

## Perbandingan Kinerja Algoritma Convolutional Neural Network (CNN) dan Recurrent Neural Network (RNN) pada Analisis Sentimen Pemilu Presiden 2024

Ghaita Zahira Rahma<sup>1</sup>, Miftahurrahma Rosyda<sup>\*2</sup>

<sup>1,2</sup>Informatika, Universitas Ahmad Dahlan, Indonesia

Email: <sup>1</sup>[ghaita2100018206@webmail.uad.ac.id](mailto:ghaita2100018206@webmail.uad.ac.id), <sup>2</sup>[miftahurrahma.rosyda@tif.uad.ac.id](mailto:miftahurrahma.rosyda@tif.uad.ac.id)

### Abstrak

Penelitian ini mengevaluasi kinerja algoritma *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN) dalam analisis sentimen opini masyarakat terkait Pemilihan Umum Presiden 2024 di Indonesia. Pentingnya penelitian ini terletak pada upaya memahami pola sentimen masyarakat terhadap isu politik melalui pendekatan pembelajaran mendalam. Data dikumpulkan dari opini pengguna Twitter menggunakan kata kunci tertentu selama periode November 2023 hingga April 2024. Metode yang digunakan meliputi *crawling* data, *preprocessing*, pembobotan TF-IDF, serta penyeimbangan data dengan metode *SMOTETomek*. Model dievaluasi menggunakan *K-Fold Cross Validation*. Hasil penelitian menunjukkan bahwa CNN dengan *resampling* mencapai rata-rata akurasi validasi 99%, sedangkan RNN hanya mencapai 77,13%. Tanpa *resampling*, CNN memiliki rata-rata akurasi 94% dibandingkan dengan RNN sebesar 87,04%. CNN terbukti unggul dalam presisi dan efisiensi waktu, terutama pada dataset tidak seimbang. Temuan ini memberikan wawasan penting tentang pemanfaatan algoritma pembelajaran mendalam untuk analisis sentimen politik dan dapat menjadi dasar bagi pengembangan metode serupa di masa depan.

**Kata kunci:** Analisis Sentimen, Convolutional Neural Network, Pemilu 2024, Recurrent Neural Network, Twitter.

### *Performance Comparison of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) Algorithms in the Sentiment Analysis of the 2024 Presidential Election*

#### Abstract

*This study evaluates the performance of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) algorithms in analyzing public sentiment regarding Indonesia's 2024 Presidential Election. The significance of this research lies in its effort to understand public sentiment patterns on political issues through deep learning approaches. Data was collected from Twitter users' opinions using specific keywords during the period from November 2023 to April 2024. The methods employed include data crawling, preprocessing, TF-IDF weighting, and data balancing using the SMOTETomek method. The models were evaluated using K-Fold Cross Validation. The results showed that CNN with resampling achieved an average validation accuracy of 99%, while RNN only reached 77.13%. Without resampling, CNN achieved an average accuracy of 94%, compared to RNN's 87.04%. CNN demonstrated superiority in precision and time efficiency, especially on imbalanced datasets. These findings provide valuable insights into the use of deep learning algorithms for political sentiment analysis and could serve as a foundation for the development of similar methods in the future.*

**Keywords:** 2024 Election, Convolutional Neural Network, Sentiment Analysis, Recurrent Neural Network, Twitter.

## 1. PENDAHULUAN

Indonesia sebagai negara demokratis menegaskan pentingnya partisipasi aktif rakyat dalam proses pengambilan keputusan politik. Sistem pemerintahan berbasis suara rakyat menjadi fondasi utama untuk menjaga keseimbangan kekuasaan di negara ini. Pemilihan Umum Presiden, yang diadakan setiap lima tahun sekali secara serentak di seluruh Indonesia, merupakan salah satu mekanisme untuk mewujudkan partisipasi tersebut. Pada periode 2024-2029, Komisi Pemilihan Umum (KPU) telah menetapkan pemungutan suara dilaksanakan pada tanggal 14-15 Februari 2024, sebagaimana diumumkan melalui konferensi pers oleh Ketua KPU, Hasyim Asy'ari, pada 13 November 2023 [1]. Komisi Pemilihan Umum telah menetapkan tiga pasangan kandidat untuk Pemilu Presiden dan Wakil Presiden 2024, yaitu pasangan nomor urut 1 Anies Baswedan, nomor

urut 2 Prabowo Subianto - Gibran Rakabuming Raka, dan pasangan nomor urut 3 Ganjar Pranowo - Mahfud M.D [2].

Hasil Pemilu Presiden 2024 telah menetapkan bahwa Prabowo Subianto berhasil meraih suara terbanyak dalam Pilpres 2024. Berdasarkan hasil rapat Pleno KPU pada 20 Maret 2024, pasangan Prabowo Subianto-Gibran Rakabuming Raka memperoleh 96.214.691 suara (58,47 persen). Sementara itu, pasangan Anies Baswedan-Muhaimin Iskandar mendapatkan 40.971.906 suara (25,23 persen), dan pasangan Ganjar Pranowo-Mahfud MD memperoleh 27.040.878 suara (16,30 persen)[3]. Penetapan ini disambut berbagai tanggapan dari masyarakat yang disampaikan melalui media sosial, seperti Twitter atau X. Sebagai platform digital populer, Twitter menjadi wadah ekspresi opini masyarakat, termasuk terkait isu politik. Berdasarkan laporan, pengguna Twitter di Indonesia mencapai 18,45 juta pada Januari 2022, menempatkan Indonesia pada peringkat kelima dunia dalam jumlah pengguna platform ini [4]. Opini yang muncul di media sosial dapat memengaruhi persepsi publik terhadap kandidat, sehingga penting untuk menganalisis sentimen yang terkandung dalam opini tersebut.

Analisis sentimen merupakan proses untuk mengumpulkan dan mengevaluasi pandangan, ide, serta persepsi masyarakat tentang berbagai topik, produk, layanan, maupun isu tertentu. Informasi dari opini publik ini dapat dimanfaatkan oleh perusahaan, pemerintah, maupun individu sebagai dasar dalam pengambilan keputusan dan penyusunan strategi[5]. Dalam konteks ini, pendekatan Pengolahan Bahasa Alami (NLP) digunakan untuk membantu mesin memahami teks manusia [6]. Metode ini melibatkan transformasi bahasa alami ke dalam format yang dapat diproses oleh komputer [7]. Dua algoritma yang sering digunakan untuk analisis sentimen adalah *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN). CNN, yang awalnya digunakan untuk pengolahan gambar, kini juga diterapkan dalam NLP untuk mengenali pola pada teks [8]. Sementara itu, RNN, dengan memori internalnya, cocok untuk memproses data berurutan seperti teks [9].

Penelitian sebelumnya telah menunjukkan bahwa CNN lebih efektif dalam waktu eksekusi dibandingkan dengan RNN dalam beberapa kasus analisis sentiment tertentu [10]. Namun, beberapa penelitian masih belum secara eksplisit mengevaluasi dampak metode *resampling*, seperti *SMOTETomek*, dalam meningkatkan akurasi model pada dataset yang tidak seimbang [11]. Berbeda dengan penelitian sebelumnya, yang belum secara eksplisit mengevaluasi dampak metode *resampling* seperti *SMOTETomek* terhadap peningkatan akurasi model pada dataset yang tidak seimbang, penelitian ini membahas penerapan teknik serupa untuk meningkatkan performa model analisis sentimen dalam menghadapi data yang tidak seimbang, terutama dalam konteks Pemilu Presiden 2024. Hal ini mendukung pentingnya eksplorasi metode *resampling* untuk meningkatkan keakuratan analisis sentiment[12]. Berdasarkan potensi yang telah dibuktikan oleh penelitian tersebut, penelitian ini bertujuan untuk membandingkan efektivitas metode CNN dan RNN dalam analisis sentimen terhadap opini masyarakat di Twitter terkait Pemilihan Umum Presiden 2024. Penelitian ini diharapkan memberikan kontribusi signifikan dalam pemilihan metode terbaik untuk analisis sentimen politik, serta memberikan wawasan baru bagi pengembangan teknologi dalam memahami opini publik di media sosial.

## 2. METODE PENELITIAN

Tahapan yang digunakan pada metode penelitian ini dijelaskan dalam Gambar 1. Berikut penjelasan tahapannya.

### 2.1. Preprocessing

Tahapan *preprocessing* yang dilakukan pada penelitian ini mencakup beberapa *step* seperti berikut.

#### 1. *Cleaning Character*

Proses *Cleaning Character* ini digunakan sebagai alat atau program untuk penghapusan nama pengguna, kata 'RT', *hashtag*, URL, simbol, emoji, dan spasi kosong dari data tweet. Pada tahapan ini menggunakan fungsi dari pustaka ekspresi reguler.

#### 2. *Case Folding*

Pada tahapan *Case Folding* ini akan mengubah semua huruf dalam tweet menjadi huruf kecil atau non-kapital.

#### 3. *Tokenizing*

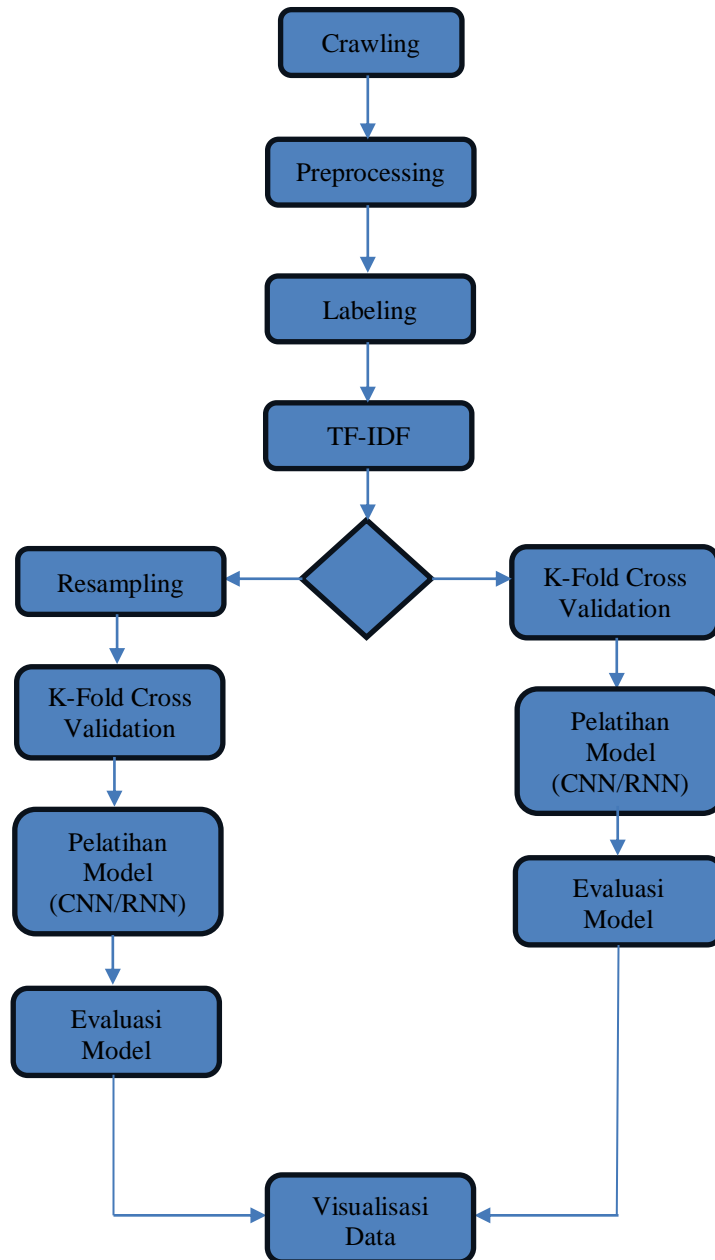
Proses *Tokenizing* menggunakan fungsi dari pustaka *Natural Language Toolkit* (NLTK) yang melakukan proses pemecahan string data pada tweet menjadi kata-kata yang terpisah atau token.

#### 4. *Stopwords Removal*

Tahapan *Stopwords Removal* merupakan tahapan untuk menghilangkan kata-kata yang tidak signifikan atau bisa dibilang kurang deskriptif. Pada tahapan ini juga menggunakan dan memanfaatkan fungsi dari pustaka *Natural Language Toolkit* (NLTK).

5. *Stemming*

*Stemming* merupakan tahapan dalam melakukan penghapusan kata-kata dengan imbuhan menjadi bentuk dasarnya.



Gambar 1. Tahapan Penelitian

2.2. Pelabelan Data

Dalam penelitian ini, proses pelabelan dilakukan secara otomatis menggunakan metode VADER (*Valence Aware Dictionary and sEntiment Reasoner*). Dikembangkan oleh C.J. Hutto dan Eric Gilbert pada tahun 2014, metode ini menggabungkan pendekatan berbasis manusia dengan validasi empiris yang didasarkan pada analisis kualitatif dan pengalaman manusia. Algoritma VADER digunakan untuk analisis sentimen karena kemampuannya dalam mengukur emosi dalam teks. Dalam proses penentuan polaritas, setiap kata dianalisis dan diberi label positif, negatif, atau netral. Penentuan label untuk teks dilakukan berdasarkan skor polaritas yang dihitung dengan cara menormalisasi total skor positif dan negatif[13].

### 2.3. Resampling

Teknik resampling (SMOTE-Tomek) merupakan kombinasi antara Teknik oversampling yang dilakukan dengan SMOTE dan Teknik undersampling Tomek Links[14]. Proses ini dilakukan dengan menggabungkan kemampuan dari SMOTE yang berupa peningkatan dari kelas minoritas, sedangkan data dari kelas mayoritas yang terlalu dekat dengan kelas minoritas dihapus menggunakan metode Tomek Links[15].

### 2.4. K-Fold Cross Validation

*K-Fold Cross Validation (K-Fold CV)* adalah teknik evaluasi model yang membagi dataset menjadi K bagian atau lipatan (*fold*) dengan jumlah yang sama. Metode ini bekerja dengan memanfaatkan setiap bagian secara bergantian sebagai data validasi, sedangkan bagian lainnya digunakan sebagai data latih. Proses ini dilakukan secara berulang hingga seluruh bagian data telah digunakan sebagai data validasi setidaknya satu kali. Teknik ini dirancang untuk memastikan evaluasi model yang lebih akurat dan mengurangi bias dengan memanfaatkan keseluruhan data secara bergantian untuk pelatihan dan validasi[16].

### 2.5. Pelatihan dengan Model

Convolutional Neural Network (CNN) merupakan jenis arsitektur jaringan saraf tiruan yang banyak digunakan untuk pengolahan gambar dan teks. CNN mampu secara otomatis mengekstraksi serta mempelajari fitur penting dari data masukan, sehingga efektif dalam tugas klasifikasi, deteksi, dan segmentasi. Sebagai bentuk khusus dari feed-forward neural network, CNN memiliki lapisan tersembunyi (*hidden layer*) yang terdiri dari operasi konvolusi dan pooling, yang berperan dalam proses analisis data untuk menghasilkan hasil yang optimal[17]. RNN (*Recurrent Neural Network*) merupakan jenis jaringan saraf tiruan yang mampu menangani data sekuensial, seperti teks. Keunggulan utamanya adalah kemampuannya memodelkan hubungan temporal dalam data. Namun, RNN menghadapi kelemahan, termasuk *vanishing* dan *exploding gradients*, terutama saat memproses ketergantungan jangka panjang dalam data. Untuk mengatasi hal ini, model seperti *Long Short-Term Memory (LSTM)* dan *Gated Recurrent Unit (GRU)* dirancang dengan mekanisme gates khusus yang membantu menyimpan dan mengatur informasi penting dalam data secara lebih efektif[18]. CNN lebih unggul dalam efisiensi waktu yang digunakan dibanding RNN karena tidak memiliki proses umpan balik atau *sequential* yang sering kali menyebabkan waktu komputasi menjadi lebih lama[6]. Kemudian, berdasarkan penelitian yang dilakukan oleh Husada dan Toba, penelitian yang menggunakan CNN cenderung memiliki kinerja yang lebih konsisten pada data yang tidak seimbang terutama ketika menerapkan teknik *resampling* sedangkan pada RNN cenderung kesulitan dalam menangani kategori dengan jumlah data lebih kecil meskipun menggunakan teknik *resampling*[19].

Pada penelitian ini, model CNN (*Convolutional Neural Network*) yang dibuat memiliki beberapa lapisan konvolusi dan pooling untuk mengekstraksi fitur dari data *input*, yang dalam hal ini kemungkinan besar berupa urutan kata pada teks. Lapisan pertama adalah *Conv1D* dengan 128 *filter* dan ukuran *kernel* 3, diikuti oleh *MaxPooling1D* untuk mengurangi dimensi fitur. Proses ini diulang dengan lapisan konvolusi kedua yang menggunakan 64 *filter*, diikuti dengan *pooling* lagi. Setelah itu, data diratakan menggunakan *Flatten*, dan melalui beberapa lapisan *Dense* untuk menghasilkan *output*. Fungsi *aktivasi* ReLU digunakan untuk memberikan non-linearitas pada model, sementara *softmax* digunakan pada lapisan terakhir untuk klasifikasi multi-kelas. *Dropout* diterapkan di beberapa titik untuk menghindari *overfitting*, dengan tingkat 50% di setiap lapisan *dropout*.

Sementara itu, model RNN (*Recurrent Neural Network*) yang menggunakan LSTM (*Long Short-Term Memory*) terdiri dari beberapa lapisan yang lebih berfokus pada pengolahan data sekuensial. Dimulai dengan lapisan *Embedding* untuk mengonversi kata menjadi representasi vektor, model kemudian dilanjutkan dengan dua lapisan LSTM yang masing-masing memiliki 128 unit dan 64 unit. Lapisan pertama menggunakan *return\_sequences=True*, sehingga urutan data diteruskan ke lapisan LSTM berikutnya. Seperti pada model CNN, *Dropout* diterapkan untuk mencegah *overfitting*, dan lapisan *Dense* dengan 64 neuron dan fungsi aktivasi ReLU ditambahkan sebelum lapisan terakhir. Sama halnya dengan CNN, lapisan terakhir pada model RNN menggunakan *softmax* untuk menghasilkan *output* klasifikasi multi-kelas. RNN lebih efektif untuk menangani ketergantungan temporal dalam data sekuensial, seperti teks atau urutan waktu.

Tabel 1. Tabel Perbandingan

Model	Arsitektur	Parameter Utama	F1 Score (Validation)	F1 Score (Test)
CNN	1D CNN + SMOTETomek	Filter: 64, 128, 256 Kernel size: 3x3	0.53	0.58

		Activation: ReLU Optimizer: Adam Loss: Categorical Crossentropy Epoch: 10 Batch Size: 32		
RNN	LSTM + SMOTETomek	Units: 64, 128 Activation: Tanh Optimizer: Adam Loss: Categorical Crossentropy Epoch: 10 Batch Size: 32	0.55	0.57

### 2.6. Evaluasi Model

Evaluasi ditampilkan dalam sebuah tabel dengan beberapa entitas yang disebut *confusion matrix*. Entitas-entitas ini mencakup akurasi, presisi, recall, dan F1-Score[20]. Evaluasi model dilakukan dengan menguji model sesuai dengan jenis model yang sedang dianalisis[21].

### 2.7. Visualisasi Data

Pada tahapan visualisasi data penelitian ini adalah menggunakan *World Cloud* yang merupakan teknik dalam melakukan visualisasi data berdasarkan *keywords* yang sering muncul dalam data yang telah tersedia. *World Cloud* sendiri akan memudahkan dalam memahami informasi dan gambaran visual mengenai topik yang sering dijadikan pembahasan oleh pengguna Twitter atau X dalam *keywords* yang telah ditentukan dalam tiap klasifikasi.

## 3. HASIL DAN PEMBAHASAN

Pada penelitian ini yang dimulai dengan melakukan pengumpulan data menggunakan teknik *crawling* dengan metode *Harvest-Crawling* dengan *keywords* “pilpres 2024”, “pemilu 2024”, dan “presiden 2024”. Setelah melakukan *crawling*, data yang sudah didapatkan akan melalui tahapan *preprocessing*, dalam tahapan ini terdapat beberapa proses tahapan lagi yaitu proses *cleaning character*, *case folding*, *tokenizing*, *stopwords removal* dan *stemming*. Kemudian pada tahapan selanjutnya yaitu proses *labelling* yang menggunakan VADER dan menghasilkan total label negatif sebanyak 46 opini, label positif sebanyak 80 opini, dan label netral sebanyak 1.155 opini. Setelah dilakukan proses *labelling*, data akan melalui tahapan pembobotan dengan TF-IDF dan dilakukan pembagian data 80% untuk data *training* dan 20% untuk data *testing*. Kemudian setelah melalui tahapan pembagian data, dilakukan penyeimbangan data menggunakan teknik *resampling*, yaitu gabungan dari teknik *oversampling SMOTE* dan juga *undersampling Tomek*. Setelah melalui tahapan penyeimbangan data yang memperoleh label negatif, positif dan netral sebanyak 1.155, data kemudian akan melalui tahapan pembagian lebih lanjut menggunakan *K-Fold Cross Validation* yang kemudian akan dilatih pada tiap Fold yang memiliki nilai K adalah 5. Pada tahapan selanjutnya, data akan melalui proses pelatihan model dengan algoritma CNN (*Convolutional Neural Network*) dan juga RNN (*Recurrent Neural Network*), pada tahapan ini dilakukan perbandingan untuk algoritma mana yang dirasa paling efektif dan akurat dalam melakukan analisis sentimen dengan proses penyeimbangan data maupun tidak menggunakan penyeimbangan data. Tahapan akhir dari penelitian ini adalah proses *world cloud*, dimana *keywords* data yang paling sering muncul akan mulai ditampilkan.

### 3.1. Studi Literatur

#### 3.1.1. Crawling Data

*Crawling* adalah teknik otomatis yang digunakan untuk mengumpulkan informasi dari web berdasarkan kata kunci yang ditentukan oleh pengguna. *Crawler*, yaitu program yang dirancang dengan algoritma tertentu, bertugas melakukan proses *crawling* dengan memindai halaman-halaman sesuai dengan alamat web dan kata kunci yang dapat diatur oleh pengguna[15].

Tahapan *Crawling* pada penelitian ini menggunakan metode *harvest-crawling* dengan pengambilan opini menggunakan *keywords*, “pilpres 2024, pemilu 2024, presiden 2024” dan memperoleh total 1.281 data *crawling*. Berikut adalah contoh data hasil tahapan *crawling* yang menggunakan teknik *harvest-crawling*.

Tabel 2. Hasil Data Crawling

Tanggal	Opini
Sun Jun 23 07:03:08 +0000 2024	Eh oon makan siang gratis buat anak sekolah bukan anak kuliah SD s/d SLTA Negeri itu setahuku sudah gratis semua. Kampanyenya bela 03 kok masih aja dipakai utk nolak kebijakan pemenang pilpres 2024 03 itu di banting 16% lol tolol <a href="https://t.co/K87K7aurVz">https://t.co/K87K7aurVz</a>
Sat Jun 22 23:24:24 +0000 2024	@langitan99 Inilah yg diinginkan oleh kaum Islam phobia bangsa Indonesia yg biadab seperti israel saling kriminal dan saling membunuh makannya Anies dikalahkan pilpres 2024 kalau wni sudah habis warga Singapore yg tempati Indonesia
Sat Jun 22 13:47:54 +0000 2024	Menjalankan Aturan seenaknya. Kemaren PilPres WalKot ikut Kontestasi ga ada ancaman apa?. Krn anak Presiden kan? Teros Pejabat yg kek gini tereak Profesional?  Tito Karnavian Ancam Pecat Penjabat Kepala Daerah yang Tak Mundur saat Maju Pilkada 2024 <a href="https://t.co/edqvvcCik">https://t.co/edqvvcCik</a>

### 3.2. Preprocessing

Pada tahapan preprocessing, didalamnya terdapat beberapa proses yang digunakan untuk melakukan pembersihan data, diantaranya adalah sebagai berikut.

#### 3.2.1. Cleaning Character

Pada proses *cleaning character* akan dilakukan proses penghapusan nama pengguna, angka, kata 'RT', *hashtag*, URL, simbol emoji dan spasi kosong dari data tweet.[22]. Hasil dari proses *cleaning character* yang berhasil dilakukan untuk menghilangkan *noise* dalam dataset dapat dilihat pada contoh Tabel 2 Hasil *Cleaning Character*.

Tabel 3. Hasil *Cleaning Character*

Sebelum <i>Cleaning Character</i>	Sesudah <i>Cleaning Character</i>
Menjalankan Aturan seenaknya. Kemaren PilPres WalKot ikut Kontestasi ga ada ancaman apa?. Krn anak Presiden kan? Teros Pejabat yg kek gini tereak Profesional?  Tito Karnavian Ancam Pecat Penjabat Kepala Daerah yang Tak Mundur saat Maju Pilkada 2024	menjalankan aturan seenanya kemaren pilpres walkot ikut kontestasi ga ada ancaman apa krn anak presiden kan teros pejabat yg kek gini tereak profesional tito karnavian ancam pecat penjabat kepala daerah yang tak mundur saat maju pilkada

#### 3.2.2. Case Folding

Pada tahapan lanjutan setelah *cleaning character*, dilakukan tahapan *case folding* yang digunakan untuk melakukan konsistensi antara besar kecilnya huruf yang digunakan dalam dataset. Contoh dari *case folding* sendiri dapat terlihat dari Tabel 3 Hasil *Cleaning Character* yang sudah termasuk program untuk meakukan *case folding*.

#### 3.2.3. Tokenizing

Pada tahapan *tokenizing* yang digunakan untuk memecah string menjadi kata-kata terpisah atau token. Pada tahapan *tokenizing* ini hasil yang bisa ditampilkan adalah contoh pada Tabel 4 Hasil *Tokenizing*.

Tabel 4. Hasil *Tokenizing*

Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
menjalankan aturan seenaknya kemaren pilpres walkot ikut kontestasi ga ada ancaman apa krn anak presiden kan teros pejabat yg kek gini tereak profesional tito karnavian ancam pecat penjabat kepala daerah yang tak mundur saat maju pilkada	['menjalankan', 'seenaknya', 'aturan', 'kemaren', 'pilpres', 'walkot', 'ikut', 'kontestasi', 'ga', 'ada', 'ancaman', 'apa', 'krn', 'anak', 'presiden', 'kan', 'teros', 'pejabat', 'yg', 'kek', 'gini', 'tereak', 'profesional', 'karnavian', 'ancam', 'tito', 'pecat', 'penjabat', 'kepala', 'daerah', 'yang', 'mundur', 'saat', 'maju', 'pilkada']

#### 3.2.4. Stopwords Removal

Tahapan *stopwords removal* merupakan tahapan yang dilakukan untuk menghilangkan kata-kata yang kurang deskriptif ataupun kata-kata yang tidak signifikan. Pada tahapan *stopwords removal* yang dilakukan pada

penelitian ini memanfaatkan fungsi yang disediakan dalam *toolkit* NLTK. Untuk hasil dari *stopwords removal* sendiri dapat dilihat dalam Tabel 5 Hasil *Stopwords Removal*.

Tabel 5. Hasil Stopwords Removal

Sebelum Stopwords Removal	Sesudah Stopwords Removal
['menjalankan', 'aturan', 'seenaknya', 'kemaren', 'pilpres', 'walkot', 'ikut', 'kontestasi', 'ga', 'ada', 'ancaman', 'apa', 'krn', 'anak', 'presiden', 'kan', 'teros', 'pejabat', 'yg', 'kek', 'gini', 'terepak', 'profesional', 'tito', 'karnavian', 'ancam', 'pecat', 'penjabat', 'kepala', 'daerah', 'mundur', 'saat', 'maju', 'yang', 'tak', 'pilkada']	['menjalankan', 'aturan', 'kemaren', 'pilpres', 'walkot', 'kontestasi', 'ga', 'ancaman', 'krn', 'anak', 'presiden', 'teros', 'pejabat', 'yg', 'kek', 'gini', 'terepak', 'profesional', 'tito', 'karnavian', 'ancam', 'pecat', 'penjabat', 'kepala', 'daerah', 'mundur', 'maju', 'pilkada']

### 3.2.5. Stemming

Pada tahapan akhir dari *preprocessing* ini adalah *stemming*, tahapan *stemming* ini akan mengubah kata-kata dengan imbuhan menjadi kata-kata yang kembali pada bentuk dasarnya. Tahapan ini akan menghasilkan pembersihan dataset dari kata-kata imbuhan dan kembali pada tatanan bentuk kata dasarnya seperti pada Tabel 6 Hasil *Stemming*.

Tabel 6. Hasil Stemming

Sebelum Stemming	Sesudah Stemming
['menjalankan', 'aturan', 'kemaren', 'pilpres', 'walkot', 'kontestasi', 'ga', 'ancaman', 'krn', 'anak', 'presiden', 'teros', 'pejabat', 'yg', 'kek', 'gini', 'terepak', 'profesional', 'tito', 'karnavian', 'ancam', 'penjabat', 'pecat', 'kepala', 'daerah', 'mundur', 'maju', 'pilkada']	['menjalankan', 'aturan', 'kemaren', 'pilpr', 'walkot', 'kontestasi', 'ga', 'ancaman', 'krn', 'anak', 'presiden', 'tero', 'pejabat', 'yg', 'kek', 'gini', 'terepak', 'profesion', 'tito', 'karnavian', 'ancam', 'pecat', 'penjabat', 'kepala', 'daerah', 'mundur', 'maju', 'pilkada']

### 3.3. Pelabelan Data

Tahapan pelabelan data pada penelitian ini menggunakan VADER (*Valence Aware Dictionary and sEntiment Reasoner*), ada beberapa skor acuan untuk menetapkan label, diantaranya adalah skor  $\geq 0.05$  untuk skor positif,  $\leq -0.05$  untuk skor negatif dan selain dua skor tersebut akan masuk kedalam label netral. Hingga dalam proses labeling sebelum dilakukan *resampling* ini mendapatkan hasil sebanyak 1.155 data untuk label netral, 80 data untuk label positif dan 46 data untuk label negatif.

### 3.4. TF-IDF

Dalam tahapan penelitian ini dilanjutkan dengan melakukan proses TF-IDF untuk melakukan perhitungan atau pembobotan frekuensi kemunculan kata dalam dokumen tertentu. Pada proses TF-IDF mendapatkan hasil berupa pembobotan tiap kata yang ada pada dataset seperti pada contoh Tabel 6 Hasil TF-IDF.

Tabel 7. Hasil TF-IDF

Contoh Opini X	Hasil dari TF-IDF	
	Kata	Nilai TD-IDF
eh oon makan siang grati anak sekolah anak kuliah sd sd slta negeri setahuku grati kampanyenya bela aja dipakai utk nolak kebijakan pemenang pilpr bant lol tolol	aja	0.13586400475044766
	anak	0.26434389696863503
	bant	0.20542066937831982
	bela	0.20542066937831982
	dipakai	0.19458464820349114
	eh	0.18689636882662675
	grati	0.3438813671970859
	kampanyenya	0.20542066937831982
	kebijakan	0.1624085754892062
	kuliah	0.20542066937831982
	lol	0.20542066937831982
	makan	0.17194068359854295
	negeri	0.16837206827493367
	nolak	0.20542066937831982
	oon	0.19458464820349114
	pemenang	0.10008222452541675

	pilpr	0.029313785210724308
	sd	0.3737927376532535
	sekolah	0.19458464820349114
	setahuku	0.20542066937831982
	siang	0.15986141888834554
	slta	0.20542066937831982
	tolol	0.18093287604089928
	utk	0.11886725803829253

### 3.5. Resampling

Tahap penelitian selanjutnya adalah melalui proses *resampling*, pada tahapan ini, data yang sudah diproses akan mengalami penyeimbangan data. Pada proses ini menggunakan metode *SMOTETomek* yaitu gabungan dari proses *oversampling* dan juga *undersampling*. Pada tahapan sebelum dilakukannya *resampling*, pembagian data setelah tahap *labelling* mengalami ketidakseimbangan yang besar, sedangkan setelah melalui tahapan *resampling* menggunakan *SMOTETomek* terjadi keseimbangan data dengan label netral, positif dan negatif sebanyak 1.115.

### 3.6. K-Fold Cross Validation

Pada tahapan *K-Fold Cross Validation* yang digunakan untuk melakukan evaluasi dan sebagai metode uji kinerja model dengan menggunakan sejumlah data yang sudah disiapkan dan membagi data latihan menjadi K sub data. Jumlah K yang digunakan pada penelitian ini adalah 5 dan dengan *epoch* 10.

### 3.7. Pelatihan Model dengan Algoritma

Pada penelitian ini, perbandingan algoritma yang digunakan adalah antara *Convolutional Neural Network* dan juga *Recurrent Neural Network*. Algoritma CNN yang digunakan pada penelitian ini menggunakan metode arsitektur 1D *Convolutional Neural Network*, sedangkan untuk RNN sendiri menggunakan arsitektur LSTM (*Long Short-Term Memory*). Dalam proses *Convolutional Neural Network* pada penelitian ini menggunakan aktivasi *relu* pada layer *convolution* kemudian pada layer *output* menggunakan aktivasi *softmax*. Sedangkan, pada metode algoritma *Recurrent Neural Network* yang menggunakan arsitektur LSTM. Pada perbandingan hasil atau rata-rata akurasi antara model algoritma CNN dan juga RNN yang menggunakan penyeimbangan data dapat terlihat pada Tabel 8 Hasil Rata-rata Akurasi tiap model dengan dan tanpa *Resampling*.

Tabel 8. Hasil Rata-rata Akurasi tiap model dengan dan tanpa *Resampling*

	CNN dengan <i>resampling</i>	RNN dengan <i>resampling</i>	CNN tanpa <i>resampling</i>	RNN tanpa <i>resampling</i>
Rata-rata Akurasi Validasi dari 5 Fold	0.99	0.7713	0.94	0.8704

Tabel 8 Hasil Rata-rata Akurasi tiap model dengan atau tanpa *resampling* menunjukkan bahwa pengaruh dari *resampling* bisa menjadi signifikan untuk mendeteksi akurasi validasi beberapa model algoritma yaitu CNN (*Convolutional Neural Network*) sedangkan untuk model algoritma RNN (*Recurrent Neural Network*) hasil akurasi rata-rata menunjukkan hasil yang lebih baik dengan tanpa *resampling* daripada menggunakan *resampling*.

### 3.8. Hasil Evaluasi Model

Pada tahapan evaluasi model dalam penelitian ini menggunakan *confusion matrix* dengan kategori pembagian dalam perhitungan *confusion matrix* adalah Negatif, Netral dan Positif. Hasil *confusion matrix* pada tiap fold adalah pada Tabel 9 hingga Tabel 10.

Tabel 9. Hasil *Confusion Matrix* CNN

Confusion Matrix CNN dengan Resampling						Confusion Matrix CNN tanpa Resampling			
Fold	Metrik Evaluasi	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
1	Negatif	1.00	1.00	1.00	232	0.67	0.17	0.27	12
	Positif	0.99	1.00	1.00	205	1.00	0.71	0.83	21



	Netral	1.00	0.99	1.00	256	0.93	1.00	0.96	224
	Akurasi			1.00	693			0.93	257
2	Negatif	0.99	1.00	0.99	218	0.83	0.45	0.59	11
	Positif	0.98	1.00	0.99	244	0.64	0.88	0.74	8
	Netral	1.00	0.97	0.98	231	0.97	0.98	0.97	237
	Akurasi			0.99	693			0.95	256
3	Negatif	1.00	1.00	1.00	247	1.00	0.55	0.71	11
	Positif	0.99	1.00	0.99	230	0.75	0.60	0.67	20
	Netral	1.00	0.98	0.99	216	0.94	0.98	0.96	225
	Akurasi			0.99	693			0.93	256
4	Negatif	0.99	1.00	1.00	235	0.56	0.56	0.56	9
	Positif	0.98	1.00	0.99	230	0.77	0.67	0.71	15
	Netral	1.00	0.97	0.98	228	0.96	0.97	0.97	232
	Akurasi			0.99	693			0.94	256
5	Negatif	0.98	1.00	0.99	223	0.67	0.25	0.36	8
	Positif	1.00	1.00	1.00	246	1.00	0.68	0.81	22
	Netral	1.00	0.98	0.99	224	0.95	1.00	0.97	226
	Akurasi			0.99	693			0.95	256
Rata-rata Akurasi Validasi 5 Fold		0.99			0.94				

Dari Tabel 9 Hasil *Confusion Matrix* CNN dengan *resampling* dapat terlihat bahwa evaluasi model menggunakan CNN menunjukkan performa klasifikasi yang sangat baik, dengan *Precision*, *Recall*, dan *F1-Score* mendekati sempurna (0.99–1.00) pada semua kategori sentimen (Negatif, Positif, Netral) dan di setiap *fold* validasi. Performa yang konsisten baik ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang tinggi dan tidak terjadi *overfitting* terhadap data latih. Dengan rata-rata akurasi validasi sebesar 0.99, model ini dapat diandalkan untuk melakukan tugas klasifikasi sentimen pada data yang diuji. Sedangkan untuk hasil *Confusion Matrix* CNN tanpa *resampling* masih menunjukkan performa yang baik jika dilihat pada rata-rata hasil akurasi untuk data validasi sebesar 94% pada data uji. Meskipun pada klasifikasi kelas negatif dengan *recall* lebih rendah dengan 17% hingga 71%, klasifikasi kelas netral dan positif cenderung mendapatkan hasil yang baik. *Precision* yang didapatkan dari kelas positif dan netral mencapai hingga 100% dengan *recall* untuk kelas netral hampir sempurna dari 98% hingga 100% dan *F1-Score* untuk kelas positif dan netral berkisar antara 67% hingga 97%. *Precision* untuk kelas negatif bervariasi antara 56% hingga 100%, dan *F1-Score* untuk kelas negatif berkisar antara 27% hingga 83%.

Tabel 10. Hasil *Confusion Matrix* RNN

Fold	Metrik Evaluasi	Confusion Matrix RNN dengan Resampling				Confusion Matrix RNN tanpa Resampling			
		Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
1	Negatif	0.08	0.80	0.15	5	0.00	0.00	0.00	12
	Positif	0.31	0.65	0.42	17	0.46	0.76	0.57	21
	Netral	0.98	0.71	0.83	235	0.95	0.94	0.95	224
	Akurasi			0.71	257			0.88	257
2	Negatif	0.14	0.56	0.22	9	0.00	0.00	0.00	11
	Positif	0.34	0.60	0.44	20	0.15	1.00	0.25	8
	Netral	0.97	0.79	0.87	227	1.00	0.84	0.91	237
	Akurasi			0.77	256			0.81	256
3	Negatif	0.20	0.78	0.32	9	0.00	0.00	0.00	11
	Positif	0.54	0.82	0.65	17	0.34	0.70	0.46	20
	Netral	0.99	0.84	0.91	230	0.96	0.92	0.94	225
	Akurasi			0.84	256			0.86	256
4	Negatif	0.12	0.43	0.18	7	0.00	0.00	0.00	9
	Positif	0.37	0.88	0.52	17	0.00	0.00	0.00	15
	Netral	0.98	0.80	0.88	232	0.91	1.00	0.95	232
	Akurasi			0.80	256			0.91	256
5	Negatif	0.18	0.56	0.28	16	0.00	0.00	0.00	8

Positif	0.29	0.89	0.43	9	0.48	0.64	0.55	22
Netral	0.97	0.75	0.85	231	0.94	0.95	0.94	226
Akurasi			0.75	256			0.89	256
Rata-rata Akurasi	0.7713				0.8704			
Validasi 5 Fold								

Pada Tabel 10 Hasil *Confusion Matrix* RNN dengan *resampling* ini selama proses pelatihan memiliki performa akurasi yang meningkat sejumlah 35% sedari *epoch* awal hingga lebih dari 95% pada *epoch* terakhir, sembari bersamaan nilai *loss* yang kian menurun. Hal ini menunjukkan bahwa model masih mampu mempelajari pola data secara efektif dengan evaluasi menghasilkan rata-rata 74% hingga 77% dengan label kategori sentimen *neutral* memiliki performa terbaik dengan (*Precision*: 0.97–0.99, *Recall*: 0.71–0.79). Namun, kategori "Negatif" dan "Positif" menunjukkan performa yang lebih rendah, terutama pada *Precision* dan *Recall*, yang mengindikasikan bahwa model mengalami kesulitan dalam mengklasifikasikan data dengan jumlah sampel lebih sedikit. Secara keseluruhan, hasil evaluasi menunjukkan bahwa model cukup andal untuk klasifikasi sentimen pada data yang didominasi kategori "Netral," tetapi membutuhkan optimasi lebih lanjut untuk meningkatkan performa pada kategori minoritas. Sedangkan pada dataset pelatihan moel tanpa menggunakan *resampling* juga menghasilkan keluaran yang berbeda pula. Hasil *Confusion Matrix* RNN tanpa *resampling* menunjukkan hasil akurasi rata-ratanya sebesar 87.04% dimana kinerja untuk kelas label negatif sangat buruk, dengan *precision*, *recall* dan *f1-score* yang selalu 0 pada tiap fold. Sedangkan kelas positif memiliki performa yang bervariasi, dengan *precision* yang relatif rendah (0.00 hingga 0.48) dan *recall* yang lebih baik (hingga 1.00 di *fold 2*), namun *F1-Score* tetap rendah (0.00 hingga 0.57), dan kelas netral menunjukkan hasil yang jauh lebih baik, dengan *precision* dan *recall* hampir selalu mendekati 1, menghasilkan *F1-Score* yang sangat tinggi (0.91 hingga 0.95). Meskipun akurasi keseluruhan cukup baik, terutama untuk kelas netral, model ini kesulitan dalam mengklasifikasikan kelas negatif dan positif secara memadai, yang menunjukkan adanya ketidakseimbangan kelas dalam data.

### 3.9. Hasil Visualisasi Data

Pada tahapan terakhir dalam penelitian ini yaitu tahapan visualisasi data yang akan menggunakan world cloud, pada hasil world cloud yang dapat terlihat yang tepatnya menggunakan model CNN dengan teknik *resampling*, menunjukkan tatanan kata yang paling sering dibah, kosakata “pemilu”, nama salah satu calon presiden pun ikut muncul dalam visualisasi data ini yaitu “prabowo subianto” terdapat juga kosakata “kecurangan” yang terdapat didalamnya serta kosakata “wakil presiden” yang marak disebut juga memenuhi tampilan hasil visualisasi data pada pemodelan CNN dengan *resampling* ini. Kemudian pada pemodelan RNN dengan *resampling*, berbeda dengan model sebelumnya untuk model RNN dengan *resampling* ini lebih sering terindikasi kosakata “presiden pilpr”, “pemilu presiden”, selain pada beberapa kosakata yang masih muncul dalam pemodelan ini, terdapat beberapa kosakata yang tidak terdapat di world cloud sebelumnya, berupa “hak angket” dan juga “kecurangan pemilu” seperti pada Gambar 2 World Cloud CNN dan RNN dengan *Resampling*.

Hasil *world cloud* yang diperoleh dari proses *resampling* berbeda pula dengan hasil *world cloud* yang diperoleh dari proses tanpa *resampling*. Pada hasil *world cloud* tanpa *resampling* untuk pemodelan dengan CNN menunjukkan beberapa kosakata yang muncul terdapat “konstitusi mk”, terdapat juga kosakata “bansos” dan masih ada beberapa kosakata tetap yang akan muncul seperti “pilpre”, “pemilu” dan juga “pemilu presiden” seperti pada Gambar 3 *World Cloud* CNN dan RNN tanpa *resampling*.

Sedangkan dalam kosakata *world cloud* pada proses pemodelan RNN tanpa *resampling* dapat juga terlihat beberapa kosakata yang sering muncul masih sama dengan beberapa hasil *world cloud* sebelumnya, tetapi intensitas dalam keluarnya kosakata dalam *world cloud* pada proses pemodelan RNN tanpa *resampling* ini menunjukkan beberapa kosakata yang sebelumnya juga belum pernah hadir dalam *world cloud* sebelumnya yaitu salah satunya adalah nama calon presiden, “anies” dan “anies baswedan”, selain itu pola muncul kosakata yang ada dalam *world cloud* yang melalui proses RNN tanpa adanya *resampling* ini cukup banyak dan kurang beraturan seperti pada Gambar 3 *World Cloud* CNN dan RNN tanpa *Resampling*.



Gambar 2. World Cloud CNN (kiri) dan RNN(kanan) dengan Resampling



Gambar 3. World Cloud CNN (kiri) dan RNN(kanan) tanpa Resampling

#### 4. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa analisis sentimen terkait Pemilihan Umum Presiden 2024 menggunakan algoritma *Convolutional Neural Network (CNN)* dan *Recurrent Neural Network (RNN)* mampu mengklasifikasikan opini masyarakat di aplikasi Twitter atau X dengan tingkat akurasi yang bervariasi. CNN menunjukkan keunggulan dengan rata-rata akurasi validasi sebesar 99% menggunakan *resampling* dan 94% tanpa *resampling*, dibandingkan RNN yang memiliki akurasi rata-rata 77.13% dengan *resampling* dan 87.04% tanpa *resampling*. Temuan ini menegaskan bahwa CNN lebih unggul dalam presisi, efisiensi waktu, dan generalisasi pada dataset tidak seimbang. Untuk penelitian selanjutnya, disarankan memperluas cakupan data ke platform media sosial lain, menggunakan kata kunci yang lebih beragam, serta menerapkan metode deep learning terbaru seperti model berbasis Transformer (BERT, GPT) untuk meningkatkan akurasi dan kedalaman analisis sentimen. Selain itu, peningkatan kualitas dataset dengan teknik pemfilteran lanjutan dapat memberikan hasil yang lebih representatif dan akurat.

#### DAFTAR PUSTAKA

- [1] I. Syahrohim, S. D. Saputra, R. W. Saputra, V. H. Pranatawijaya, and R. Priskila, “PERBANDINGAN ANALISIS SENTIMEN SETELAH PILPRES 2024 DI TWITTER MENGGUNAKAN ALGORITMA MACHINE LEARNING,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, Apr. 2024, doi: 10.23960/jitet.v12i2.4249.
- [2] T. R. Dhanny Susetyawidianta Dan Galang Geraldny, “POLITICAL BRANDING PRABOWO-GIBRAN DALAM PEMILU PRESIDEN 2024: ANALISA INTERAKSIONISME SIMBOLIK DI MEDIA SOSIAL INSTAGRAM DAN,” vol. 3, no. 1, 2024, [Online]. Available: <https://jurnal.untagsby.ac.id/index.php/sintesa>
- [3] Ning Ayunda Chofifi and Eny Kusdarini, “Perkembangan Putusan Mahkamah Konstitusi dalam Menangani Perkara Perselisihan Hasil Pemilihan Umum Presiden,” *Jurnal Hukum IUS QUIA IUSTUM*, vol. 31, no. 2, pp. 408–433, Sep. 2024, doi: 10.20885/iustum.vol31.iss2.art8.
- [4] Rita Kusumawati, “Analisis Sentimen Pengguna Twitter Mengenai Pemilihan Presiden (Pilpres) Tahun 2024 Dengan Metode Naïve Bayes Classifier,” 2022, Accessed: Oct. 01, 2024. [Online]. Available: [https://digilib.uin-suka.ac.id/id/eprint/53735/1/18106050008\\_BAB-I\\_IV-atau-V\\_DAFTAR-PUSTAKA.pdf](https://digilib.uin-suka.ac.id/id/eprint/53735/1/18106050008_BAB-I_IV-atau-V_DAFTAR-PUSTAKA.pdf)
- [5] M. Wankhade, A. C. S. Rao, and C. Kulkarni, “A survey on sentiment analysis methods, applications, and challenges,” *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

- 
- [6] D. W. Otter, J. R. Medina, and J. K. Kalita, "A Survey of the Usages of Deep Learning for Natural Language Processing," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 2, pp. 604–624, Feb. 2021.
- [7] A. Rajput, "Natural language processing, sentiment analysis, and clinical analytics," pp. 79–97, Feb. 2019.
- [8] E. Y. Hidayat and D. Handayani, "Penerapan 1D-CNN untuk Analisis Sentimen Ulasan Produk Kosmetik Berdasar Female Daily Review," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 3, pp. 153–163, Jan. 2023, doi: 10.25077/teknosi.v8i3.2022.153-163.
- [9] B. Laurensz, A. Sentimen, and E. Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19 (Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic)," 2021.
- [10] D. A. N. Taradhita and I. K. G. D. Putra, "Hate speech classification in Indonesian language tweets by using convolutional neural network," *Journal of ICT Research and Applications*, vol. 14, no. 3, pp. 225–239, 2021, doi: 10.5614/itbj.ict.res.appl.2021.14.3.2.
- [11] F. Gurcan and A. Soyly, "Learning from Imbalanced Data: Integration of Advanced Resampling Techniques and Machine Learning Models for Enhanced Cancer Diagnosis and Prognosis," *Cancers (Basel)*, vol. 16, no. 19, Oct. 2024, doi: 10.3390/cancers16193417.
- [12] M. Humayun, D. Javed, N. Jhanjhi, M. Fahaad Almufareh, and S. Naif Almuayqil, "Deep Learning Based Sentiment Analysis of COVID-19 Tweets via Resampling and Label Analysis," *Computer Systems Science and Engineering*, vol. 47, no. 1, pp. 575–591, 2023, doi: 10.32604/csse.2023.038765.
- [13] J. Jumanto, M. A. Muslim, Y. Dasril, and T. Mustaqim, "Accuracy of Malaysia Public Response to Economic Factors During the Covid-19 Pandemic Using Vader and Random Forest," *Journal of Information System Exploration and Research*, vol. 1, no. 1, pp. 49–70, Dec. 2022, doi: 10.52465/joiser.v1i1.104.
- [14] Yosua Alberth Sir and Agus H H Soepranoto, "Pendekatan Resampling Data Untuk Menangani Masalah Ketidakseimbangan Kelas," *Semantic Scholar*, Mar. 2022.
- [15] D. Shabrina Assyifa and A. Luthfiarta, "SMOTE-Tomek Re-sampling Based on Random Forest Method to Overcome Unbalanced Data for Multi-class Classification," *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 9, no. 2, pp. 151–160, Jul. 2024, doi: 10.25139/inform.v9i2.8410.
- [16] I. K. Trisiawan and Y. Yuliza, "Penerapan Multi-Label Image Classification Menggunakan Metode Convolutional Neural Network (CNN) Untuk Sortir Botol Minuman," *Jurnal Teknologi Elektro*, vol. 13, no. 1, p. 48, Feb. 2022, doi: 10.22441/jte.2022.v13i1.009.
- [17] B. A. Yuniarossy *et al.*, "ANALISIS SENTIMEN TERHADAP ISU FEMINISME DI TWITTER MENGGUNAKAN MODEL CONVOLUTIONAL NEURAL NETWORK (CNN)," vol. 5, no. 1, 2024, doi: 10.46306/lb.v5i1.
- [18] M. E. Basiri, M. A. S. Nemati, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, 2021, doi: 10.1016/j.future.2020.08.005.
- [19] I. N. Husada and H. Toba, "Pengaruh Metode Penyeimbangan Kelas Terhadap Tingkat Akurasi Analisis Sentimen pada Tweets Berbahasa Indonesia," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, Aug. 2020, doi: 10.28932/jutisi.v6i2.2743.
- [20] M. Rahardi, A. Majid, A. Aminuddin, F. Abdulloh, and R. Nugroho, "Sentiment Analysis of Covid-19 Vaccination using Support Vector Machine in Indonesia," *International Journal of Advanced Computer Science and Applications*, vol. 13, 2022.
- [21] S. Syafrizal, M. Afdal, and R. Novita, "Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 10–19, Dec. 2023, doi: 10.57152/malcom.v4i1.983.
- [22] Putri Lannidya Parameswari and Prihandoko Prihandoko, "PENGGUNAAN CONVOLUTIONAL NEURAL NETWORK UNTUK ANALISIS SENTIMEN OPINI LINGKUNGAN HIDUP KOTA DEPOK DI TWITTER," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 27, no. 1, pp. 29–42, 2022.