

ANALISIS SENTIMEN KOMENTAR POLITIK DI MEDIA SOSIAL X DENGAN PENDEKATAAN DEEP LEARNING

Ayu Liawati ¹, Riri Narasati ², Dodi Solihudin ³, Cep Lukman Rohmat ⁴, Sandy Eka Permana ⁵

^{1,2,3} Teknik Informatika, ⁴ Rekaya Perangkat Lunak, ⁵ Manajemen Informatika, STMIK IKMI Cirebon
Jl. Perjuangan No.10B-Majasem, Kesambi Karya Mulya, Kota Cirebon, Jawa Barat Indonesia
ayuliawati9e@gmail.com

ABSTRAK

Media sosial, sebagai *platform* utama interaksi dan berbagi informasi, memberikan peluang bagi Masyarakat untuk mengungkapkan sudut pandang mereka. Salah satu contoh yang populer adalah media sosial X, suatu jejaring sosial yang memungkinkan penggunanya berbagi pesan singkat yang mencakup teks, gambar, video, dan tautan. Dampak media sosial sangat terasa, khususnya dalam konteks politik di mana penyebaran informasi dan alur komunikasi semakin mudah. Meskipun demikian, adanya peristiwa negatif seperti penyebaran informasi palsu dan ujaran kebencian memberikan tantangan yang perlu diatasi melalui analisis sentimen. Media sosial memainkan peran penting dalam membentuk opini publik. Proses penggalangan dukungan melalui media sosial X telah mempercepat penyebaran pesan politik. Oleh karena itu, analisis sentimen menjadi penting untuk memahami pandangan dan sikap masyarakat di media sosial terkait isu-isu tertentu. Maka berfokus pada analisis sentimen terhadap komentar politik di media sosial X dengan menggunakan metode Convolutional Neural Network (CNN) pemanfaatan model pelatihan *Word2Vec*. Analisis ini menunjukkan bahwa model yang digunakan mampu memberikan hasil analisis sentimen dengan kinerja yang memuaskan, memberikan pemahaman yang lebih mendalam mengenai pandangan masyarakat terhadap permasalahan politik. Implikasi hasil ini dapat memberikan arahan bagi pengambil keputusan politik dan praktis yang tertarik memanfaatkan teknik *deep learning* dalam menganalisis sentimen, bukan hanya dalam konteks politik, melainkan juga dalam berbagai disiplin lainnya. Hasil yang telah dilakukan dengan menggunakan algoritma *Convolution Neural Network* menunjukkan nilai akurasi terbaik yaitu 81%.

Kata kunci : analisis sentimen, komentar politik, CNN, media sosial X, deep learning

1. PENDAHULUAN

Media sosial adalah sebuah *platform* yang sering dimanfaatkan oleh masyarakat untuk berinteraksi, berbagi informasi, serta menyampaikan pandangan mereka. Menurut *fokustekno* salah satu contoh media sosial yang sedang populer adalah X. sebuah layanan jejaring sosial yang memungkinkan pengguna untuk mengirim dan membaca pesan singkat yang dikenal sebagai "tweet." Tweet dapat berupa teks, gambar, video, atau tautan. Media sosial X ini memungkinkan pengguna untuk mengikuti akun-akun yang menarik bagi mereka dan membangun jejaring sosial dengan individu yang memiliki minat atau pandangan serupa. Selain itu, X juga memungkinkan pengguna untuk mengikuti perkembangan berita dan tren terkini, seperti politik, olahraga, kesehatan, hiburan, dan ilmu pengetahuan [1]. Oleh karena itu, dibutuhkan terobosan dalam peningkatan fasilitas pembelajaran guna menjaga keterkinian secara berkelanjutan [2].

Media sosial X memainkan peran dalam membentuk opini publik. Proses penggalangan dukungan melalui media sosial X telah mempercepat penyebaran pesan politik. Oleh karena itu, media sosial X dalam ranah komunikasi politik tidak dapat dihindari, tetapi juga sebagai pendorong partisipasi aktif masyarakat. Pengguna media sosial X dapat berkontribusi secara aktif dan memberikan tanggapan terbuka, baik dalam menyebarkan informasi maupun memberikan respons secara online dengan kecepatan

yang luar biasa. Dengan demikian, kontribusi media sosial dalam mengembangkan keterlibatan warga menjadi semakin krusial dalam konteks dinamika komunikasi politik modern [3].

Penelitian sebelumnya mengenai analisis twitter bahasa Indonesia berbasis *Word2Vec*, menggunakan *Deep Convolutional Neural Network* menunjukkan kinerja yang optimal, dengan tingkat akurasi mencapai 76,40% [4]. Penelitian yang lain menunjukkan bahwa model pendekatan *deep learning* dengan menerapkan metode klasifikasi Convolutional Neural Network. Hasil dari penelitian ini mencerminkan tingkat akurasi sebesar 72.50%, [5]. Penelitian selanjutnya dalam melakukan evaluasi sentiment terhadap produk kategori tools & home di platform Amazon, pendekatan model CNN dan LSTM. akurasi model LSTM mencapai sekitar 60,74%, sementara model CNN mencapai sekitar 57,99%. [6].

Penelitian ini memanfaatkan informasi dari pengguna media sosial X untuk menganalisis sentimen terhadap komentar politik dengan tujuan mengidentifikasi respons yang diberikan oleh pengguna tentang komentar politik yang sedang berlangsung [7]. Penelitian ini menggunakan metode *Convolutional Neural Network (CNN)* pemanfaatan model pelatihan *Word2Vec*, khususnya menggunakan pendekatan model skip-gram [8]. dengan menggunakan model *Word2Vec* untuk menganalisis komentar politik di media sosial X. Diharapkan dapat memberikan kontribusi bagi pengembangan ilmu

pengetahuan dan teknologi dalam bidang analisis sentimen khususnya pada domain politik.

Pengumpulan data yang dilakukan untuk mengetahui klasifikasi sentimen positif dan negatif terhadap tanggapan pengguna *X* terhadap komentar politik [9] dengan menerapkan metode *Convolutional Neural Network (CNN)*. Konsep *deep learning* mengacu pada cabang machine learning yang memanfaatkan model matematis yang kompleks guna memahami representasi tingkat tinggi dari data [4]. Dalam model *Word2Vec*, variabel menjadi proses pembentukan vector representatif yang mana menggabungkan serta menganalisis relasi yang terdapat di antara kata-kata dalam dimensi vector.

Implikasi hasil ini dapat memberikan pemahaman terkait performa model *deep learning* dalam menganalisis sentimen dari komentar politik di media sosial *X*. Hasil dari analisis ini dapat menjadi pedoman bagi peneliti atau praktisi yang tertarik menggunakan pendekatan *deep learning* dalam analisis sentimen teks, baik dalam ranah politik maupun bidang lain, dan memberikan wawasan tentang pandangan dan sikap masyarakat terhadap isu-isu politik yang tengah berkembang.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen adalah proses untuk mengidentifikasi dan mengekstraksi opini, emosi, atau sikap yang tersirat atau tersurat dalam teks dan diklasifikasikan menjadi opini positif, negatif atau netral. Dalam melakukan penelitian yang berfokus pada analisis sentimen, kita dapat menggambarkan respon atau tanggapan pengguna terhadap topik yang sedang menjadi fokus perbincangan utama [10]. Lingkup analisis sentimen mencakup beragam bidang, melibatkan teknik pengolahan bahasa alami, komputasi linguistik, dan text mining dengan maksud untuk menganalisis pandangan, sentimen, evaluasi, sikap, penilaian, dan ekspresi emosional seseorang, baik sebagai pembicara maupun penulis, yang terkait dengan berbagai topik, produk, layanan, organisasi, individu, atau kegiatan tertentu [11].

2.2. Deep learning

Artificial intelligence (AI), yang lebih umum dikenal sebagai kecerdasan buatan, merangkum bagian dari konsep machine learning yang memberikan kapabilitas kepada komputer untuk meraih pengetahuan dari data baru tanpa kebutuhan akan program yang secara eksplisit. Tujuan utama yang dikedepankan adalah pengembangan program komputer yang memiliki kemampuan untuk menganalisis data dan merancang model yang dapat diaplikasikan secara efektif dalam menyelesaikan masalah tertentu [12].

2.3. Convolution Neural Network (CNN)

CNN merupakan bagian dari salah satu jenis arsitektur *deep learning* yang sering digunakan dalam

pengolahan data gambar dan teks. *CNN* termasuk komponen dari jaringan saraf mendalam (*Deep Neural Network*). Nama *Convolutional Neural Network (CNN)* memberikan petunjuk bahwa jaringan saraf ini terlibat dalam operasi matematika *convolution*, yang mencerminkan suatu proses operasi linear. *Convolutional Neural Network*, pada prinsipnya, adalah suatu kategori *Neural Network* yang menggunakan operasi *convolution* sebagai pendekatan perkalian matriks, setidaknya terdapat pada satu lapisan tertentu. Struktur *CNN* melibatkan lapisan *convolution* dan *pooling*, memungkinkan proses ekstraksi fitur dari citra [13].

2.4. Word2Vec

Word2Vec merupakan salah satu teknik atau metode dalam pemrosesan bahasa alami (*Natural Language Processing* atau *NLP*). *Word2Vec* mencatat sebuah langkah maju dalam proses ekstraksi fitur kata secara semantik dari suatu korpus. *Word2Vec* menggambarkan setiap kata yang bersifat unik melalui kumpulan nilai numerik yang dikenal sebagai vektor. Seleksi vektor dilakukan melalui fungsi matematis, mencerminkan tingkat kemiripan semantik antara kata-kata yang diwakili oleh vektor tersebut [8].

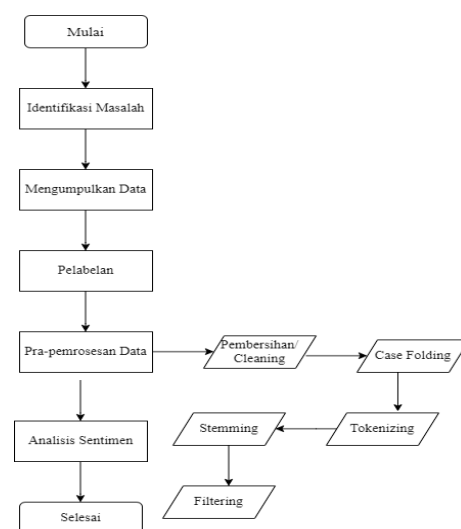
2.5. Natural Language Processing (NLP)

Pemrosesan Bahasa Alam atau *Natural Language Processing (NLP)* adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. Terbukti efektif tugas Pemrosesan bahasa alamiah *Natural Language Processing (NLP)* dan berhasil mencapai performa yang optimal dalam melakukan klasifikasi kalimat [14].

3. METODE PENELITIAN

3.1. Tahapan Metode Penelitian

Berikut ini tahapan metode yang digunakan pada Gambar 1 dibawah ini:



Gambar 1. Tahapan Metode Penelitian

3.2. Pengumpulan Data

Pengumpulan data dilakukan melalui Google Collab dengan memakai bahasa pemrograman *Python*, menggunakan parameter pencarian “Politik” yang diambil pada rentang waktu September 2023 sampai sekarang. Dengan jumlah data komentar sebanyak 1300 yang belum diseleksi atau belum diolah. Dan data disimpan dalam bentuk tampilan *file excel*. Berikut tampilan dari pengumpulan data dapat dilihat pada Gambar 2 dibawah ini:

id	created_at	id_user	full_text	quote_tweet	reply_tweet	retweet_tweet	favorited_tweet	lang	user_id_str	conversation_id_str	username
0	Thu Mar 13 13:46:12 +0800 2023	1731607349506261	@GinRomli Kebenaran akan Selalu menemukan jalannya... Fitnahmu terlalu tinggi hingga tidak menapak ke Bumi Satu persatu rakyat yang sudah jenuh dan Jengah dengan Politik Fitnah muncul ke permukaan bicara fakta dan data (Bukannya Katanya) https://t.co/2e3fZBEEF					id	17308116326412248	173288458225270	kinchawal20
1	Thu Mar 13 13:46:12 +0800 2023	1731607349506261	@GinRomli Kebenaran akan Selalu menemukan jalannya... Fitnahmu terlalu tinggi hingga tidak menapak ke Bumi Satu persatu rakyat yang sudah jenuh dan Jengah dengan Politik Fitnah muncul ke permukaan bicara fakta dan data (Bukannya Katanya) https://t.co/2e3fZBEEF					id	17308116326412248	173288458225270	kinchawal20
2	Thu Mar 13 13:46:12 +0800 2023	1731607349506261	@GinRomli Kebenaran akan Selalu menemukan jalannya... Fitnahmu terlalu tinggi hingga tidak menapak ke Bumi Satu persatu rakyat yang sudah jenuh dan Jengah dengan Politik Fitnah muncul ke permukaan bicara fakta dan data (Bukannya Katanya) https://t.co/2e3fZBEEF					id	17308116326412248	173288458225270	kinchawal20
3	Thu Mar 13 13:46:12 +0800 2023	1731607349506261	@GinRomli Kebenaran akan Selalu menemukan jalannya... Fitnahmu terlalu tinggi hingga tidak menapak ke Bumi Satu persatu rakyat yang sudah jenuh dan Jengah dengan Politik Fitnah muncul ke permukaan bicara fakta dan data (Bukannya Katanya) https://t.co/2e3fZBEEF					id	17308116326412248	173288458225270	kinchawal20

Gambar 2. Pengumpulan Data

3.3. Labeling Sentimen

Pada tahap ini dilakukan proses labeling manual berupa sentimen positif dan negatif yang di simpan kedalam *file excel*, berikut merupakan tampilan labeling ditampilkan pada Gambar 3 dibawah ini:

A	B
Negatif	Kebenaran akan Selalu menemukan jalannya... Fitnahmu terlalu tinggi hingga tidak menapak ke Bumi Satu persatu rakyat yang sudah jenuh dan Jengah dengan Politik Fitnah muncul ke permukaan bicara fakta dan data (Bukannya Katanya) https://t.co/2e3fZBEEF
Negatif	@GinRomli Ngapa emangnya dgn politik identitas ?? Kok yg memalsukan identitas ngga loe komenin?
Positif	@Nurhadi0842723 Andai bersyukur dan gak balelo mungkin lebih cemerlang karier politik pasca pensiun ya ... ☺
Positif	Ketua DPP PDIP, Puan Maharani, meihat tak ada cawe-cawe Presiden Joko Widodo (Jokowi) ke partai politik (parpol) lain. Puan meyakini setiap ketua umum parpol memiliki strategi kemenangan masing-masing. https://t.co/uxv6lcau57
Positif	@TeddiGus Yang sibuk menetralsir ucapan Jokowi kemarin ya. Itu blom apa2 bro. Saya mau nunggu deklarasi capres lain selain Ganjar yg dihadiri JKW nanti. Sebenarmyalakian udah tau arahnya pakde kemana tapi kalian seolah2 TDK mau tau utk memaksakan syahwat politik pribadi kalian aja
Positif	@OposisiCerdas Semua partai politik pasti berkata atas kebutuhan RAKYAT™
Negatif	@strawbaeri Problematik anjir perasaan dulu cuma ngurusin politik aja
Positif	Survei Indikator: PDIP Masih Unggul Karena Figur Jokowi Survei terbaru lembaga Indikator Politik Indonesia menyebut PDIP masih mendominasi paling atas. Berkat pengaruh Jokowi 22 persen, dan pengaruh Megawati hanya 2,4 persen. https://t.co/EV14rtUu2

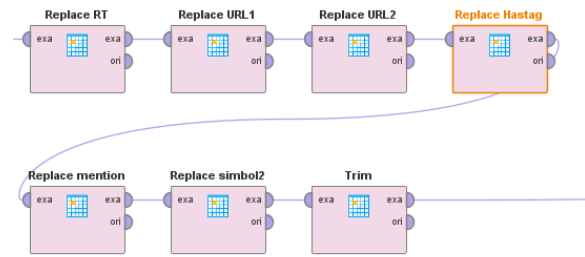
Gambar 3 Labeling

3.4. Pemrosesan Data

Dalam pemrosesan data ini di bagi menjadi beberapa tahapan, berikut tahapan pemrosesan data:

1. Pembersihan/Cleaning

Pada tahapan cleaning ini dilakukan pengerjaan dua kali, tahapan *cleaning* pertama menggunakan tool Rapid Miner versi 10.3, dari data yang semula 1300 dan setelah data di *cleaning* menjadi 1253 data. Dibagian *cleaning* pertama dilakukan menghilangkan bagian seperti RT(retweet), URL, hastag, mention, simbol-simbol, dan duplikat. Dan tahapan cleaning kedua ini menggunakan Google Colab dengan memakai bahasa pemrograman *Python* dengan menghilangkan emoji/ikon ekspresi. Berikut merupakan tampilan cleaning pertama dan kedua ditampilkan pada Gambar 4 dan 5 dibawah ini:



Gambar 4. Cleaning Pertama

full_text	Sentimen	test_clean
Pesanten selama ini telah mjd institusi yg berakar pada madyarakat dgn tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *	Positif	pesanten selama ini telah mjd institusi yg berakar pada madyarakat dgn tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *
Ada makhluk yg betarik mengabdikan Tan Malaka tapi jdi GIGILO POLITIK Benar-benar bangke 🤔🤔🤔 Sehat selalu mas 🙏	Negatif	ada makhluk yg betarik mengabdikan tan malaka tapi jdi gigilo politik benar-benar bangke sehat selalu mas 🙏
Pesanten selama ini telah mjd institusi yg berakar pada msykt dg n tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *	Positif	pesanten selama ini telah mjd institusi yg berakar pada msykt dg n tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *
Pesanten selama ini telah mjd institusi yg berakar pada msykt dg n tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *	Positif	pesanten selama ini telah mjd institusi yg berakar pada msykt dg n tepo menyarkan nilai-nilai keagamaan yg rahmatan li alamin Pesanten hrs mengedepankan politik kebangsaan bkn dukungan/dukung n dijadikan slogan kampanye *
Kampanye politik bukan janji yg hrs ditepati	Positif	kampanye politik bukan janji yg hrs ditepati

Gambar 5. Cleaning Kedua

2. Case Folding

Case Folding merupakan metode untuk menyederhanakan data teks, menormalisasi teks dengan merubah seluruh teks yang semula menjadi huruf acak menjadi huruf kecil hanya huruf A-Z. Berikut ini tampilan hasil dari proses *case folding* yang telah dilakukan melalui Tabel 1 dibawah ini:

Tabel 1. Case Folding

No.	Teks	Sesudah
1.	Lihat rekam jejak Anies di DKI paling banyak umbar janji manis tanpa realisasi Menghalalkan segala cara demi meraih kemenangan Anies otak dan pelaku politik SARA	lihat rekam jejak anies di dki paling banyak umbar janji manis tanpa realisasi menghalalkan segala cara demi meraih kemenangan anies otak dan pelaku politik sara

3. Tokenizing

Tokenisasi adalah pemecahan suatu kalimat atau teks menjadi unit-unit kecil atau kata-kata terpisah yang disebut sebagai token yang merupakan tahapan dalam tokenisasi. Berikut ini tampilan hasil dari proses tokenisasi yang telah dilakukan melalui Tabel 2 dibawah ini:

Tabel 2 Tokenizing

No.	Teks	Sesudah
1.	Lihat rekam jejak Anies di DKI paling banyak umbar janji manis tanpa realisasi Menghalalkan segala cara demi meraih kemenangan Anies otak dan pelaku politik SARA	[lihat, rekam, jejak, anies, dki, umbar, janji, mani, realisasi, menghalalkan, meraih, kemenangan, ani, otak, pelaku, politik, sara]

4. Stemming

Stemming adalah mengubah kata-kata menjadi kalimat bentuk dasarnya, proses ini menjadi bagian penting dalam pemrosesan teks dalam penelitian.

Dan menghasilkan representasi kata-kata yang lebih seragam. Berikut ini tampilan hasil dari proses *stemming* yang telah dilakukan melalui Tabel 3 dibawah ini:

Tabel. 3 *Stemming*

No.	Teks	Sesudah
1.	Lihat rekam jejak Anies di DKI paling banyak umbar janji manis tanpa realisasi Menghalalkan segala cara demi meraih kemenangan Anies otak dan pelaku politik SARA	lihat rekam jejak anies dki umbar janji manis realisasi menghalalkan meraih kemenangan anies otak pelaku politik sara

3.5. Filtering/Stopword Removal

Stopword Removal adalah tahapan pemrosesan teks yang bertujuan untuk menghilangkan kata-kata yang umum dan sering muncul dalam sebuah bahasa, tetapi tidak memberikan pengaruh signifikan terhadap makna atau inti informasi. Berikut ini tampilan hasil dari proses *Stopword Removal* yang telah dilakukan melalui Tabel 4 dibawah ini:

Tabel. 4 *Stopword Removal*

No.	Teks	Sesudah
1.	Lihat rekam jejak Anies di DKI paling banyak umbar janji manis tanpa realisasi Menghalalkan segala cara demi meraih kemenangan Anies otak dan pelaku politik SARA	lihat rekam jejak anies dki umbar janji manis realisasi menghalalkan meraih kemenangan anies otak pelaku politik sara

3.6. Representasi Data

```
from sklearn.feature_extraction.text import TfidfVectorizer
df['text_clean'] = df['text_clean'].astype(str)
tfidf = TfidfVectorizer()
tweet = df['text_clean'].values.tolist()
tfidf_vect = tfidf.fit(tweet)
x = tfidf_vect.transform(tweet)
y = df['Sentimen']
print(x[0:2])
```

(0, 6686)	0.15073519691792012	(0, 903)	0.20880957187399093
(0, 6280)	0.18823731005985578	(0, 555)	0.1762033083435689
(0, 6123)	0.15105593141222087	(0, 136)	0.1762033083435689
(0, 5616)	0.15732747967866217	(0, 92)	0.17307626090865888
(0, 5503)	0.20880957187399093	(1, 6686)	0.10727062578201049
(0, 5229)	0.1762033083435689	(1, 6080)	0.15259417581723703
(0, 5049)	0.031209334438194517	(1, 6062)	0.31432701816698827
(0, 4944)	0.34615252181731776	(1, 5615)	0.2215099221341267
(0, 4518)	0.11798226451960192	(1, 5574)	0.2463397241879507
(0, 4332)	0.17024750241770908	(1, 5049)	0.04442029656309566
(0, 4014)	0.22084357359027781	(1, 3802)	0.31432701816698827
(0, 3861)	0.1762033083435689	(1, 3428)	0.21499815333241257
(0, 3757)	0.1762033083435689	(1, 3364)	0.31432701816698827
(0, 3307)	0.22084357359027781	(1, 3349)	0.31432701816698827
(0, 3233)	0.1762033083435689	(1, 2309)	0.1417236076108077
(0, 2609)	0.17024750241770908	(1, 1878)	0.31432701816698827
(0, 2600)	0.1762033083435689	(1, 730)	0.31432701816698827
(0, 2522)	0.14835593451194667	(1, 531)	0.2971990039141733
(0, 2245)	0.17307626090865888	(1, 426)	0.2971990039141733
(0, 2229)	0.22084357359027781	(1, 19)	0.13308565117984347
(0, 2114)	0.1762033083435689		
(0, 1584)	0.1762033083435689		
(0, 1526)	0.1836621949426429		
(0, 1305)	0.15563104652943374		
(0, 1194)	0.22084357359027781		

Gambar 6. Metode *TF-IDF*

Pada Representasi Data menggunakan metode *TF-IDF* (*Term Frequency-Inverse Document Frequency*), *TF-IDF* adalah metode statistik yang diterapkan dalam pengolahan data dan mengubah teks ke bentuk angka, langkah ini dilakukan untuk mendukung penerapan machine learning dan arsitektur *deep learning*. Berikut ini di tampilkan hasil dari proses representasi data yang telah dilakukan menggunakan metode *TF-IDF* melalui Gambar 6

3.7. Pembagian Data

Pada tahap pembagian data ini peneliti data yang digunakan dibagi menjadi data *training* atau data *validation* dan data *testing*. Jumlah data yang sudah melalui tahap prosesing yang akan dipakai yaitu 1253. Dengan pembagian 90% (1127) data *training* dan 10% (126) untuk data *testing*. Berikut pembagian data dapat dilihat pada Gambar 6 sebagai berikut:

```
"train set has total (0) entries with (1.1f)% negative, (2.1f)% positive".format(len(x_train),
(len(x_train[y_train == 0]) / (len(x_train)*1.)) *100,
(len(x_train[y_train == 1]) / (len(x_train)*1.)) *100)

"train set has total 1127 entries with 0.1% negative, 0.9% positive"

"test set has total (0) entries with (1.1f)% negative, (2.1f)% positive".format(len(x_test),
(len(x_test[y_test == 0]) / (len(x_test)*1.)) *100,
(len(x_test[y_test == 1]) / (len(x_test)*1.)) *100)

"test set has total 126 entries with 0.5% negative, 0.5% positive"
```

Gambar 7. Pembagian Data

4. HASIL DAN PEMBAHASAN

4.1. Evaluasi Model

Pada tahap evaluasi model ini peneliti menggunakan *CNN* (*Convulation Neural Network*) dengan model *Word2Vec*, Kemudian pengukuran akurasi yang didapatkan menggunakan metode *K-Fold Cross Validation*, pembangunan model *CNN* diawali dengan pengaturan parameter awal, yang mencakup parameter *dropout* dan tingkat pembelajaran (*learning rate*). Rangkaian pertama pada penelitian ini yaitu menentukan Representasi Vektor terbaik untuk model pelatihan dan validasi, dapat dilihat pada Tabel 5 sebagai berikut:

Tabel 5. Representasi Vektor

No	Representasi Vektor	Akurasi	
		Training	Validasi
1.	Representasi Vektor Python Library keras		
	Epoch-1	0,6522	0,6746
	Epoch-2	0,8962	0,7302
	Epoch-3	0,9814	0,7857
	Epoch-4	0,9973	0,7381
	Epoch-5	0,9991	0,7619
2.	Penyempurnaan Representasi Vektor		
	Epoch-1	0,6486	0,7381
	Epoch-2	0,8385	0,7540
	Epoch-3	0,9193	0,7778
	Epoch-4	0,9645	0,7937
	Epoch-5	0,9867	0,7778

Selanjutnya dalam rangkaian kedua yaitu mengkonkatenasikan hasil *Max Pooling Bigram*, *Tigram* dan *Fourgram* pada model *CNN* serta menggabungkan hasilnya. Ini berpengaruh terhadap

hasil nilai akurasi. Berikutnya rangkaian ketiga yaitu mencari nilai *Learning Rate* dan *Dropout* terbaik dapat dilihat pada Tabel 6, 7 dan 8 sebagai berikut:

Tabel 6. *Learning Rate 0,0001 dan Dropout 0,2*

No.	Kombinasi Parameter	Akurasi	
		Training	Validasi
1.	Learning Rate=0,0001 ; Dropout=0,2		
	<i>K-Fold</i> = 1	1	0,7920
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1
2.	Learning Rate=0,001 ; Dropout=0,2		
	<i>K-Fold</i> = 1	1	0,8009
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1
3.	Learning Rate=0,01 ; Dropout=0,2		
	<i>K-Fold</i> = 1	1	0,7832
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1

Tabel 7. *Learning Rate 0,0001 dan Dropout 0,3*

No.	Kombinasi Parameter	Akurasi	
		Training	Validasi
1.	Learning Rate=0,0001; Dropout=0,3		
	<i>K-Fold</i> = 1	0,9989	0,7965
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1
2.	Learning Rate=0,001 ; Dropout=0,3		
	<i>K-Fold</i> = 1	0,9978	0,7965
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1
3.	Learning Rate=0,01 ; Dropout=0,3		
	<i>K-Fold</i> = 1	1	0,7832
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1

Tabel 8. *Learning Rate 0,0001 dan Dropout 0,4*

No.	Kombinasi Parameter	Akurasi	
		Training	Validasi
1.	Learning Rate=0,0001 ; Dropout=0,4		
	<i>K-Fold</i> = 1	0,9989	0,7876
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1

No.	Kombinasi Parameter	Akurasi	
		Training	Validasi
2.	Learning Rate=0,001 ; Dropout=0,4		
	<i>K-Fold</i> = 1	0,9989	0,8053
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1
3.	Learning Rate=0,01 ; Dropout=0,4		
	<i>K-Fold</i> = 1	1	1
	<i>K-Fold</i> = 2	1	1
	<i>K-Fold</i> = 3	1	1
	<i>K-Fold</i> = 4	1	1

Model *CNN* yang diterapkan merupakan hasil dari penggabungan parameter, di mana akurasi dari *training* dan akurasi validasinya pada Tabel diatas yang menunjukkan kombinasi parameter *learning rate* dan juga *dropout* yang terbaik, Dimana nilai *learning rate* yaitu pada nilai 0,01 dan *dropout* pada nilai 0,2 untuk akurasi terbaik yaitu dilihat pada *k-fold* nya. Maka kombinasi parameter *learning rate* dan *dropout* terbaik yang dilakukan penggabungan (konkatenasi) antara *bigram*, *trigram*, dan *fourgram* untuk memperkuat representasi dan pemahaman yang lebih baik. Tingkat akurasi dari data dengan menerapkan model *CNN best weights* adalah 81%. Berikut tampilan dari akurasi dapat dilihat pada Gambar 8 sebagai berikut:

```
#load model sesuai nama file diatas dengan model bobot terbaik
from keras.models import load_model
loaded_CNN_model = load_model('CNN_best_weights.01-1.0000.hdf5')
loaded_CNN_model.evaluate(x=x_te, y=y_te)

8/8 [=====] - 1s 32ms/step - loss: 0.0091 - accuracy: 1.0000
[0.009145944379270077, 1.0]

I Model Evaluation dengan Data testing

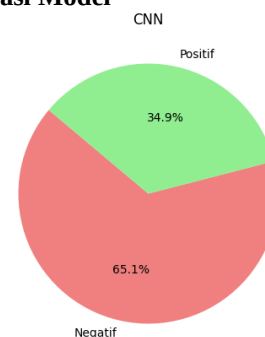
sequences_test = tokenizer.texts_to_sequences(x_test)
x_test_seq = pad_sequences(sequences_test, maxlen=55)

#load CNN model
loaded_CNN_model.evaluate(x=x_test_seq, y=y_test)

4/4 [=====] - 1s 35ms/step - loss: 0.5294 - accuracy: 0.8175
[0.5294057726860046, 0.817460298538200]
```

Gambar 8. Nilai Akurasi

4.2. Visualisasi Model

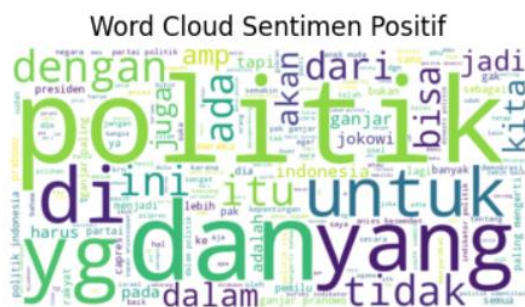


Gambar 9. Diagram Polaritas Presentasi Sentimen

Visualisasi data dalam penelitian ini yaitu proses menggambarkan informasi atau hasil temuan dalam bentuk yaitu diagram dan wordcloud sentiment positif

dan negatif. Berikut ini di tampilkan hasil dari visualisasi data yang telah dilakukan menggunakan *Google Colab* melalui Gambar 9.

Dari hasil analisis sentimen termanifestasi dalam bentuk polaritas positif dan negatif, dengan kemampuan perhitungan persentase label untuk masing-masing sentimen seiring dengan 35,5% untuk sentimen positif dan 64,5% untuk sentimen negatif. Untuk mengidentifikasi kata-kata yang mendominasi kemunculan di setiap sentimen, metode wordcloud diadopsi untuk menampilkan sejumlah kata yang terkandung dalam kolom sentimen. Dengan demikian, di bawah ini disajikan representasi visual dari wordcloud yang merepresentasikan kedua sentimen pada Gambar 10 dan 11 sebagai berikut:



Gambar 10. Wordcloud Sentimen Positif



Gambar 11. Wordcloud Sentimen Negatif

4.3. Penerapan Metode CNN

Tahapan awal dimulai dengan pengumpulan data, dalam pengumpulan data terdapat beragam metode yang dapat digunakan dalam proses pengumpulan data, salah satunya pengambilan data menggunakan teknik *crawling* dengan penerapan bahasa pemrograman *python*. dengan menentukan parameter pencarian “Politik” Kemudian dilakukan tahapan labeling sentimen pada tahap proses pelabelan (*labeling*), *labeling* dalam penelitian ini dilakukan secara manual. Kemudian dilakukan pemrosesan data yang dilakukan menggunakan dua *tools* dan dilanjut pada tahap representasi data yang menggunakan *TF-IDF*, kemudin dilakukan pembagian data yaitu data *testing* dan data *training*, selanjutnya dilakukaan evaluasi model dengan gambar diagram dan *wordcloud* berupa kata positif dan negatif. Dalam tahap penerapan metode ini menggunakan metode *CNN* dengan parameter *dropout* dan *learning rate*

yang dapat mempengaruhi kinerja dari model klasifikasi *tweet* berbahasa Indonesia yang mengandung unsur politik dengan nilai *learning rate* 0,01 dan nilai *dropout* 0,2 dengan menggunakan model *Word2vec* memperoleh hasil akurasi 81%.

4.4. Evaluasi Algoritma CNN

Pada tahap evaluasi model ini peneliti menggunakan *CNN (Convolution Neural Network)* dengan model *Word2vec*. Dalam pengujian terhadap testing data, algoritma *Deep Convolutional Neural Network (CNN)*. Kemudian proses validasi hasil yang telah diperoleh dengan mengavaluasi pengukuran akurasi yang didapatkan menggunakan metode *K-Fold Cross Validation* pada algoritma *CNN*, bertujuan untuk mengukur efektivitas model klasifikasi yang digunakan. Perbandingan dilakukan antara label data asli dan label hasil prediksi untuk melibatkan proses evaluasi. Pada pengujian terhadap data training, akan dievaluasi akurasi dari data tersebut dengan parameter yang sudah ditentukan.

4.5. Implementasi Algoritma CNN

Pada tahap implementasi algoritma *CNN* ini dilakukan pengujian terhadap *dataset* komentar *X* yang belum dilabeli, data uji yang berisi komentar *X* yang belum dilabeli akan dianalisis sentimennya dengan menggunakan *source code*. *Dataset* yang sudah melalui evaluasi dari pembobotan hasil training data yang berisi komentar *X* yang telah melalui tahap pemrosesan data dan sudah diberikan label polaritas sentimen positif dan negatif. disimpan dalam *file joblib*, data yang sudah melalui evaluasi diterapkan pada model baru, Selanjutnya model baru mengalami tahapan prediksi dengan menggunakan fungsi yang telah dirancang, dan hasil prediksi sentimen positif atau negatif dipaparkan. Misalkan ada *tweet* seperti "Bodoh politik riang gembira lha kompor model koq dipercaya mencla mencl koyo dele." Setelah diterapkan pada model yang sudah dievaluasi maka hasil prediksinya negatif dan ada *tweet* seperti ini "Di tahun segini ada yang percaya omongan tokoh politik tuh ko bisasih." Hasil prediksi menunjukkan bahwa *tweet* tersebut ber label negatif. Dengan rangkaian ini menjelaskan secara teliti implementasi model *Convolutional Neural Network (CNN)* untuk memprediksi sentimen pada teks-*tweet* yang baru muncul, sehingga mampu memberikan pemahaman mendalam terkait dengan opini positif atau negatif terhadap suatu topik.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang di lakukan dapat disimpulkan bahwa penerapan pada metode *CNN* ini berhasil melakukan analisis sentimen komentar Politik di media sosial *X* dengan menggunakan metode *Convolutional Neural Networks (CNN)* mendapatkan hasil akurasi 81%. Pada tahap evaluasi algoritma *CNN* ini dapat disimpulkan bahwa dengan menggunakan metode *CNN (Convolution Neural Network)* dengan model *Word2vec* ini berhasil dengan

data yang digunakan dibagi menjadi data *training*/ data *validation*. Pada tahap implementasi algoritma CNN dapat disimpulkan bahwa implementasi ini dikatakan berhasil yang dilakukan pengujian terhadap *dataset* komentar *X* yang sudah dievaluasi akan diterapkan pada model baru yang datanya yang belum dilabeli, model baru yang berisi komentar *X* yang belum dilabeli akan dianalisis sentimennya dengan menggunakan *source code*.

Diharapkan bahwa penelitian selanjutnya akan memanfaatkan dua atau lebih metode untuk tujuan meningkatkan akurasi model. Model belum mampu mengidentifikasi kata-kata yang disingkat yang saling terhubung, sehingga mempengaruhi tingkat akurasi perhitungan. Oleh sebab itu, penelitian selanjutnya diperlukan penambahan fitur khusus guna mengulangi permasalahan ini.

DAFTAR PUSTAKA

- [1] D. T. Hermanto, A. Setyanto, dan E. T. Luthfi, "Algoritma LSTM-CNN untuk Binary Klasifikasi dengan Word2vec pada Media Online," *Creat. Inf. Technol. J. (Citec Journal)*, 2021, [Daring]. Tersedia pada: <https://citec.amikom.ac.id/main/index.php/citec/article/view/264>
- [2] E. Tohidi, A. Faqih, dan R. Narasati, "Pengembangan Media Pembelajaran Berbasis Permainan Edukasi Dan Interaktif Pada Matakuliah Akuntansi Keuangan," *J. Edukasi (Ekonomi, Pendidik. dan Akuntansi)*, vol. 9, no. 2, hal. 139, 2021, doi: 10.25157/je.v9i2.6275.
- [3] S. Alam, "Penggunaan Media Sosial Sebagai Alat Komunikasi Politik," *Avant Garde*, vol. 9, no. 1, hal. 67, 2021, doi: 10.36080/ag.v9i1.1257.
- [4] H. Juwiantho dkk., "Sentiment Analysis Twitter Bahasa Indonesia Berbasis Word2vec Menggunakan Deep Convolutional Neural Network Sentiment Analysis Twitter Indonesian Language Word2vec-Based Using Deep Convolutional Neural Network.," vol. 7, no. 1, hal. 181–188, 2020, doi: 10.25126/jtiik.202071758.
- [5] G. T. Pangestu dan M. Rosyda, "Sentiment Analysis Tweet Pilkada 2020 Saat Pandemi COVID-19 di Media Sosial Twitter Menggunakan Metode 1D Convolutional Neural Network," *J. Media Inform. Budidarma*, 2022, [Daring]. Tersedia pada: <http://www.stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/3765>
- [6] S. Ahmad, A. M. Ridwan, dan G. D. Setyawan, "Analisis Sentimen Product Tools & Home Menggunakan Metode CNN Dan LSTM," *TEKNOKOM*, 2023, [Daring]. Tersedia pada: <https://teknokom.unwir.ac.id/index.php/teknokom/article/view/154>
- [7] M. R. F. Kamarula dan N. Rochmawati, "Perbandingan CNN dan Bi-LSTM pada Analisis Sentimen dan Emosi Masyarakat Indonesia Di Media Sosial Twitter Selama Pandemi Covid-19 yang Menggunakan Metode Word2Vec," *J. Informatics Comput. Sci.*, 2022, [Daring]. Tersedia pada: <https://ejournal.unesa.ac.id/index.php/jinacs/article/view/50063>
- [8] S. Khomsah, "Sentiment Analysis On YouTube Comments Using Word2Vec and Random Forest Sentimen Analisis pada Opini YouTube Menggunakan Word2Vec dan Random Forest," *J. Inform. dan Teknol. Inf.*, vol. 18, no. 1, hal. 61–72, 2021, doi: 10.31515/telematika.v18i1.4493.
- [9] N. Amalia, T. Suprpti, dan G. Dwilestari, "Analisis Sentimen Pengguna Twitter Terhadap Pelaksanaan Kurikulum MBKM," *E-Link J. Tek. Elektro dan Inform.*, vol. 18, no. 1, hal. 57, 2023, doi: 10.30587/e-link.v18i1.5335.
- [10] A. B. Sasmita, B. Rahayudi, dan L. Muflikhah, "Analisis Sentimen Komentar pada Media Sosial Twitter tentang PPKM Covid-19 di Indonesia dengan Metode Naïve Bayes," vol. 6, no. 3, hal. 1208–1214, 2022.
- [11] A. M. Zuhdi, E. Utami, dan S. Raharjo, "Analisis sentiment twitter terhadap capres Indonesia 2019 dengan metode K-NN," *J. Inf.*, 2019, [Daring]. Tersedia pada: <https://informa.poltekindonesia.ac.id/index.php/informa/article/view/73>
- [12] E. F. S. Hutagalung dan P. Sitompul, "Implementasi Deep Learning Menggunakan Metode Cnn Untuk Klasifikasi Jenis Ulos Batak Toba," *Student Sci. Creat. J.*, 2023, [Daring]. Tersedia pada: <https://journal.amikveteran.ac.id/index.php/sscj/article/view/1541>
- [13] J. E. Widyaya dan S. Budi, "Pengaruh Preprocessing Terhadap Klasifikasi Diabetic Retinopathy dengan Pendekatan Transfer Learning Convolutional Neural Network," *J. Tek. Inform. Dan Sist. Inf.*, 2021, [Daring]. Tersedia pada: <https://journal.maranatha.edu/index.php/jutisi/article/view/3327>
- [14] S. N. Listyarini dan D. A. Anggoro, "Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network (CNN)," *J. Pendidik. dan Teknol. Indones.*, vol. 1, no. 7, hal. 261–268, 2021, doi: 10.52436/1.jpti.60.
- [15] FokusTekno. (15 Agustus 2023). *Twitter: Kenapa Jadi "X"? Ini Penjelasan dan Perubahan Fiturnya*. Fokus. Diakses pada 16 Desember 2023, dari <https://www.fokus.co.id/teknol/twitter-kenapa-jadi-x-ini-penjelasan-dan-perubahan-fiturnya/>