

Ekspansi Fitur dengan FastText untuk Analisis Sentimen di Media Sosial X Menggunakan Recurrent Neural Network dan Convolutional Neural Network

1st Robith Naufal Razzak

Fakultas Informatika

Universitas Telkom

Bandung, Indonesia

robithnaufal@students.telkomuniversity.ac.id

2nd Erwin Budi Setiawan

Fakultas Informatika

Universitas Telkom

Bandung, Indonesia

erwinbudisetiawan@telkomuniversity.ac.id

Abstrak - Dalam era digital, media X telah menjadi platform penting bagi masyarakat untuk ekspresi dan pertukaran informasi, terutama selama pemilihan umum, memungkinkan analisis sentimen untuk memahami opini publik secara real-time. Namun, tantangan utama adalah pengolahan teks yang tidak terstruktur dan kompleks, yang sering menghasilkan analisis yang kurang akurat. Untuk mengatasi ini, penelitian ini mengembangkan model hybrid deep learning yang mengintegrasikan Recurrent Neural Networks (RNN) dan Convolutional Neural Networks (CNN). Selain itu, penggunaan ekspansi fitur FastText memungkinkan model untuk mengatasi inkonsistensi kosakata dan meningkatkan pemahaman kontekstual, dengan mengidentifikasi dan menggantikan kata-kata yang kurang representatif dengan alternatif yang memiliki kemiripan semantik. Dua dataset digunakan untuk membangun similarity corpus dengan dataset media X sebesar 62.955 dan dataset IndoNews sebesar 126.673. Sistem menunjukkan tingkat akurasi maksimum pada model hybrid deep learning RNN-CNN yaitu sebesar 73,00% dengan peningkatan 2,50% terhadap model RNN dan 3,00% terhadap model CNN. Hasil ini menunjukkan bahwa integrasi RNN dan CNN, bersama dengan teknik ekspansi fitur FastText, meningkatkan kemampuan analisis sentimen, memberikan metode yang lebih efisien dan akurat dalam mengklasifikasikan sentimen publik.

Kata kunci - hybrid, RNN, CNN, fasttext, analisis sentimen

I. PENDAHULUAN

A. Latar Belakang

Dalam era digital saat ini, media sosial menjadi platform utama bagi masyarakat untuk berbagi informasi, berinteraksi, dan mengungkapkan pendapat mereka [1]. Media X, sebagai salah satu platform media sosial yang sangat populer, menyediakan sumber data yang kaya akan informasi dan sentimen publik [2]. Saat pemilihan umum, masyarakat menggunakan media ini untuk menyampaikan pandangan dan opini mereka, memberi

dukungan, atau mengkritik berbagai aspek terkait dengan pemilu. Data yang dihasilkan dari platform ini menjadi sumber informasi berharga untuk memahami pola pikir dan sentimen masyarakat terhadap calon, partai politik, dan isu-isu pemilu [3]. Pentingnya analisis sentimen dalam konteks ini adalah kemampuannya untuk memproses dan mengklasifikasikan teks dari cuitan yang tidak terstruktur menjadi data yang terstruktur dengan jelas, membedakan antara sentimen positif, negatif, dan netral [4]. Oleh karena itu, pengembangan model analisis sentimen yang akurat sangat diperlukan untuk memastikan keandalan dan relevansi hasil analisis, yang bisa digunakan untuk memahami dinamika pemilihan umum, memprediksi tren, dan mengidentifikasi perubahan sentimen secara langsung.

Saat ini, metode *deep learning* hingga *hybrid deep learning* telah menjadi pilihan utama dalam penelitian terkait analisis sentimen. Penelitian yang dilakukan oleh Muhammad Umer dan rekan-rekan menggunakan model klasifikasi CNN dan *word embedding* FastText berhasil mencapai nilai akurasi tertinggi sebesar 96% [5]. Dalam penelitian oleh M. A. Saputra dan E. B. Setiawan, model *deep learning* RNN dengan ekspansi fitur FastText mencapai nilai akurasi tertinggi sebesar 97,75% [6]. Rhanoui M dan rekan-rekannya melakukan penelitian menggunakan metode *hybrid deep learning* CNN-BiLSTM yang berhasil memperoleh nilai akurasi 90,66% [7]. Sementara itu, penelitian yang dilakukan oleh Ankita Sharma dan rekan-rekan tentang analisis sentimen pada media sosial X menggunakan kombinasi CNN-LSTM mendapatkan nilai akurasi 94% [8].

Ide terbaru dalam pengembangan analisis sentimen adalah penggunaan model *hybrid deep learning* dan ekspansi fitur untuk mengatasi masalah ketidaksesuaian kata dalam dataset. Sejauh pengetahuan penulis, penelitian tentang analisis sentimen yang memanfaatkan *hybrid deep learning* dan ekspansi fitur masih sedikit. *Hybrid deep learning* merupakan teknik yang menggabungkan dua atau lebih metode dari *deep*

learning. Deep learning bertujuan meniru bagaimana otak manusia memprediksi hasil berdasarkan data pengguna, sehingga memungkinkan mesin untuk belajar dari pengalaman [9]. Vektor semantik memuat berbagai fitur linguistik yang seringkali memiliki kesamaan fitur antara satu dengan yang lainnya. Metode ekspansi fitur, sebagai pendekatan yang lebih baru, bertujuan untuk mengatasi masalah ketidaksesuaian kosakata dalam vektor semantik dengan mengidentifikasi dan menggantikan kata-kata yang hilang dengan kata-kata lain yang memiliki kemiripan semantik [10].

Meskipun telah banyak penelitian mengenai analisis sentimen pada media X, masih terdapat kekurangan dalam pemahaman mendalam tentang sentimen yang dipengaruhi oleh bahasa yang sangat spesifik dan kontekstual. Penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menerapkan model *hybrid deep learning* yang menggabungkan RNN dan CNN. Dengan demikian, penelitian ini dapat mengevaluasi efektivitas dari penggunaan model *hybrid deep learning* RNN-CNN bersama ekspansi fitur FastText dalam melakukan analisis sentimen pada media sosial X.

B. Topik dan Batasannya

Dalam penelitian ini, penulis menyusun dan mengembangkan sebuah sistem untuk melakukan analisis sentimen berdasarkan pada data tweet yang berfokus pada pemilu di Indonesia tahun 2024, menggunakan metode *hybrid RNN-CNN* dan ekspansi fitur FastText. Kajian ini memiliki batasan, yaitu data yang dianalisis adalah tweet berbahasa Indonesia.

C. Tujuan

Tujuan utama dari penelitian ini adalah untuk mengimplementasikan analisis sentimen pada data media sosial X dengan menggunakan model *hybrid deep learning* yang menggabungkan *Recurrent Neural Networks* (RNN) dan *Convolutional Neural Networks* (CNN), serta Ekspansi Fitur melalui FastText. Penelitian ini juga bertujuan untuk mengevaluasi dan membandingkan kinerja antara model *hybrid RNN-CNN* dengan model-model *deep learning* yang berdiri sendiri, yaitu CNN dan RNN. Selain itu, penelitian ini mengukur nilai akurasi sebelum dan setelah implementasi Ekspansi Fitur dengan FastText, untuk menentukan efektivitas dari teknik ekspansi dalam meningkatkan akurasi model.

II. STUDI TERKAIT

Analisis sentimen adalah metode pengolahan teks yang digunakan untuk mengklasifikasikan sentimen menjadi positif, negatif, atau netral [4]. Bidang analisis sentimen memiliki potensi besar untuk dikembangkan, terutama karena pertumbuhan data yang pesat di Internet, seperti blog dan forum [11]. Penelitian terkait analisis sentimen telah berkembang, mulai dari penggunaan *deep learning* hingga *hybrid deep learning*. Berdasarkan berbagai penelitian sebelumnya, referensi tersebut akan digunakan sebagai pedoman dalam penelitian ini. Dalam penelitian [6] mengenai penggunaan *deep learning* pada

analisis sentimen berbasis aspek terhadap ulasan film di media X, digunakan metode klasifikasi RNN serta teknik ekstraksi fitur TF-IDF, ekspansi fitur FastText, dan SMOTE untuk meningkatkan akurasi penelitian. Hasilnya menunjukkan bahwa pada aspek plot, diperoleh akurasi sebesar 77,24% dan *F1 score* sebesar 77,19%; pada aspek akting, diperoleh akurasi sebesar 96,59% dan *F1 score* sebesar 96,59%; serta pada aspek sutradara, diperoleh akurasi sebesar 97,75% dan *F1 score* sebesar 97,74%.

Kemudian, dalam penelitian [12] yang berjudul "Sentiment Analysis using Recurrent Neural Network", berbagai teknik *machine learning* dan *deep learning* telah diterapkan dan dievaluasi, termasuk perbandingan antara CNN dengan ekspansi fitur word2vec, *Naïve Bayes*, RNN Conv, dan RNN dengan ekspansi fitur word2vec. Hasil penelitian tersebut menunjukkan bahwa RNN yang dikombinasikan dengan ekspansi fitur word2vec menghasilkan akurasi tertinggi sebesar 91,88%. Metode CNN dengan ekspansi fitur word2vec berada di peringkat kedua dengan akurasi 89,23%, diikuti oleh RNN Conv dengan akurasi 88,77%. Sementara itu, *Naïve Bayes* menempati posisi terakhir dengan akurasi yang jauh lebih rendah, yaitu 44,1%.

Dalam penelitian [5], penggunaan *deep learning* untuk klasifikasi teks menggunakan metode CNN dan *word embedding* FastText telah diterapkan pada tujuh dataset, yaitu Twitter US Airline Sentiment, AG's News, Yahoo! Answers, Amazon Review Full, Amazon Review Polarity, Yelp! Reviews Full, dan Yelp! Reviews Polarity. Hasilnya, model yang diusulkan berhasil mencapai akurasi tertinggi dengan nilai 0,96 pada dataset AG's News, precision tertinggi dengan nilai 0,94 pada dataset AG's News dan Yelp Reviews Polarity, recall tertinggi dengan nilai 0,95 pada dataset Yelp Reviews Polarity, serta F-score tertinggi dengan nilai 0,95 pada dataset Yelp Reviews Polarity. Dari penelitian tersebut, dapat disimpulkan bahwa *word embedding* FastText mampu meningkatkan akurasi.

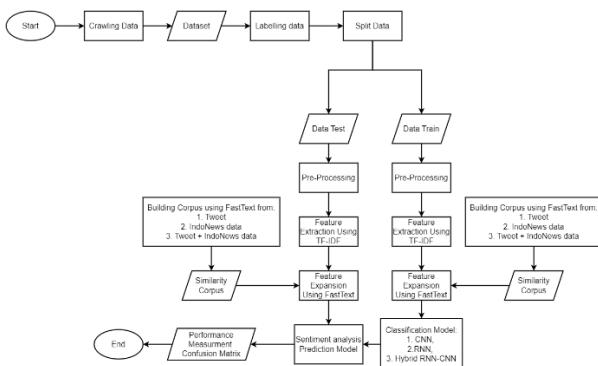
Pada penelitian [7], beberapa metode klasifikasi digunakan untuk analisis sentimen pada tingkat dokumen (*document-level*), yaitu CNN, LSTM, BiLSTM, CNN-LSTM, dan CNN-BiLSTM, dengan hasil akurasi masing-masing sebesar 88,00%; 85,87%; 86,40%; 90,13%; dan 90,66%. Sementara itu, dalam penelitian [8], beberapa metode klasifikasi diterapkan untuk analisis sentimen terhadap #BlackLivesMatter pada X, yaitu *Random Forest*, CNN, LSTM, BiLSTM, BERT base, BERT large, dan CNN-LSTM. Hasil akurasi dari metode klasifikasi tersebut berturut-turut adalah 77,35%; 79,46%; 76,21%; 77,58%; 87,45%; 89,13%; dan 94%. Dari kedua penelitian tersebut, dapat disimpulkan bahwa model *hybrid deep learning* memiliki akurasi yang lebih tinggi dibandingkan dengan model *deep learning* tunggal.

Berdasarkan penelitian-penelitian tersebut, diketahui bahwa model *hybrid* dapat meningkatkan akurasi dalam analisis sentimen. Kombinasi RNN-CNN dan penggunaan ekspansi fitur FastText memiliki potensi besar untuk diterapkan dalam analisis sentimen. Dalam penelitian ini, ekspansi fitur diterapkan pada vektor semantik untuk menggantikan kata-kata yang tidak ada

dengan kata-kata yang memiliki kesamaan semantik, menggunakan korpus yang dibangun dengan bantuan FastText. Berdasarkan pengetahuan peneliti hingga saat ini, penelitian yang mengeksplorasi analisis sentimen dengan menggabungkan teknik *hybrid deep learning* dan ekspansi fitur, khususnya pada konten X di Indonesia, masih sedikit. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi efektivitas penggunaan model *hybrid deep learning* RNN-CNN dan ekspansi fitur FastText dalam analisis sentimen.

III. SISTEM YANG DIBANGUN

Pembangunan sistem analisis sentimen ini dimulai dengan melakukan *crawling data* dari media sosial X. Setelah proses *crawling* selesai, data yang berhasil dikumpulkan diberi label secara manual sebelum dilanjutkan ke tahap pemisahan data. Setelah data dipisahkan, dilakukan tahap *preprocessing*. Setelah semua tahap tersebut selesai, dilakukan ekstraksi fitur, yang kemudian digunakan untuk ekspansi fitur. Dalam tahap klasifikasi, sistem ini menerapkan beberapa metode *deep learning* dan *hybrid deep learning*. Gambar 1 menunjukkan arsitektur sistem yang dibangun..



Gambar 1.

Alur Kerja Sistem Analisis Sentimen Menggunakan RNN-CNN

A. Crawling Data

Crawling data adalah proses pengumpulan data yang akan digunakan sebagai dataset dalam penelitian ini. Dataset yang digunakan berasal dari media sosial X, berupa cuitan yang dikumpulkan menggunakan *keyword* terkait pemilu 2024. Proses crawling data ini menghasilkan 62.955 data, dengan distribusi *keyword* pada dataset dijelaskan pada Tabel 1.

TABEL 1.
Distribusi Data

Kata Kunci	Total
Pilpres2024	7.326 (11,63%)
AniesMuhamin	10.341 (16,42%)
PrabowoGibran	25.420 (40,73%)
PolitikIndonesia	4.013 (6,37%)
GanjarMahfud	9.778 (15,53%)
DebatPilpres	6.077 (9,65%)

Jumlah Data	62.955
-------------	--------

B. Labeling data

Data yang diperoleh melalui proses *crawling* belum dikategorikan sebagai positif, negatif, atau netral. Oleh karena itu, pelabelan data menjadi proses penting untuk menetapkan kategori dari data tersebut [13]. Proses pelabelan akan dilakukan secara manual oleh enam anggota tim berdasarkan sentimen pada setiap cuitan yang akan digunakan dalam proses klasifikasi. Tingkatan sentimen yang digunakan dalam penelitian ini dibagi menjadi tiga, yaitu positif, netral, dan negatif. Distribusi label data dijelaskan pada Tabel 2.

TABEL 2.
Distribusi Label Dataset

Label	Total
Positif	21.114 (33,53%)
Netral	20.644 (32,79%)
Negatif	21.197 (33,63%)

C. Split Data

Split data adalah proses pembagian dataset menjadi dua bagian, yaitu *data train* dan *data test*. *Data train* digunakan untuk melatih model klasifikasi yang diajukan, sedangkan *data test* digunakan untuk pengujian atau validasi model yang telah dilatih sebelumnya pada *data train*. Tujuan dari *split data* adalah untuk melatih model pada sebagian data dan kemudian menguji kinerjanya pada data yang belum pernah dilihat sebelumnya, sehingga dapat mengevaluasi kemampuan model secara objektif.

D. Pre-Processing

Tahap *preprocessing* data dilakukan karena data yang didapat dari proses *crawling* merupakan data mentah. Proses ini bertujuan untuk membersihkan data dari dataset agar dapat digunakan pada tahap berikutnya. *Library Sastrawi* dan *NLTK* digunakan sebagai alat dalam *preprocessing ini*. Berikut adalah beberapa tahap yang digunakan dalam *preprocessing* :

1. Data Cleaning

Proses *data cleaning* dilakukan untuk membersihkan dataset dari elemen-elemen yang tidak relevan dengan analisis sentimen, seperti *username*, URL, angka, tanda baca, *hashtag*, *mention*, *emoticon*, dan simbol-simbol lainnya. Langkah ini penting untuk memastikan bahwa data yang digunakan dalam analisis hanya terdiri dari informasi yang mendukung proses klasifikasi sentimen, sehingga model dapat bekerja dengan lebih efektif dan akurat.

2. Case Folding

Semua huruf kapital (*uppercase*) diubah menjadi huruf kecil (*lowercase*) pada proses ini.

3. Stopword Removal

Proses *stopword removal* bertujuan untuk menghapus kata-kata yang sering digunakan tetapi tidak memberikan makna yang signifikan dalam analisis teks. Dalam proses ini, data *stopword* diambil dari *library* nltk.

4. Stemming

Proses *stemming* digunakan untuk mengubah kata-kata ke bentuk dasar. Tujuan dari proses ini adalah untuk mengurangi variasi kata yang pada dasarnya memiliki arti atau kata dasar yang sama. *Library* sastrawi digunakan pada proses ini.

5. Normalization

Normalization bertujuan mengubah kata-kata yang tidak terdaftar dalam kamus (*out-of-vocabulary*, OOV) menjadi kata-kata yang lebih umum. Proses ini sangat krusial karena mempermudah analisis bahasa dalam memahami makna leksikal. Dengan normalisasi kata-kata OOV, efisiensi dalam pemrosesan bahasa dapat dioptimalkan [14].

6. Tokenizing

Tokenizing adalah metode yang digunakan untuk memecah teks menjadi unit-unit yang lebih kecil seperti kata, frasa, atau simbol, yang disebut token. Proses ini memudahkan analisis dan pemrosesan data yang lebih efisien.

E. Ekstraksi Fitur

Ekstraksi fitur adalah metode untuk mengambil informasi yang relevan dari data mentah [15]. Dalam penelitian ini, teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) dipilih sebagai metode ekstraksi fitur. TF-IDF sering digunakan dalam klasifikasi teks karena kemampuannya dalam mengidentifikasi pentingnya sebuah kata dalam dokumen secara efektif. Metode ini terdiri dari dua komponen utama: *Term Frequency* (TF) yang mengukur frekuensi kemunculan suatu kata dalam dokumen tertentu, menandakan pentingnya kata tersebut, dan *Inverse Document Frequency* (IDF) yang menghitung berapa banyak dokumen yang mencakup kata tersebut, menunjukkan pentingnya kata dalam keseluruhan korpus. Nilai TF-IDF diperoleh dengan mengalikan nilai TF dengan nilai IDF [16]. TF-IDF dihitung dengan persamaan berikut:

$$TF - IDF(t, d) = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (1)$$

dengan $tf_{t,d}$ sebagai frekuensi kata t dalam dokumen d , N adalah total jumlah dokumen dalam korpus, dan df_t menunjukkan jumlah dokumen yang mengandung kata t .

F. Ekspansi Fitur menggunakan FastText

Dalam penelitian ini, model FastText digunakan untuk ekspansi fitur dengan menggunakan modul gensim. FastText, dengan memanfaatkan informasi subkata, mampu merepresentasikan kata-kata secara efektif, khususnya dalam kondisi saat *word embeddings* tradisional tidak berfungsi dengan baik [17]. Model ini memanfaatkan korpus yang terdiri dari kumpulan kata dari dataset tweet dan berita. Korpus yang digunakan meliputi korpus tweet, korpus IndoNews, serta kombinasi

dari keduanya. Tabel 3 menampilkan jumlah kosakata dari korpus yang telah dibuat.

Corpus Similarity	Jumlah Kosakata pada Korpus	
	Total	
Tweet	62.955	
IndoNews	126.673	
Tweet + IndoNews	189.628	

Korpus tersebut kemudian digunakan untuk menilai kesamaan dengan kata-kata tertentu, yang dijelaskan pada tabel 4.

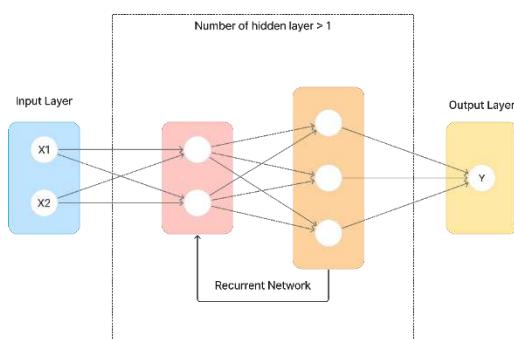
Kata	Korpus Similarity Top 5				
	Top 1	Top 2	Top 3	Top 4	Top5
capres	bapres	wacapers	wapres	capresku	capresmu

Setelah memproleh hasil similarity korpus, langkah selanjutnya adalah melakukan ekspansi fitur. Ekspansi fitur merupakan teknik yang digunakan untuk mengatasi masalah ketidaksesuaian kosakata, dengan tujuan untuk mengidentifikasi fitur yang memiliki nilai 0 dan menggantikannya dengan kata yang memiliki kemiripan. Jika kata serupa ditemukan dengan nilai 1, maka nilai yang semula 0 akan diubah menjadi 1 [10]. Sebagai contoh, terdapat kalimat "mending pilih capres nomor berapa", misal kata "capres" memiliki nilai 0 dan kata "bapres" memiliki nilai 1. Maka, nilai untuk kata "capres" dapat diganti menjadi 1 karena memiliki kesamaan dengan kata "bapres" sesuai dengan tabel 4.

G. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) merupakan arsitektur jaringan saraf tiruan yang sering digunakan dalam analisis sentimen. Keunggulan RNN terletak pada kemampuan memiliki memori internal yang menyimpan informasi penting dari data sebelumnya untuk membuat prediksi yang lebih akurat. RNN bekerja dengan desain arsitektur yang berulang, memungkinkan informasi dari data sebelumnya untuk tetap tersimpan [18]. Gambar 2 menggambarkan ilustrasi RNN, dengan *input* adalah serangkaian data dengan setiap elemennya direpresentasikan sebagai vektor. RNN dilengkapi dengan "*Hidden State*" atau memori internal yang berfungsi untuk menyimpan informasi dari serangkaian data yang telah diproses sebelumnya. Struktur dasar RNN memungkinkan untuk menerima *input* saat ini bersama dengan *hidden state* sebelumnya di setiap *time step*, menghasilkan *output* baru dan memperbarui *hidden state*, yang mendefinisikan *recurrent network*. RNN bisa menghasilkan *output* di setiap *time step* atau hanya di *time step* terakhir setelah seluruh urutan selesai diproses. *Output* ini dapat digunakan untuk berbagai tugas,

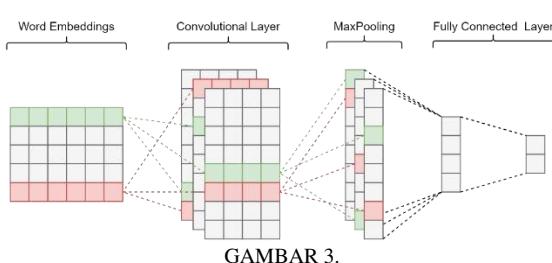
termasuk identifikasi teks, prediksi waktu, atau generasi teks..



GAMBAR 2.
Ilustrasi Recurrent Neural Network

H. Convolutional Neural Network (CNN)

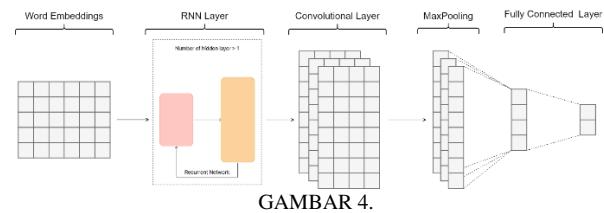
Convolutional Neural Network (CNN) adalah arsitektur *deep learning* yang berjenis *artificial neural network*. *Convolutional Neural Network* (CNN) telah terbukti menjadi alat yang efektif dalam analisis sentimen, khususnya dalam pemrosesan teks dan dokumen [19]. Penggunaan CNN dalam analisis sentimen memungkinkan sistem untuk secara otomatis mengekstrak ciri-ciri penting dari teks, memahami hubungan antar kata, dan mengidentifikasi sentimen yang terkandung dalam suatu dokumen [20]. Gambar 3 menggambarkan susunan lapisan CNN yang meliputi beberapa lapisan berbeda: *Convolutional layer*, *Pooling layer*, dan *Fully Connected Layer*. *Convolutional layer* bertugas menentukan hasil keluaran *neuron* yang terkoneksi dari lapisan *input* [21]. *Pooling layer* berfungsi mengurangi ukuran sampel, sehingga memungkinkan data diwakili dalam ukuran yang lebih kecil dan memudahkan penanganan *overfitting* [22]. Lapisan terakhir adalah *Fully Connected Layer*, di mana setiap *neuron* di lapisan ini terhubung satu sama lain [22].



GAMBAR 3.
Ilustrasi Convolutional Neural Network

I. Hybrid Model

Hybrid deep learning adalah model terbaru yang dihasilkan dari penggabungan dua atau lebih metode *deep learning* yang berbeda [23]. Pada tahap ini, sistem *hybrid* dibangun dengan menggabungkan RNN dan CNN untuk analisis sentimen. Pada model ini RNN adalah model pertama yang menerima *input*, yang kemudian *output* dari RNN digunakan untuk melatih model CNN. Pendekatan *hybrid* ini dirancang untuk memperkuat kemampuan model dalam memahami struktur kalimat dan merespons dinamika perubahan urutan kata.



GAMBAR 4.
Ilustrasi Hybrid Model

J. Performance Evaluation

Proses evaluasi dilakukan menggunakan *Confusion Matrix*. *Confusion matrix* menjelaskan jumlah data uji yang benar diklasifikasi dan yang salah diklasifikasi oleh model, seperti yang dijelaskan pada tabel 5.

TABEL 5.
Confusion Matrix

Actual Values	Predicted Values		
	Positive	Neutral	Negative
Positive	TP(Positive)	FN(Neutral)	FN(Negative)
Neutral	FP(Positive)	TN(Neutral)	FN(Negative)
Negative	FP(Positive)	FN(Neutral)	TN(Negative)

Dengan menggunakan *confusion matrix*, kita dapat menghitung berbagai evaluasi kinerja model seperti *accuracy*, *precision*, *recall*, dan *F1-score*, dengan formula sebagai berikut:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Accuracy adalah ukuran yang menunjukkan persentase prediksi yang tepat, termasuk prediksi positif dan negatif, dari total data. Sebaliknya, *precision* adalah ukuran yang menilai berapa banyak prediksi positif yang benar dari semua prediksi positif yang dilakukan. *Recall* menentukan berapa banyak prediksi positif yang benar terhadap total kasus positif sebenarnya dalam dataset. *F1-score*, yang merupakan metrik terakhir, merupakan kombinasi dari rata-rata tertimbang antara *precision* dan *recall* yang memberikan penilaian keseluruhan terhadap efektivitas model [24].

IV. EVALUASI

Dalam studi ini, dilakukan serangkaian pengujian yang terstruktur dalam lima skenario berbeda. Berikut adalah rincian dari masing-masing skenario pengujian:

1. Skenario I, menguji rasio *split data* untuk memperoleh nilai dasar atau *baseline*.
2. Skenario II, menguji integrasi *max feature* dalam ekstraksi fitur TF-IDF yang bertujuan untuk memperoleh nilai *max feature* dengan akurasi terbaik.
3. Skenario III, menguji penggunaan n-gram dalam ekstraksi fitur TF-IDF untuk mendapatkan konfigurasi n-gram dengan akurasi terbaik.
4. Skenario IV, melakukan pengujian dengan mengaplikasikan ekspansi fitur menggunakan FastText untuk memperoleh model dengan akurasi terbaik setelah melakukan ekspansi fitur.
5. Skenario V, melakukan pengujian dengan model *hybrid* untuk mendapatkan model dengan hasil akurasi terbaik.

Setiap skenario akan dieksekusi sebanyak lima kali. Hasil pengujian dari setiap skenario akan diukur berdasarkan rata-rata nilai akurasi yang dicapai.

A. Skenario 1

Dalam Skenario ini, dilakukan pengujian untuk mencari rasio *split data* terbaik untuk menghasilkan akurasi tertinggi. Rasio *split data* yang diuji adalah 90:10, 80:20, dan 70:30 (data latih:data uji) dengan menggunakan *all feature* (30103). Tabel 6 merupakan hasil dari skenario 1.

Tabel 6. Skenario Split Data

Split Ratio	Akurasi (%)	
	RNN	CNN
90:10	67,8 0	67,3 0
80:20	67,1 9	67,1 5
70:30	66,1 7	65,6 5

Berdasarkan tabel 6, dapat diketahui bahwa rasio pembagian 90:10 menghasilkan nilai akurasi terbaik untuk kedua model yang menjadikan rasio pembagian ini menjadi baseline pada penelitian ini.

B. Skenario 2

Dalam Skenario ini pengujian dilakukan dengan mencari nilai *max feature* yang dapat menghasilkan nilai akurasi tertinggi. Nilai *max feature* yang diuji adalah 5.000, 10.000, 15.000, 20.000, 25.000, dan 30.103 (*all features*). Tabel 7 menunjukkan hasil dari skenario 2.

**TABEL 7.
Skenario Max Feature**

Max Feature	Akurasi(%)	
	RNN	CNN
5.000	67,64(-0,16)	67,38(+0,08)
10.000	67,79(-0,10)	67,30(+0,00)
15.000	67,91(+0,11)	67,70(+0,40)

20.000	68,07(+0,27)	67,35(+0,05)
25.000	67,42(-0,38)	67,62(+0,32)
30.103 (<i>baseline</i>)	67,80	67,30

Berdasarkan tabel 7, nilai *max feature* terbaik yang menghasilkan nilai akurasi terbaik pada model RNN adalah 20.000, sedangkan pada model CNN nilai *max feature* menghasilkan nilai akurasi terbaik adalah 15.000. Oleh karena itu, RNN dengan rasio pembagian 90:10 dan *max feature* 20.000 lalu CNN dengan rasio pembagian 90:10 dan *max feature* 15.000 ini digunakan untuk skenario selanjutnya.

C. Skenario 3

Dalam Skenario ini dilakukan serangkaian pengujian untuk mengidentifikasi konfigurasi n-gram yang memberikan tingkat akurasi terbaik. Konfigurasi yang diuji dalam skenario ini mencakup unigram, bigram, trigram, serta kombinasi dari unigram-bigram (uni-bigram) dan unigram-bigram-trigram (uni-bi-trigram). Tabel 8 merupakan hasil dari skenario 3.

**TABEL 8.
Skenario Konfigurasi n-gram**

n-gram	Akurasi (%)	
	RNN	CNN
Unigram	68,07(+0,27)	67,70(+0,40)
Bigram	65,30(-2,50)	64,66(-2,64)
Trigram	58,07(-9,73)	58,45(-8,85)
Uni-Bigram	68,59(+0,79)	68,30(+1,00)
Uni-Bi-Trigram	68,29(+0,49)	67,44(+0,14)

Berdasarkan Tabel 8, Uni-Bigram mendapatkan nilai akurasi tertinggi dibandingkan dengan nilai n-gram lainnya yaitu Unigram, Bigram, Trigram, dan Uni-Bi-Trigram pada kedua model. Dengan hasil tersebut Uni-Bigram digunakan dalam skenario berikutnya.

D. Skenario 4

Dalam skenario keempat, pengujian dilakukan dengan menerapkan teknik ekspansi fitur menggunakan FastText pada kedua model. Teknik ini diimplementasikan untuk memperkaya dataset berdasarkan korpus similarity yang telah dibangun sebelumnya. Ada tiga jenis korpus yang digunakan, yaitu korpus Tweet, korpus IndoNews, dan gabungan dari keduanya yang disebut korpus Tweet + IndoNews, dengan tingkat similarity top 1, top 5, top 10, dan top 15. Tabel 9 menampilkan hasil dari skenario keempat ini.

**TABEL 9.
Skenario Ekspansi Fitur dengan FastText**

Model	Rank	Akurasi (%)		
		Tweet	IndoNews	Tweet + IndoNews
RNN	Top 1	69,53(+1,73)	69,73(+1,93)	70,00(+2,20)
	Top 5	68,94(+1,14)	68,47(+0,67)	69,61(+1,81)

	Top 10	66,89(-0,91)	67,42(-0,38)	68,62(+0,82)
	Top 15	65,65(-2,15)	67,65(-0,15)	68,60(+0,80)
CNN	Top 1	69,06(+1,76)	68,87(+1,57)	69,39(+2,09)
	Top 5	68,41(+1,11)	68,67(+1,37)	68,77(+1,47)
	Top 10	66,94(-0,36)	68,27(+0,97)	68,62(+1,32)
	Top 15	65,97(-1,33)	68,33(+1,03)	67,89(+0,59)

Tabel 9 menunjukkan penerapan ekspansi fitur FastText menggunakan korpus Tweet + IndoNews dengan similarity top 1 menghasilkan akurasi terbaik dibandingkan dengan korpus Tweet dan korpus Indonews serta *similarity rank* lainnya pada kedua model. Akurasi model RNN meningkat 2,20% dari *baseline*. Sementara pada model CNN akurasi meningkat 2,09% dari *baseline*. Dapat disimpulkan bahwa ekspansi fitur meningkatkan akurasi.

E. Skenario 5

Dalam skenario ini, pengujian dijalankan menggunakan model *hybrid deep learning* RNN-CNN dengan ekspansi fitur FastText. Skenario 5 mengimplementasikan pengujian yang serupa dengan skenario 4, namun dengan memanfaatkan model *hybrid*.

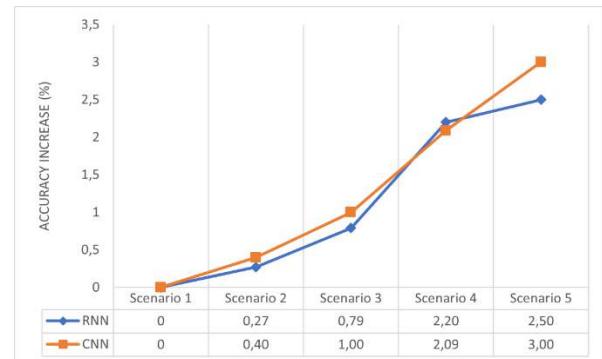
TABEL 10.
Akurasi Model Hybrid

Model	Baseline Score	Rank	Akurasi (%)		
			Tweet	IndoNews	Tweet + IndoNews
Hybrid Model	RNN(67.80)	1	69,65(+1,85)	69,71(+1,91)	70,30(+2,50)
		5	68,80(+1,00)	68,86(+1,06)	69,73(+1,93)
		10	67,25(-0,55)	67,99(+0,19)	69,25(+1,45)
		15	65,88(-1,92)	67,43(-0,37)	68,43(+0,63)
RNN-CNN	CNN(67.30)	1	69,65(+2,35)	69,71(+2,41)	70,30(+3,00)
		5	68,80(+1,50)	68,86(+1,56)	69,73(+2,43)
		10	67,25(-0,05)	67,99(+0,69)	69,25(+1,95)
		15	65,88(-1,42)	67,43(+0,13)	68,43(+1,13)

Berdasarkan tabel 10, penggunaan model *hybrid* menunjukkan peningkatan akurasi dibanding dengan *baseline* penelitian. Peningkatan akurasi terbaik pada model *hybrid* didapat dari top 1 similarity pada korpus Tweet + IndoNews.

F. Analisis Hasil Pengujian

Gambar 3 menunjukkan peningkatan akurasi dari tiap skenario yang telah dilakukan.



GAMBAR 5.
Peningkatan Akurasi tiap Skenario

Berdasarkan gambar 3, skenario 2 menunjukkan peningkatan akurasi pada kedua model yang diuji dibandingkan dengan skenario 1 yang menjadi *baseline* dalam penelitian ini. Model RNN dengan nilai *max feature* sebesar 20.000 memberikan peningkatan akurasi tertinggi sebesar 0,27%, sedangkan model CNN dengan nilai *max feature* sebesar 15.000 memberikan peningkatan tertinggi sebesar 0,40%. Dalam skenario 3, konfigurasi uni-bigram menunjukkan peningkatan akurasi tertinggi pada kedua model dibandingkan dengan konfigurasi n-gram lainnya, dengan peningkatan akurasi untuk model RNN dan CNN masing-masing sebesar 0,79% dan 1,00%. Pada skenario 4, penambahan ekspansi fitur FastText pada setiap model yang diuji menunjukkan peningkatan akurasi pada kedua model, yaitu 2,20% pada model RNN dan 2,09% pada model CNN, keduanya menggunakan korpus Tweet + IndoNews top 1. Peningkatan akurasi tertinggi dicapai pada skenario 5, yaitu model *hybrid* menggunakan rasio *split data* 90:10, *max feature* sebesar 20.000, konfigurasi n-gram uni-bigram, dan ekspansi fitur FastText dengan korpus Tweet + IndoNews top 1. Dibandingkan dengan *baseline*, model *hybrid* menunjukkan peningkatan sebesar 2,50% untuk RNN dan 3,00% untuk CNN. Secara keseluruhan, model *hybrid* menunjukkan peningkatan akurasi tertinggi di semua skenario yang telah diuji.

Dalam penelitian ini, uji signifikansi statistik digunakan untuk memvalidasi perubahan akurasi. Kriteria penentuan signifikansi statistik adalah P-Value, perbandingan dinyatakan signifikan secara statistik apabila P-Value kurang dari 0,05, dan sangat signifikan secara statistik apabila kurang dari 0,01. Selanjutnya, Z-Value digunakan dengan confidence level 95%, yang memerlukan Z-Value lebih dari 1,96 untuk menyatakan perbandingan signifikan secara statistik.

Hasil uji signifikansi statistik pada tabel 11 menunjukkan bahwa perubahan dari skenario 3 ke skenario 4 meningkat signifikan secara statistik pada kedua model. Hal ini membuktikan penggunaan ekspansi fitur meningkatkan performa model secara signifikan. Kemudian perubahan dari skenario 1 (model *baseline*) ke skenario 5 (model terbaik) yang dilakukan dalam

penelitian ini juga meningkat sangat signifikan secara statistik

TABEL 11.
Uji signifikansi statistik semua skenario

Model		S1→ S2	S2→ S3	S3→ S4	S4→ S5	S1→ S5
CN N	Z-Value	0,84 71	1,89 12	4,52 70	2,57 84	13,76 55
	P-Value	0,39 68	0,05 85	0,00 00	0,00 99	0,000 0
	Significant?	Fals e	Fals e	True	True	True
RN N	Z-Value	1,16 87	2,42 77	6,54 78	1,60 52	9,310 3
	P-Value	0,24 25	0,01 51	0,00 00	0,10 84	0,000 0
	Significant?	Fals e	True	True	Fals e	True

V. KESIMPULAN

Penelitian yang dilakukan berhasil menunjukkan efektivitas model *hybrid* RNN-CNN dengan tambahan ekspansi fitur FastText, yang dibangun dari 62.955 Tweet dan 126.673 data IndoNews. Hasil penelitian menunjukkan peningkatan signifikan dalam akurasi, model *hybrid* memberikan kinerja yang lebih unggul dibandingkan dengan penggunaan model-model *deep learning* seperti CNN dan RNN.

Berdasarkan hasil uji, model *hybrid* dengan konfigurasi *split data* 90:10, *max feature* sebesar 20.000, konfigurasi n-gram uni-bigram, dan penggunaan ekspansi fitur FastText dengan korpus Tweet + IndoNews top 1 mencapai akurasi 73,00%, dengan peningkatan 2,50% terhadap model RNN dan 3,00% terhadap model CNN. Peningkatan ini menunjukkan bahwa kombinasi RNN dan CNN dalam model *hybrid*, bersama dengan strategi ekspansi fitur, berhasil memberikan pemahaman yang lebih dalam dan akurat terhadap sentimen dalam data. Ekspansi fitur menggunakan FastText terbukti efektif dalam meningkatkan akurasi, menunjukkan pentingnya penambahan informasi leksikal dalam analisis sentimen yang akurat.

Dengan mempertimbangkan hasil tersebut, sangat disarankan untuk menerapkan model ini pada dataset yang lebih luas dan dari berbagai platform media sosial untuk menilai adaptabilitas dan kekuatan model. Penelitian lebih lanjut dapat melibatkan pengujian metode ekspansi fitur lain seperti GloVe atau Word2Vec untuk membandingkan keefektifannya dalam analisis sentimen. Selain itu, optimasi arsitektur model hybrid juga dapat menjadi fokus penelitian berikutnya untuk meningkatkan performa.

REFERENSI

- [1] Kaplan, Andreas & Haenlein, Michael. 2010. Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*. 53. 59-68. 10.1016/j.bushor.2009.09.003.
- [2] Bollen, Johan & Mao, Huina & Zeng, Xiao-Jun. 2010. Twitter Mood Predicts the Stock Market. *Journal of Computational Science*. 2.10.1016/j.jocs.2010.12.007.
- [3] Budiharto, W., Meiliana, M. 2018. Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *J Big Data* 5, 51. <https://doi.org/10.1186/s40537-018-0164-1>.
- [4] Tul, Qurat & Ali, Mubashir & Riaz, Amna & Noureen, Amna & Kamranz, Muhammad & Hayat, Babar & Rehman, Aziz Ur. 2017. Sentiment Analysis Using Deep Learning Techniques: A Review. *International Journal of Advanced Computer Science and Applications*. 8. 10.14569/IJACSA.2017.080657.
- [5] Umer, Muhammad & Imtiaz, Zainab & Ahmad, Muhammad & Nappi, Michele & Medaglia, Carlo & Choi, Gyu Sang & Mehmood, Arif. (2022). Impact of convolutional neural network and FastText embedding on text classification. *Multimedia Tools and Applications*. 82. 10.1007/s11042-022-13459-x.
- [6] M. A. Saputra and E. B. Setiawan, "Aspect Based Sentiment Analysis Using Recurrent Neural Networks (RNN) on Social Media Twitter," 2023 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia, 2023, pp. 265-270, doi:10.1109/ICoDSA58501.2023.10276768..
- [7] Rhanoui M, Mikram M, Yousfi S, Barzali S. A CNN-BiLSTM Model for Document-Level Sentiment Analysis. *Machine Learning and Knowledge Extraction*. 2019; 1(3):832-847. <https://doi.org/10.3390/make1030048>.
- [8] Sharma, Ankita & Rani, Shalli & Bashir, Ali & Alhudhaif, Adi & Koundal, Deepika & RN, Emine. 2022. An efficient CNN-LSTM model for sentiment detection in #BlackLivesMatter. *Expert Systems with Applications*. 193. 116256. 10.1016/j.eswa.2021.116256.
- [9] Muniasamy, A., & Alasiry, A. (2020). Deep Learning: The Impact on Future eLearning. *International Journal of Emerging Technologies in Learning (iJET)*, 15(01), pp. 188–199. <https://doi.org/10.3991/ijet.v15i01.11435>.
- [10] Setiawan, Erwin & Widayantoro, Dwi & Surendro, Kridanto. 2016. Feature expansion using word embedding for tweet topic classification. 1-5. 10.1109/TSSA.2016.7871085.
- [11] B. Keith Norambuena, E. F. Lettura, and C. M. Villegas, “Sentiment analysis and opinion mining applied to scientific paper reviews,” *Intelligent Data Analysis*, vol. 23, no. 1, pp. 191-214, 2019, doi: 10.3233/IDA-173807.
- [12] L. Kurniasari and A. Setyanto, “Sentiment Analysis using Recurrent Neural Network,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Mar. 2020. doi: 10.1088/1742-6596/1471/1/012018.
- [13] J. Hartmann, J. Huppertz, C. Schamp, and M. Heitmann, “Comparing automated text

- classification methods," International Journal of Research in Marketing, vol. 36, no. 1, pp. 20–38, Mar. 2019, doi: 10.1016/j.ijresmar.2018.09.009.
- [14] Salsabila, Nikmatun & Winatmoko, Yosef & Septiandri, Ali & Jamal, Ade. (2018). Colloquial Indonesian Lexicon. 226-229. 10.1109/IALP.2018.8629151.
- [15] Alessa, Ali & Faezipour, Miad. (2018). Tweet Classification Using Sentiment Analysis Features and TF-IDF Weighting for Improved Flu Trend Detection. 10.1007/978-3-319-96136-1_15.
- [16] A. O. Salau and S. Jain, "Feature Extraction: A Survey of the Types, Techniques, Applications," in 2019 International Conference on Signal Processing and Communication (ICSC), IEEE, Mar. 2019, pp. 158–164. doi: 10.1109/ICSC45622.2019.8938371..
- [17] Jianming, Zhang. "Weibo Text Sentiment Classification Model Based on FastText-BERT-Attention." 2 (2024):29-35. doi: 10.62517/jbdc.202401104.
- [18] Aufa, Rizki & Prasetyowati, Sri & Sibaroni, Yuliant. (2023). The Effect of Feature Weighting on Sentiment Analysis TikTok Application Using The RNN Classification. Building of Informatics, Technology and Science (BITS). 5. 10.47065/bits.v5i1.3597.
- [19] Nurdin, Arliyanti & Seno aji, Bernadus & Bustamin, Anugrayani & Abidin, Zaenal. (2020).
- PERBANDINGAN KINERJA WORD EMBEDDING WORD2VEC, GLOVE, DAN FASTTEXT PADA KLASIFIKASI TEKS. Jurnal Tekno Kompak. 14. 74. 10.33365/jtk.v14i2.732.
- [20] Zhang, D., & Lee, W. S. (2018). Sensitivity analysis of convolutional neural networks for sentence classification. Knowledge-Based Systems, 139, 30–39.
- [21] O. Bayat, S. Aljawarneh, H. F. Carlak, International Association of Researchers, Institute of Electrical and Electronics Engineers, and Akdeniz Üniversitesi, "Understanding of a Convolutional Neural Network," in 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, Aug. 2017, pp. 1–6. doi: 10.1109/ICEEngTechnol.2017.8308186.
- [22] H. Khotimah, E. Budi, and I. Kurniawan, "Implementation Information Gain Feature Selection for Hoax News Detection on Twitter using Convolutional Neural Network (CNN)," 2020, doi: 10.34818/indojc.2021.5.3.506.
- [23] M. U. Salur and I. Aydin, "A Novel Hybrid Deep Learning Model for Sentiment Classification," IEEE Access, vol. 8, pp. 58080–58093, 2020, doi: 10.1109/ACCESS.2020.2982538.
- [24] Babacar, Gaye., Aziguli, Wulamu. (2019). Sentiment Analysis of Text Classification Algorithms Using Confusion Matrix. doi: 10.1007/978-981-15-1922-2_16.