelitian ini mereka menggunakan CNN untuk menganalisis sentimen untuk mengukur kesehatan mental penggunaan X dengan menerapkan algoritma deep learning yaitu CNN dan hasil menunjukkan bahwa mereka mendapatkan

akurasi sebesar 80,90% yang dimana akurasi ini termasuk cukup baik untuk menunjukkan bahwa sentimen yang diekspresikan dalam tweet dapat digunakan untuk menilai kondisi mental pengguna [4]. Dalam penelitian yang dilakukan oleh Lickha dan Reisa, dilakukan perbandingan tiga varian algoritma Naïve Bayes yaitu Multinomial, Gaussian, dan Bernoulli untuk analisis sentimen terkait gangguan depresi di platform media sosial X. Hasil studi tersebut menunjukkan bahwa Gaussian Naïve Bayes (GNB) memperoleh tingkat akurasi sebesar 88,38% [5]. Sementara itu, penelitian lain bertujuan mengidentifikasi tingkat depresi pengguna X dengan menerapkan metode klasifikasi KNearest Neighbor (KNN). Berdasarkan hasil pengujian, metode KNN pada penelitian tersebut menghasilkan akurasi sebesar 78,18%. Nilai akurasi ini diperoleh dari sistem yang dibangun melalui tahapan *pre-processing*, klasifikasi menggunakan KNN, serta evaluasi model dengan confusion matrix [6]. Penelitian ini menggabungkan BERT untuk representasi text dan ResNet50 untuk ekstraksi fitur visual dengan lima metode fusi, termasuk berbasis attention. Pada dataset MAVAS-single, model OTEModel mencapai akurasi 74,5% dan F1-score 73%, melampaui model unimodal dan fusi sederhana, membuktikan bahwa strategi fusi tepat dapat memanfaatkan informasi komplementer text dan gambar untuk meningkatkan analisis sentimen [27].

Dalam tugas akhir ini, penulis menggunakan teknik menganalisis cuitan sosial media X untuk memahami situasi di dunia nyata menggunakan model CNN untuk analisis sentimen. Saya memilih jaringan konvolusional sebagai model analisis sentimen karena dalam analisis dan klasifikasi, CNN dapat mengekstrak wilayah fitur dari semua 2 informasi. Dengan melakukan konvolusi, data dapat dikumpulkan sebagai fitur, dan hubungan antara fitur-fitur tersebut dapat dihitung [7]. Penelitian ini bertujuan untuk mengevaluasi validitas metode yang diajukan penulis.

## KAJIAN PUSTAKA

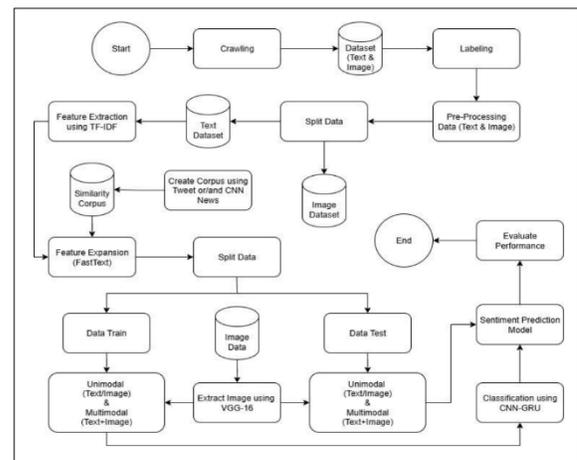
Dalam kajian Pustaka ini, terdapat beberapa paper atau jurnal yang meneliti masalah isu kesehatan mental di sosial media. Kesehatan mental adalah konsep yang luas, kompleks, dan tidak berwujud, sehingga sulit untuk diukur [8]. Platform sosial seperti X [9], Facebook, dan Instagram menyediakan banyak data yang dapat dieksplorasi oleh para peneliti untuk memahami pandangan masyarakat dan fenomena sosial. Bahasa yang dipakai pengguna di platform-platform ini dapat mencerminkan emosi dan pandangan mereka mengenai berbagai isu, termasuk politik, masalah sosial, dan kesehatan masyarakat [10]. Di samping itu, karakteristik data media sosial yang cepat membantu dalam aspek kesadaran situasi, pengawasan, dan menemukan pola perilaku serta tren berita [11]. Analisis sentimen memiliki potensi untuk secara signifikan meningkatkan kesehatan mental dan kesejahteraan dengan menyediakan alat yang ampuh untuk memantau kondisi emosional individu dan komunitas [12].

Pada penelitian pertama mengidentifikasi bahwa hasil yang diperoleh menunjukkan algoritma deep learning LSTM dan CNN memiliki kinerja yang lebih baik daripada algoritma machine learning dengan akurasi yang lebih tinggi untuk analisis sentimen [13]. Selanjutnya pada penelitian ini mereka menggunakan model arsitektur LSTM-CNN untuk mendeteksi depresi dan mampu mencapai akurasi tertinggi sebesar 97% yang mengungguli model lain seperti regresi logistik, Naïve Bayes, Random Forest, dan Decision Tree [14].

Pada penelitian ini hasil eksperimen menunjukkan bahwa fitur pemilihan melalui AOA setelah ekstraksi fitur melalui CNN berkontribusi secara signifikan terhadap kinerja klasifikasi. Sedangkan akurasi klasifikasi CNN adalah 89,717% [15]. Hasilnya menunjukkan bahwa analisis kontekstual FastText meningkatkan akurasi klasifikasi di mana mayoritas postingan Reddit. Klasifikasi biner dari data ini melaporkan 0,78 AUC dan 0,71 bobot Skor F1 [16]. Penelitian ini menunjukkan bahwa penggunaan model hybrid CNN-GRU mampu meningkatkan performa analisis sentimen dibandingkan dengan model tunggal CNN maupun GRU. Dengan memanfaatkan CNN untuk mengekstraksi fitur spasial dari text dan GRU untuk menangkap dependensi sekuensial, model *hybrid* berhasil mencapai akurasi 4 82,67%, precision 0,84, recall 0,80, dan F1-score 0,83 pada data X berbahasa Inggris dan Thailand. Hasil ini menegaskan bahwa pendekatan kombinasi CNN-GRU lebih efektif dalam menangani kompleksitas data text pada media sosial dibandingkan model tunggal [22].

Secara keseluruhan, penelitian-penelitian tersebut menunjukkan bahwa pendekatan yang melibatkan berbagai metode, seperti pemanfaatan CNN dan penerapan FastText, memiliki potensi signifikan dalam analisis sentimen terkait kesehatan mental. Metode-metode ini memungkinkan pemeriksaan yang lebih mendalam terhadap pandangan publik di media sosial, khususnya X, yang sering kali menjadi platform utama untuk perbincangan mengenai kesehatan mental.

## METODOLOGI



Gambar. 1. Flowchart Analisis Sentimen Multimodal

## A. Crawling Data

Solusi pengelolaan data secara menyeluruh mengumpulkan sejumlah besar tweet yang berupa text maupun gambar untuk dilakukan analisis sentimen terhadap percakapan di X terkait isu kesehatan mental. Dengan menggunakan kata kunci, rentang waktu tertentu, atau akun pengguna, alat seperti Tweet-harvest dapat mengekstrak tweet yang relevan, kemudian menyimpannya dalam format CSV yang lengkap, termasuk stempel waktu, tautan gambar, *text* tweet, serta metadata yang mendukung analisis berbasis multimodal. Sebaran data tersebut bisa di lihat pada Tabel 1.

TABEL 1 JUMLAH DATASET DENGAN KATA KUNCI

Keyword	Quantity
broken	2,078
burnout	3,000
frustration	2,500
unloved	3,000
anxiety	2,800
selflove	3,570
mentalhealthmatters	2,500
wellbeing	2,794
toxic	2,500
Total	24,742

## B. Labeling Data

Dalam analisis sentimen multimodal, khususnya pada pembahasan kesehatan mental di sosial media X, proses pelabelan data memegang peran krusial. Setelah data dikumpulkan, setiap tweet yang memuat text sekaligus gambar dikategorikan ke dalam sentimen positif dan negatif. Penandaan dilakukan pada masing-masing jenis data, di mana ekspresi emosinya bisa saja sejalan ataupun bertolak belakang. Untuk menjaga keseragaman, para annotator yang berjumlah 5 orang bekerja secara kolaboratif dan menyamakan persepsi melalui diskusi. Langkah yang sistematis ini menghasilkan himpunan data yang dapat dipercaya, 13 yang menjadi fondasi penting dalam pelatihan serta pengujian model hibrida CNN-GRU agar mampu menafsirkan ragam ekspresi emosi secara lebih akurat. Penjelasan label tersebut dijabarkan pada Tabel 2.

TABEL 2 DISTRIBUSI LABEL SENTIMEN

Category	Label	Quantity
Text	Positive	12,429
	Negative	12,313
Image	Positive	12,429
	Negative	12,313
Total		24,742

## C. Pre-Processing Data

Pre-processing mencakup penghilangan komponen yang tidak penting seperti URL, emoji, dan simbol, pengubahan kata

menjadi huruf kecil, dan penghapusan kata-kata yang kurang berperan (stopwords). Dalam beberapa situasi, metode lemmatization diterapkan untuk mengembalikan kata ke bentuk dasarnya. Untuk gambar, *pre-processing* melibatkan penyesuaian ukuran agar sesuai dengan masukan model CNN + GRU, seperti mengubah gambar ke format grayscale atau RGB yang konsisten. Tujuan dari proses ini adalah untuk menyederhanakan data agar lebih relevan dan siap untuk dianalisis. Stopwords adalah kata-kata umum yang biasanya diabaikan dalam analisis text (misalnya, "dan", "atau", "ini") dan ada library yang digunakan yaitu scikit-learn. Stemming adalah menghilangkan akhiran kata untuk mendapatkan bentuk dasar kasar tanpa mempertimbangkan makna atau aturan gramatikal. Ada beberapa library yang digunakan adalah NLTK (*Natural Language Toolkit*). Pre-processing data gambar dengan cara mengubah ukuran menjadi 128 x 128 piksel sehingga mendapatkan hasil resolusi tetap dan seragam. Mengubah format gambar menjadi grayscale dan mendapatkan hasil gambar menjadi hitam putih. Normalisasi nilai piksel menjadi antara 0 dan 1 sehingga lebih stabil digunakan untuk model CNN + GRU.

## E. Ekstraksi Fitur TF-IDF

Representasi text diubah ke dalam bentuk angka dengan menggunakan metode seperti TF-IDF, yang menilai relevansi kata berdasarkan frekuensinya dalam dokumen tertentu. Metode ini menghasilkan representasi vektor kata yang mencerminkan hubungan makna di antara kata-kata, sehingga membantu model untuk mengerti arti dan konteks dari *text*.

## F. Ekspansi Fitur FastText

Setelah fitur text didapat, perluasan fitur dilakukan dengan menggunakan FastText. Tujuan dari proses ini adalah untuk memperluas representasi semantik dari kata dengan menciptakan kata-kata yang memiliki makna yang terkait dengan kata di korpus awal. Misalnya, kata "depresi" bisa diperluas dengan sinonim seperti "stres" atau istilah lain yang relevan, tergantung pada korpus tambahan yang diterapkan. Ini membantu menyelesaikan masalah ketidakcocokan kosakata, khususnya ketika ada variasi dalam cara berbahasa yang digunakan oleh pengguna pada media sosial X. Penelitian ini menggunakan 3 korpus dari berbagai sumber seperti yang ada di Tabel 3.

Tabel 3. Korpus yang digunakan FastText

Corpus	CNN News	Tweets	Tweets + CNN News
Total	63,512	24,742	88,254

## G. Ekstraksi Gambar

Untuk aspek visual, fitur-fitur diambil menggunakan model VGG-16 yang dibuat untuk menganalisis gambar. Model ini menemukan pola-pola visual yang penting, seperti ekspresi wajah atau latar belakang yang bisa menunjukkan masalah kesehatan mental [30]. Setelah kedua jenis fitur, yaitu text dan visual, berhasil diambil, langkah selanjutnya adalah

menyatukan keduanya menjadi satu kumpulan fitur multimodal. Penyatuan ini menjamin bahwa analisis melihat konteks verbal dan visual secara bersamaan.

## H. Penerapan Sentimen Hybrid CNN-GRU

Data multimodal yang telah diolah selanjutnya digunakan dalam model CNN-GRU untuk pengklasifikasian sentimen dan identifikasi masalah terkait kesehatan mental. CNN-GRU dibuat untuk mengenali pola dalam data dan membuat prediksi berdasarkan pola-pola tersebut. Model ini dilatih dengan menggunakan dataset berlabel agar dapat mengenali sentimen seperti positif dan negatif. Serta mendeteksi tanda-tanda kesehatan mental yang relevan.

## I. Evaluasi Model

Setelah model membuat prediksi, kinerjanya dinilai dengan menggunakan metrik seperti akurasi, presisi, *recall*, dan F1-score. Jika evaluasi menunjukkan hasil yang baik, model bisa digunakan untuk menganalisis data baru atau dijadikan sebagai bagian dari penelitian. Namun, jika kinerjanya masih kurang, langkah-langkah perbaikan dilakukan, seperti meningkatkan proses *pre-processing* dan ekstraksi fitur. Proses ini dilakukan berulang kali hingga model mendapatkan tingkat akurasi dan keandalan yang diharapkan.

## EVALUASI DAN DISKUSI

Empat model algoritma klasifikasi yang akan diuji dalam penelitian ini adalah CNN, GRU, CNN-GRU, dan GRU-CNN. Hasil pengujian akan dijelaskan dan diuraikan berdasarkan beberapa skenario berikut.

### A. Skenario 1

Pada skenario pertama, eksperimen bertujuan untuk mendapatkan akurasi dari model dasar menggunakan metode CNN dan GRU tanpa menggunakan ekspansi fitur FastText

TABEL 3 HASIL SPLIT RATIO

Split Ratio	Accuracy (%)	
	CNN	GRU
90:10	85.79	86.79
80:20	<b>86.29</b>	86.83
70:30	86.01	<b>86.89</b>

TABEL 4 HASIL MAX FEATURE CNN

Data Ratio	Max Feature	Accuracy %
		CNN
80:20	5000	86.24
	<b>10000</b>	<b>86.44</b>
	15000	85.39

TABEL 5 HASIL MAX FEATURE GRU

Data Ratio	Max Feature	Accuracy %
		GRU
70:30	<b>5000</b>	<b>86.89</b>
	10000	86.88
	15000	86.87

Berdasarkan hasil pengujian pada Tabel 5, model GRU menunjukkan performa yang sedikit lebih unggul dibandingkan CNN pada seluruh skenario, dengan akurasi tertinggi 86,89% pada split ratio 70:30, sedangkan CNN mencapai akurasi tertinggi 86,29% pada split ratio 80:20. Keunggulan GRU ini dapat dijelaskan oleh kemampuannya dalam memproses data sekuensial melalui mekanisme gating yang efektif mempertahankan konteks dan mengabaikan informasi yang kurang relevan. Pada analisis sentimen, urutan kata dan hubungan antar kata memiliki pengaruh besar terhadap pemahaman makna, sehingga GRU mampu memanfaatkan representasi sekuensial secara lebih optimal. Hasil pada pengujian max feature juga memperkuat temuan ini, di mana GRU memperoleh akurasi tertinggi pada 10.000 fitur yang dijelaskan pada Tabel 10, sementara pada Tabel 9 menunjukkan CNN optimal pada 5.000 fitur. Hal ini menunjukkan bahwa GRU dapat memanfaatkan kosakata yang lebih besar untuk memperkaya konteks, sedangkan CNN lebih efektif dengan jumlah fitur yang lebih terbatas. Konsistensi keunggulan GRU pada seluruh pengujian mengindikasikan bahwa pendekatan pemodelan sekuensial memberikan fleksibilitas dan kemampuan generalisasi yang lebih baik dibandingkan pendekatan ekstraksi pola lokal yang digunakan CNN.

### B. Skenario 2

Pada skenario kedua, tujuan eksperimen adalah mencapai akurasi tertinggi dengan menggunakan metode CNN dan GRU yang dilengkapi dengan ekstraksi fitur TF-IDF untuk mencari n-gram terbaik.

TABEL 6 HASIL SPLIT RATIO

N-Gram	Accuracy (%)	
	CNN	GRU
Unigram (1, 1)	86.29	<b>86.89</b>
Bigram (2, 2)	75.39	86.62
Trigram (3, 3)	63.37	86.68
Unigram + Bigram (1, 2)	86.26	86.70
Unigram + Bigram + Trigram (1, 1)	86.39	86.77

Berdasarkan hasil pada Tabel 6, penggunaan unigram menghasilkan performa terbaik bagi GRU model, dengan akurasi tertinggi 86,89% pada GRU dan 86,39% pada CNN. Performa GRU cenderung meningkat lebih signifikan dibandingkan CNN ketika menggunakan kombinasi berbagai

N-gram. Hal ini mengindikasikan bahwa GRU mampu memanfaatkan informasi tambahan dari rentang konteks kata yang lebih luas secara lebih efektif, berkat kemampuannya dalam memodelkan dependensi sekuensial. Sebaliknya, CNN yang lebih berfokus pada pola lokal tidak memperoleh peningkatan sebesar GRU ketika konteks diperluas melalui kombinasi N-gram.

### C. Skenario 3

Tujuan pada skenario ketiga adalah mendapatkan akurasi model dasar dengan menggunakan fitur similarity dari berbagai sumber (Tweet, CNN News, Tweet + CNN News) untuk metode CNN dan GRU.

TABEL 7 HASIL FASTTEXT CNN DAN GRU

Model	Rank	Accuracy (%)			
		Baseline	Tweet	News	Tweet+News
CNN	Top 1	86.39	87.21	<b>87.04</b>	87.57
	Top 5		87.69	86.83	<b>87.71</b>
	Top 10		<b>87.75</b>	86.66	87.70
	Top 15		87.43	86.62	87.65
	Top 20		87.45	-	-
	Top 25		-	-	-
GRU	Top 1	87.79	87.52	<b>87.42</b>	87.42
	Top 5		87.24	87.25	<b>87.72</b>
	Top 10		<b>87.53</b>	87.36	87.49
	Top 15		87.49	87.19	87.71
	Top 20		87.15	-	-
	Top 25		-	-	-

Berdasarkan hasil pada Tabel 7, CNN mencapai akurasi tertinggi sebesar 87,75% pada *corpus Tweet* di rank Top 10, sedangkan GRU memperoleh akurasi tertinggi 87,72% pada *corpus gabungan Tweet dan News* di rank Top 5. Perbedaan ini menunjukkan bahwa CNN lebih sensitif terhadap peningkatan akurasi ketika menggunakan representasi kata FastText, khususnya pada data yang bersifat tunggal seperti *corpus Tweet*. Sementara itu, GRU cenderung menghasilkan performa yang stabil pada berbagai konfigurasi, namun sedikit lebih rendah dibanding CNN dalam kondisi tertentu. Sensitivitas CNN ini dapat dikaitkan dengan kemampuannya mengekstraksi pola lokal yang lebih terfokus pada konteks spesifik *corpus*, sedangkan GRU lebih mengandalkan pemodelan sekuensial yang menjaga konsistensi performa meskipun variasi data dan peringkat fitur berbeda.

### D. Skenario 4

Tujuan pada skenario keempat adalah mendapatkan akurasi model dasar dengan menggunakan fitur *similarity* dari berbagai

sumber (Tweet, CNN News, Tweet + CNN News) untuk metode Hybrid CNN + GRU dan Hybrid GRU + CNN.

TABEL 8 HASIL FASTTEXT HYBRID CNN-GRU DAN GRU-CNN

Model	Rank	Accuracy (%)		
		Tweet	News	Tweet+News
CNN-GRU	Top 1	87.42	87.38	87.53
	Top 5	<b>87.77</b>	<b>87.50</b>	<b>87.69</b>
	Top 10	87.23	87.11	87.61
	Top 15	87.15	87.23	87.34
	Top 20	-	-	-
	Top 25	-	-	-
GRU-CNN	Top 1	86.88	<b>87.46</b>	<b>87.30</b>
	Top 5	86.84	87.00	87.00
	Top 10	87.19	87.34	87.00
	Top 15	<b>87.30</b>	86.96	86.77
	Top 20	87.27	-	-
	Top 25	87.11	-	-

Berdasarkan hasil pada Tabel 8, metode CNN-GRU memperoleh akurasi tertinggi pada *corpus tweet* dengan nilai 87,77% pada peringkat TOP 5, sedangkan metode GRU-CNN mencapai akurasi tertinggi 87,46% pada *corpus news* di peringkat TOP 1. Perbedaan ini menunjukkan bahwa penggunaan *corpus tweet* cenderung memberikan performa yang lebih baik dibandingkan sumber data lain. Hal ini kemungkinan disebabkan oleh karakteristik *tweet* yang umumnya padat, langsung menyampaikan opini, dan memiliki variasi bahasa yang lebih ekspresif, sehingga lebih mudah ditangkap pola sentimennya oleh model. Sementara itu, *corpus news* biasanya bersifat lebih formal dan netral, sehingga informasi emosional yang relevan untuk analisis sentimen lebih sedikit dan memerlukan pemrosesan yang lebih kompleks untuk diidentifikasi.

### E. Skenario 5

Pada skenario terakhir, hasil pengujian menunjukkan pada Tabel 9 bahwa integrasi modalitas *text* dan gambar mampu memberikan peningkatan kinerja klasifikasi sentimen dibandingkan penggunaan *text* saja. Meskipun selisih akurasi yang diperoleh relatif kecil, yaitu 0,12% lebih tinggi dibanding model berbasis *text* tunggal, capaian akurasi 87,89% tersebut merupakan hasil terbaik secara keseluruhan. Hal ini mengindikasikan bahwa informasi visual dari gambar dapat memberikan kontribusi tambahan yang melengkapi representasi *text*, sehingga model memiliki pemahaman yang lebih kaya terhadap konteks data. Temuan ini selaras dengan

prinsip analisis sentimen multimodal, di mana kombinasi berbagai sumber informasi dapat memperkuat prediksi, terutama pada kasus ketika salah satu modalitas memiliki keterbatasan dalam menangkap nuansa emosional atau konteks tertentu.

TABEL 9 EVALUASI MODALITAS TUNGGAL DAN MULT

Feature	Accuracy (%)
Text-based	87,77
Image-based	70,22
<b>Multimodal (Text+Image)</b>	<b>87,89</b>

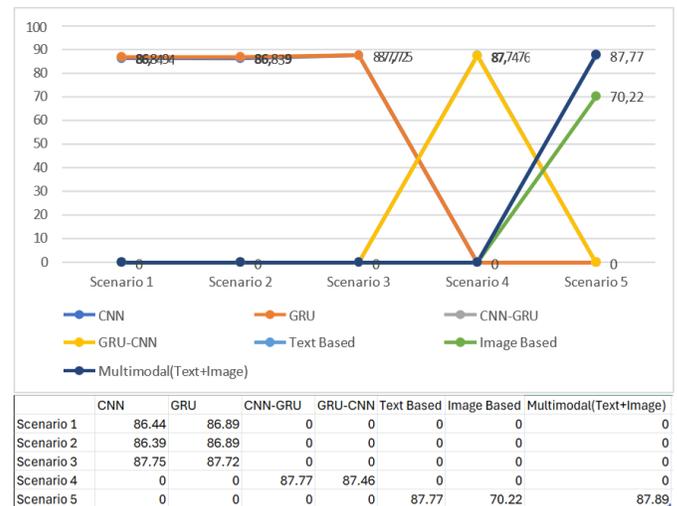
## F. DISKUSI

Dalam penelitian ini dilakukan beberapa skenario pengujian untuk menemukan model yang paling optimal. Hasil akurasi model berbasis text dan model multimodal pada lima skenario tersebut ditunjukkan pada Gambar 2. Skenario pertama dijadikan sebagai baseline, dengan akurasi awal 86,44% untuk CNN dan 86,89% untuk GRU. Pada skenario kedua, akurasi CNN sedikit menurun menjadi 86,39%, sementara GRU tetap di angka 86,89%. Selanjutnya, pada skenario ketiga akurasi CNN naik sebesar 1,36%, sedangkan GRU meningkat juga sebesar 0,83%. Pada skenario keempat, model hibrida menunjukkan hasil yang lebih baik dibanding model tunggal, yaitu 87,77% untuk CNN-GRU dan 87,46% untuk GRU-CNN. Terakhir, di skenario kelima dilakukan pengujian gabungan text dan gambar untuk melihat efektivitas model multimodal. Hasilnya, akurasi fusi multimodal mencapai 87,89%.

Untuk mengetahui apakah peningkatan tersebut signifikan, dilakukan uji statistik menggunakan Z-Test terhadap hasil akurasi model terbaik yang dijelaskan pada Tabel 10. Dari uji tersebut diketahui bahwa peningkatan antara skenario 1 ke 2, 3 ke 4, serta 4 ke 5 tidak signifikan. Sebaliknya, perbedaan antara skenario 2 ke 3 dan 1 ke 5 sangat signifikan ( $p < 0,01$ ). Hasil ini membuktikan bahwa penggabungan text dan gambar pada model multimodal menghasilkan peningkatan kinerja yang paling berarti.

TABEL 10 UJI SIGNIFIKAN

Transisi Skenario	Z-Value	P-Value	Signifikan?
S1 → S2	-1.0	0.32	Tidak
S2 → S3	4.13	3.60	Ya
S3 → S4	0.86	0.39	Tidak
S4 → S5	1.00	0.32	Tidak
S1 → S5	15.67	0.0	Ya



Gambar. 2. Performa Akurasi Varian Pada Seluruh Skenario

## KESIMPULAN

Pada eksperimen analisis sentimen multimodal ini menggunakan 24.742 tweet gambar berbahasa Inggris yang telah diproses sebelumnya dari X, dengan label positif dan negatif. Korpus kemiripan FastText dibuat dengan menggunakan 63.512 item dari situs berita digital CNN (Cable News Network) dan dari data Tweet. Model CNN-GRU dengan mekanisme attention digunakan untuk menganalisis semua kasus, menggunakan rasio pembagian 90:10 untuk penilaian. Pengaturan ekstraksi fitur *text* terbaik adalah menggabungkan TF-IDF dengan Unigram, maksimum 15.000 fitur, dan diperluas dengan 5 kata pembanding dari korpus FastText. Ekstraksi fitur visual dengan VGG-16. Teknik multimodal menggabungkan *text* dan visual, menghasilkan akurasi tertinggi secara keseluruhan 87,89%. Hasilnya menunjukkan peningkatan sebesar 0,12% dibandingkan model berbasis text saja. Kekuatan penelitian ini adalah dalam pendekatannya yang lengkap terhadap ekstraksi fitur, yang memanfaatkan manfaat dari setiap modalitas dan tingkat fusi untuk menafsirkan dan mengidentifikasi emosi dalam wacana kesehatan mental dengan lebih baik.

## DAFTAR PUSTAKA

- [1] O. : Adisty, W. Putri, B. Wibhawa, and A. S. Gutama, "41 KESEHATAN MENTAL MASYARAKAT INDONESIA (PENGETAHUAN, DAN KETERBUKAAN MASYARAKAT TERHADAP GANGGUAN KESEHATAN MENTAL)".
- [2] S. Zahira Ardhanita, F. Catur, and P. Lestari, "Triwikrama: Jurnal Ilmu Sosial PENGARUH MEDIA SOSIAL TERHADAP KESEHATAN MENTAL REMAJA," vol. 4, no. 3, pp. 2024–91, 2023.
- [3] Kukuh Wijayanti and Qoniah Nur Wijayani, "PERANAN APLIKASI X ATAU X DALAM INTERAKSI KOMUNIKASI GUNA MEMBANTU PENYEIMBANGAN KESEHATAN MENTAL PADA REMAJA SAAT INI," JOURNAL SAINS STUDENT RESEARCH, vol. 2, no. 1, pp. 07–15, Dec. 2023, doi: 10.61722/jssr.v2i1.469.

- [4] A. Rizki and Y. Sibaroni, "ANALISIS SENTIMEN UNTUK PENGUKURAN TINGKAT DEPRESI PENGGUNA X MENGGUNAKAN DEEP LEARNING."
- [5] N. L. Lavenia and R. Permatasari, "Sentiment Analysis on X Social Media Regarding Depression Disorder Using the Naive Bayes Method," *CoreID Journal*, vol. 1, no. 2, pp. 66–74, Jul. 2023, doi: 10.60005/coreid.v1i2.14.
- [6] A. Primadhani Tirtopangarsa and W. Maharani, "Sentiment Analysis of Depression Detection on X Social Media Users Using the K-Nearest Neighbor Method Analisis Sentimen Deteksi Depresi pada Pengguna Media Sosial X dengan Menggunakan Metode K-Nearest Neighbor," pp. 13–2021.
- [7] M. Daffa, A. Fahreza, A. Luthfiarta, M. Rafid, M. Indrawan, and A. Nugraha, "Analisis Sentimen: Pengaruh Jam Kerja Terhadap Kesehatan Mental Generasi Z," *JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST)*, vol. 5, no. 1, pp. 2723–1453, 2024, doi: 10.52158/jacost.715.
- [8] V. C. Chandu, Y. Marella, G. S. Panga, S. Pachava, and V. Vadapalli, "Measuring the Impact of COVID-19 on Mental Health: A Scoping Review of the Existing Scales," Sep. 01, 2020, SAGE Publications Ltd. doi: 10.1177/0253717620946439.
- [9] M. Fachriza and H. Artikel, "Analisis Sentimen Kalimat Depresi Pada Pengguna X Dengan Naive Bayes, Support Vector Machine, Random Forest," 2023. [Online]. Available: <http://studentjournal.umpo.ac.id/index.php/komputek>
- [10] S. Jalukar, A. Ratnaparkhi, P. Shinde, S. Kunkulol, and V. Kulkarni, "SENTIMENT ANALYSIS FOR DEPRESSION DETECTION," *International Journal of Trendy Research In Engineering And Technology*, vol. 06, no. 03, 2022, doi: 10.54473/ijret.2022.6305.
- [11] J. Park, M.-H. Tsou, A. Nara, S. Cassels, and S. Dodge, "Developing a social sensing index for monitoring place-oriented mental health issues using social media (X) data," *Urban Informatics*, vol. 3, no. 1, Jan. 2024, doi: 10.1007/s44212-023-00033-5.
- [12] J. Zhang, "An Overview of the Application of Sentiment Analysis in Mental Well- being," *Applied and Computational Engineering*, vol. 8, no. 1, pp. 354–359, Aug. 2023, doi: 10.54254/2755-2721/8/20230186.
- [13] A. Bhardwaj, S. Bharany, and S. K. Kim, "Fake social media news and distorted campaign detection framework using sentiment analysis & machine learning," *Heliyon*, vol. 10, no. 16, Aug. 2024, doi: 10.1016/j.heliyon.2024.e36049.
- [14] J. Sharma and V. Tomer, "Depression detection using sentiment analysis of social media data," *AIP Conf Proc*, vol. 2481, no. 1, p. 020044, Nov. 2022, doi: 10.1063/5.0104192.
- [15] S. Aslan, S. Kiziloluk, and E. Sert, "TSA-CNN-AOA: X sentiment analysis using CNN optimized via arithmetic optimization algorithm," *Neural Comput Appl*, vol. 35, no. 14, pp. 10311–10328, May 2023, doi: 10.1007/s00521-023-08236-2.
- [16] S. Ghosal and A. Jain, "Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 1631–1639. doi: 10.1016/j.procs.2023.01.141.
- [17] D. Murthy, "Sociology of X/X: Trends, Challenges, and Future Research Directions," *Annu Rev Sociol*, vol. 17, p. 5, 2024, doi: 10.1146/annurev-soc-031021.
- [18] S. Cong and Y. Zhou, "A review of convolutional neural network architectures and their optimizations," *Artif Intell Rev*, vol. 56, no. 3, pp. 1905–1969, 2023, doi: 10.1007/s10462-022-10213-5.
- [19] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, 2022, doi: 10.1109/TNNLS.2021.3084827.
- [20] H. Alhakiem and E. Setiawan, "Aspect-Based Sentiment Analysis on X Using Logistic Regression with FastText Feature Expansion," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, pp. 840–846, Nov. 2022, doi: 10.29207/resti.v6i5.4429.
- [21] Adam, A. Z. R., & Setiawan, "Social Media Sentiment Analysis Using Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU)," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 9, no. 1, pp. 119–131, 2023, doi: 10.26555/jiteki.v9i1.25813
- [22] Ukaihongsar Wathana and Jitsakul, "Enhancing Sentiment Analysis Using Hybrid Deep Learning. In S. and J. W. and T. S. Meesad Phayung and Sodsee (Ed.)," *Proceedings of the 18th International Conference on Computing and Information Technology (IC2IT 2022)*, pp. 183–193, 2022.
- [23] Palomino, M. A., & Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Applied Sciences (Switzerland)*, vol. 12, no. 17, 2022, doi: 10.3390/app12178765
- [24] Harmandini, K. P., & Muslim, "Analysis of TF-IDF and TF-RF Feature Extraction on Product Review Sentiment," *Jurnal Dan Penelitian Teknik Informatika*, vol. 8, no. 2, 2024, doi: 10.33395/v8i2.13376
- [25] Adalakun, N. O., & Lasisi, "Deep Learning-Based Sentiment Analysis In Financial Markets Using Gated Recurrent Unit," *Andalasian International Journal of Applied Science, Engineering and Technology*, vol. 5, no. 1, pp. 27–38, 2025, doi: 10.25077/aijaset.v5i1.217
- [26] Nadeem, A., Aslam, N., Abid, M. K., & Fuzail, "Text-Based Sentiment Analysis Using CNN-GRU Deep Learning Model," In *J. inf. commun. technol. robot. appl (Vol. 14, Issue 1)*. Available: <http://www.jictra.com.pk/index.php/jictra,pISSN:2523-5729,eISSN:2523-5739>
- [27] Ren, J, "Multimodal Sentiment Analysis Based on BERT and ResNet," 2024, Available: <http://arxiv.org/abs/2412.03625>
- [28] Li, H., Lu, Y., & Zhu, H., "Multi-Modal Sentiment Analysis Based on Image and Text Fusion Based on Cross-Attention Mechanism," *Electronics (Switzerland)*, vol. 13, no. 11, 2024, doi: 10.3390/electronics13112069
- [29] Al-Tameemi, I. K. S., Feizi-Derakhshi, M.-R., Pashazadeh, S., & Asadpour, M., "A Comprehensive Review of Visual-Textual Sentiment Analysis from Social Media Networks," 2023, doi: 10.1007/s42001-024-00326-y
- [30] Malhotra, A., & Jindal, R., "Multimodal deep learning based framework for detecting depression and suicidal behaviour by affective analysis of social media post," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 6 no. 21, 2020, doi: 10.4108/eai.13-7-2018.164259