

Analisis Sentimen Pengguna X (Twitter) Terhadap Kebijakan Tapera Di Indonesia Menggunakan Metode CNN Dan BERT

Syindikha Putri Inombi¹, Dewi Rahmawaty Isa², Asriadi³, Salmun K. Nasib⁴, Isran K. Hasan⁵, Siti Nurmardia Abdussamad⁶

¹Program Studi Statistika, Universitas Negeri Gorontalo, Bone Bolango 96554, Indonesia

*Penulis korespondensi: syindikhaputri@gmail.com

Abstrak

Program Tabungan Perumahan Rakyat (TAPERA) yang dicanangkan pemerintah telah menimbulkan beragam reaksi dari masyarakat, terutama di media sosial X (*Twitter*). Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap kebijakan Tapera menggunakan metode *Convolutional Neural Network* (CNN) dan *Bidirectional Encoder Representations from Transformers* (BERT). Dataset diperoleh melalui teknik *crawling* pada media sosial X (*Twitter*), dengan total 1.790 tweet yang kemudian diproses melalui tahapan preprocessing, termasuk *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. Model CNN dan BERT kemudian dilatih dan diuji untuk mengklasifikasikan sentimen positif dan negatif. Hasil penelitian menunjukkan bahwa model BERT memiliki akurasi lebih tinggi (86%) dibandingkan CNN (85%), dengan nilai *recall*, *precision* dan *F1-score* yang lebih baik. Hasil ini menunjukkan bahwa metode BERT lebih efektif dalam memahami konteks sentimen secara keseluruhan

Kata Kunci: Analisis Sentimen, Tapera, CNN, BERT, Social Media X (*Twitter*)

Abstract

The government's Housing Savings Program (TAPERA) has sparked various public reactions, particularly on social media platform X (*Twitter*). This study aims to analyze user sentiment toward the TAPERA policy using the *Convolutional Neural Network* (CNN) and *Bidirectional Encoder Representations from Transformers* (BERT) methods. The dataset was collected using a *crawling* technique on X (*Twitter*), comprising a total of 1,790 tweets. These data were processed through preprocessing stages, including *cleaning*, *case folding*, *normalization*, *tokenization*, *stopword removal*, and *stemming*. The CNN and BERT models were then trained and tested to classify sentiments as positive or

Article History

Received: April 2025

Reviewed: April 2025

Published: April 2025

Plagirism Checker No 234

Prefix DOI : Prefix DOI :
10.3483/trigonometri.v1i1.800

Copyright : Author

Publish by : Trigonometri



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

negative. The findings indicate that the BERT model outperformed CNN, achieving a higher accuracy of 86% compared to CNN's 85%, along with superior recall, precision, and F1-score values. These results suggest that the BERT method is more effective in comprehensively understanding sentiment context.

Keywords: Sentiment Analysis, Tapera, CNN, BERT, Social Media X (Twitter)

1. Pendahuluan

Perumahan merupakan kebutuhan dasar manusia setelah pangan dan sandang. Pemerintah Indonesia bertanggung jawab dalam menyediakan perumahan yang terjangkau, namun masih menghadapi tantangan seperti keterbatasan lahan dan kondisi rumah yang tidak memadai [1]. Urbanisasi yang terus meningkat semakin memperburuk keterbatasan lahan, terutama bagi masyarakat berpenghasilan rendah yang kesulitan mendapatkan kredit perumahan [2]. Sebagai solusi, Pemerintah Indonesia menerbitkan Peraturan Pemerintah (PP) Nomor 21 Tahun 2024 tentang Tabungan Perumahan Rakyat (Tapera). Program ini bertujuan untuk menyediakan pendanaan perumahan yang berkelanjutan [3]. Namun, kebijakan ini memicu berbagai reaksi masyarakat, baik dukungan maupun kritik, yang banyak diungkapkan melalui media sosial, khususnya X (Twitter), yang memiliki lebih dari 24,69 juta pengguna aktif di Indonesia pada tahun 2024 [4].

Reaksi masyarakat terhadap kebijakan Tapera masih beragam dan belum sepenuhnya dipahami secara sistematis. Analisis sentimen diperlukan untuk mengidentifikasi opini publik secara lebih mendalam. Metode *deep learning* seperti *Convolutional Neural Network* (CNN) dan *Bidirectional Encoder Representations from Transformers* (BERT) telah terbukti efektif dalam analisis sentimen, namun belum banyak penelitian yang membandingkan kedua metode ini dalam konteks kebijakan perumahan di Indonesia. Penelitian sebelumnya telah banyak menggunakan metode CNN dan BERT dalam analisis sentimen pada berbagai isu seperti politik, ekonomi, dan kebijakan sosial. [5] menemukan bahwa CNN mencapai akurasi 94,94% dalam analisis sentimen pemindahan ibu kota, sementara [6] menemukan bahwa model BERT dapat mencapai akurasi 99% dalam analisis sentimen ulasan aplikasi digital. Namun, belum banyak penelitian yang secara khusus membandingkan kedua metode ini dalam analisis kebijakan perumahan.

Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial X (Twitter) terhadap kebijakan Tapera, membandingkan performa metode CNN dan BERT dalam mengklasifikasikan sentimen positif dan negatif, serta memberikan rekomendasi bagi pemangku kebijakan berdasarkan hasil analisis opini publik. Dengan demikian, penelitian ini memiliki kebaruan dalam penerapan CNN dan BERT untuk analisis sentimen kebijakan Tapera. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pembuat kebijakan dalam merancang strategi komunikasi yang lebih efektif terkait program perumahan rakyat di Indonesia.

2. Tinjauan Pustaka

2.1. Analisis Sentimen

Analisis sentimen merupakan suatu metode penelitian yang mengkaji berbagai opini, penilaian, evaluasi, serta respons emosional dan sikap yang berhubungan dengan berbagai aspek, seperti produk, layanan, institusi, individu, isu, kejadian, topik, dan karakteristik lainnya [7]. Sebagian besar tugas dalam analisis sentimen adalah untuk mengidentifikasi sentimen dalam sebuah teks serta mengklasifikasikan polaritasnya, baik pada tingkat dokumen maupun kalimat, sehingga dapat dikategorikan sebagai sentimen positif atau negatif [8]. Dengan memanfaatkan internet, seseorang dapat mengumpulkan ulasan tentang merek, produk, atau individu, dan menentukan apakah ulasan tersebut cenderung bersifat positif atau negatif [9].

2.2. TAPERA (Tabungan Perumahan Rakyat)

Tapera adalah tabungan yang disimpan oleh peserta berdasarkan gaji, upah, atau penghasilan bulanan dengan persentase tertentu. Dana ini nantinya dapat digunakan untuk membiayai perumahan dan/atau sebagai tabungan. Jika kepesertaan berakhir, dana yang telah dikumpulkan oleh peserta Tapera akan dikembalikan dan dapat dimanfaatkan [10].

2.3. Text Preprocessing

Text preprocessing adalah memproses data awal melalui serangkaian tahapan sehingga data tersebut dapat digunakan dengan tepat [11]. Tahapan *Text Preprocessing* adalah :

- *Cleaning* adalah tahap pembersihan untuk menghapus kata-kata yang tidak diinginkan guna mengurangi gangguan dalam proses klasifikasi
- *Case folding* adalah langkah membuat semua huruf besar dalam dataset menjadi huruf kecil.
- *Tokenizing* adalah proses memecah kalimat menjadi unit-unit data tunggal, tujuannya untuk menyederhanakan proses transformasi dengan memproses kata-kata individual dari kalimat.
- *Remove Stopword* adalah proses filtrasi di mana kata-kata yang tidak penting dihapus dari hasil tokenisasi, yaitu kata-kata yang tidak mewakili isi dokumen.
- *Stemming* adalah suatu proses yang bertujuan untuk mengidentifikasi dan mengubah berbagai bentuk kata menjadi bentuk dasarnya. Dalam proses ini, kita mencari kata dasar dengan cara menghapus imbuhan atau variasi yang terdapat pada kata tersebut.
- *Normalization* adalah serangkaian tahap *preprocessing* yang diterapkan dalam penelitian ini untuk mengonversi dataset dari format yang tidak terstruktur menjadi lebih terstruktur, sehingga mempermudah proses pengolahan data [12]

2.4. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) awalnya digunakan dalam pemrosesan gambar,

namun telah diterapkan dalam NLP untuk mengklasifikasikan teks. CNN menggunakan teknik konvolusi kalimat, paragraf, atau seluruh dokumen teks [13]. Proses utama dalam CNN meliputi konvolusi, *pooling layer*, dan *fully connected layer*, yang membantu dalam ekstraksi fitur teks yang lebih baik [14]. Rumus dasar perhitungan CNN [5]

1. Konvolusi (*Convolution*)

Rumus Konvolusi umum adalah :

$$C_i = f(W \times X_{i:i+h-1} + b) \quad (2.1)$$

Keterangan:

- C_i : hasil konvolusi diposisi i
- f : fungsi aktivasi
- W : bobot kernel
- $X_{i:i+h-1}$: potongan vektor kata-kata dari posisi i hingga i+h-1
- b : bias

2. *Pooling (Max Pooling)* Rumus *Max Pooling* adalah :

$$P_j = \max(C_{j:j+k-1}) \quad (2.2)$$

Keterangan

- P_j : hasil *max pooling* pada posisi j
- $C_{j:j+k-1}$: kelompok hasil konvolusi yang berdekatan
- k : ukuran *pooling*

3. *Fully Connected Layer*

Rumus *fully connected layer* adalah :

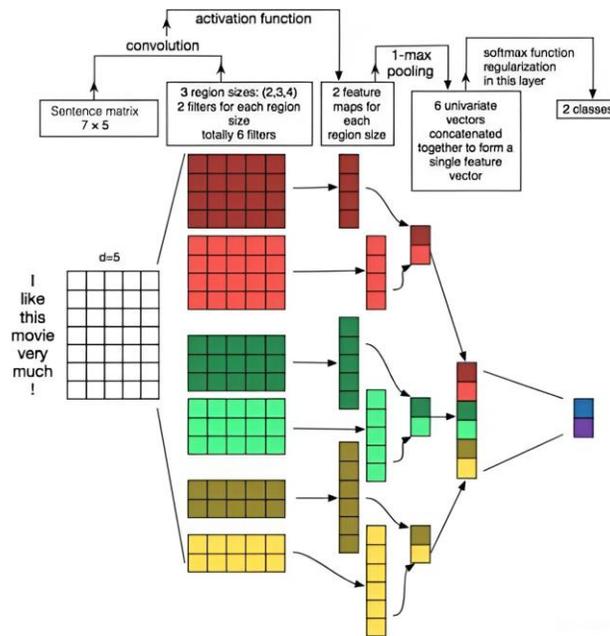
$$O = f(W' \times P + b') \quad (2.3)$$

Keterangan:

- O : keluaran dari lapisan sepenuhnya terhubung
- f : fungsi aktivasi
- W' : matriks bobot lapisan sepenuhnya terhubung
- P : vektor hasil *pooling*
- b' : bias lapisan sepenuhnya terhubung

Input untuk sebagian besar tugas pemrosesan bahasa alami adalah kalimat atau dokumen yang direpresentasikan dalam bentuk matriks. Setiap baris matriks berhubungan dengan satu token, yang biasanya merupakan sebuah kata, meskipun bisa juga berupa karakter. Ini berarti bahwa

setiap baris merupakan vektor yang menggambarkan representasi suatu kata. Biasanya, vektor-vektor ini ditimbang menggunakan metode penyematan kata yang mengindeks kata tersebut ke dalam kosakata. Ilustrasinya dapat ditemukan pada gambar 2.1, sementara untuk pemrosesan bahasa alami kami mengubah masukan berdasarkan teks menjadi matriks vektor dan menerapkan fungsi aktivasi linier pada lapisan neuron [14].



Gambar 2.1: CNN dalam Ilustrasi Pemrosesan Bahasa Alami

2.5. Bidirectional Encoder Representations from Transformers (BERT)

BERT adalah model berbasis NLP *Transformer* yang mampu memahami konteks kata dalam suatu kalimat secara lebih akurat dibandingkan metode lain. Dikembangkan oleh Delvin dkk tahun 2018, BERT telah banyak digunakan dalam berbagai tugas NLP, termasuk analisis sentimen. Model ini memungkinkan pemahaman konteks yang lebih mendalam [15]. BERT adalah model *deep learning* yang telah membawa kemajuan besar dalam berbagai tugas *Natural Language Processing* (NLP). BERT menggunakan arsitektur *Transformer* dengan 6 lapisan yang saling bertumpuk dalam *encoder* dan *decoder*. Hal Ini menyebabkan proses pelatihan yang kompleks, pengaturan yang rumit, durasi pelatihan yang lama, serta biaya tinggi. Namun, Google merilis model *pretrained* BERT sebagai sumber terbuka, sehingga penulis bisa menggunakannya tanpa harus membuat model dari awal [16].

2.6. Confusion Matrix

Evaluasi menggunakan *confusion matrix* menghasilkan nilai-nilai seperti *Accuracy*, *Precision*, *Recall*, dan *F1-score* [17]. Berikut adalah persamaan model *Confusion Matrix* [15]:

1. Accuracy

Accuracy adalah perbandingan antara jumlah prediksi yang benar, baik positif maupun negatif, dengan seluruh data yang ada

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.4)$$

2. Precision

Precision merupakan perbandingan antara *True Positive* (TP) dengan banyak data yang diprediksi positif

$$Precision = \frac{TP}{TP + FP} \quad (2.5)$$

3. Recall

Recall membandingkan nilai *True Positive* (TP) dengan banyak nilai yang benar-benar positif

$$Recall = \frac{TP}{TP + FN} \quad (2.6)$$

4. F1-score

F1-score merupakan nilai rasio rata-rata perbandingan *precision* dengan nilai *recall*

$$F1-score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (2.7)$$

3. Metode Penelitian

Data yang digunakan pada penelitian ini merupakan data yang didapat dari proses *crawling* pada media sosial X (*Twitter*) dari bulan November 2015 – Juni 2024 dengan kata kunci “Kebijakan Tapera” dan “#Tapera”. Dari proses tersebut didapatkan data tweets sebanyak 1790 data. Variabel yang digunakan pada penelitian ini adalah *tweet* kategori teks pengguna X (*Twitter*). Teknik penarikan sampel dalam penelitian ini menggunakan teknik *purposive sampling*.

Penelitian ini menggunakan metode *Convolutional Neural Network* (CNN) dan *Bidirectional Encoder Representations from Transformers* (BERT). Untuk analisis data pada penelitian ini menggunakan *software Google Colab* untuk menjalankan pemrograman *Python 3.10*. Berikut ini tahapan pada penelitian :

1. Melakukan pengumpulan data yang ada pada media sosial X (*Twitter*), dengan cara *crawling* data dengan menggunakan *search keyword* “Kebijakan Tapera” dan “#Tapera”
2. Melakukan *preprocessing* data dari X (*Twitter*), yaitu dengan cara *Cleaning, Case Folding, Normalized, Tokenizing, Stopword Removal, Stemming*
3. Melakukan pelabelan data yaitu proses pemberian label positif dan negatif pada dataset
4. Kemudian melakukan perancangan arsitektur model CNN dan Pembangunan model BERT
5. Melakukan pelatihan model yang telah dibuat, menggunakan data *training* dan data *testing*

6. Membuat *Confussion Matrix* dari hasil model menggunakan data *testing*
7. Perbandingan kinerja CNN dan BERT dilihat dari hasil *Confussion Matrix*

4. Hasil dan Pembahasan

4.1. Pengambilan Data

Pengambilan data dilakukan melalui metode *crawling* pada media sosial X (*Twitter*) dengan menggunakan perangkat lunak *tweet-harvest* versi 2.6.1. Data yang dikumpulkan mencakup *tweet* yang berisi kata kunci “Kebijakan Tapera” dan “#Tapera” dalam kurun waktu November 2015 hingga Juni 2024. Hasil pengumpulan data menghasilkan 1.790 *tweet* yang mencerminkan opini publik terhadap kebijakan TAPERERA.

4.2. Preprocessing Data

Setelah data dikumpulkan, Langkah selanjutnya adalah preprocessing untuk membersihkan dan menyiapkan data agar siap digunakan dalam analisis sentiment. Langkah-langkah nya meliputi :

- *Cleaning* : menghapus karakter yang tidak diperlukan seperti simbol, angka, tautan, dan tagar
- *Case Folding* : mengubah semua huruf menjadi huruf kecil
- *Tokenizing* : memisahkan teks menjadi kata-kata individual
- *Stopword Removal* : menghilangkan kata-kata umum yang tidak memiliki makna signifikan
- *Stemming* : mengubah kata menjadi bentuk dasarnya
- *Normalization* : menyesuaikan format kata agar seragam

Setelah preprocessing selesai, jumlah tweet tetap 1790, tetapi dalam bentuk yang lebih terstruktur untuk analisis lebih lanjut.

4.3. Pelabelan Data

Proses pelabelan data dilakukan menggunakan *Indonesian Sentiment Lexicon*, yang mengklasifikasikan *tweet* menjadi tiga kategori yakni positif, negatif, dan netral. Hasil pelabelan menunjukkan bahwa 811 *tweet* memiliki sentimen negatif, 616 *tweet* memiliki sentimen positif, dan 363 *tweet* memiliki sentimen netral.

4.4. Convolutional Neural Network (CNN)

Model CNN dirancang untuk terdiri dari sejumlah lapisan utama yang menjadi komponen inti dalam struktur model

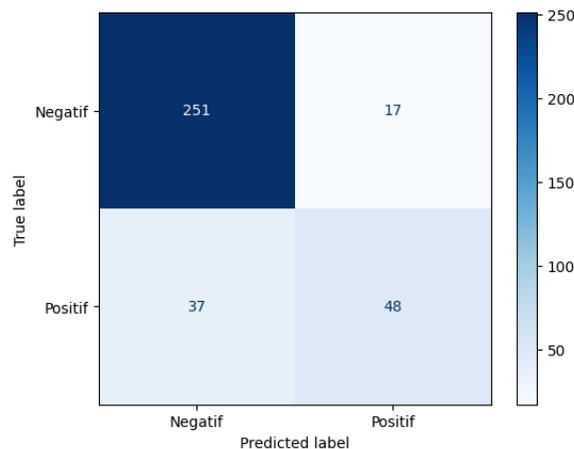
Model: "functional"

| Layer (type) | Output Shape | Param # |
|---|-----------------|---------|
| input_layer (InputLayer) | (None, 30) | 0 |
| embedding (Embedding) | (None, 30, 200) | 708,200 |
| conv1d (Conv1D) | (None, 28, 64) | 38,464 |
| global_max_pooling1d (GlobalMaxPooling1D) | (None, 64) | 0 |
| dropout (Dropout) | (None, 64) | 0 |
| dense (Dense) | (None, 1) | 65 |

Total params: 746,729 (2.85 MB)
 Trainable params: 746,729 (2.85 MB)
 Non-trainable params: 0 (0.00 B)

Gambar 4.2: Lapisan Model CNN

Model CNN dilatih dengan pembagian data latih (80%) dan data uji (20%), di mana dalam penelitian ini hanya menggunakan *tweet* sentiment negatif dan positif dengan total data 1763 data. Setelah proses pelatihan dengan menggunakan 50 *epochs* selesai, hasil evaluasi menunjukkan akurasi 85%, dengan *precision* 74%, *recall* 56%, dan *F1-score* 64%. Berikut *confusion matrix* yang diperoleh.



Gambar 4.3: Confusion Matrix CNN

Dari hasil *confusion matrix* dapat disimpulkan *False Negative* (FN) dengan 17 *tweet*, menunjukkan jumlah kesalahan prediksi ketika model mengidentifikasi data dengan label asli positif sebagai negatif, *False Positive* (FP) dengan 37 *tweet*, menunjukkan jumlah kesalahan prediksi ketika model mengidentifikasi data dengan label asli negatif sebagai positif, *True Negative* (TN) dengan 251 *tweet*, menunjukkan jumlah prediksi yang benar ketika model mengidentifikasi data dengan label asli negatif sebagai negatif, Ini menunjukkan kemampuan model mengenali sentimen negatif dengan tepat. *True Positive* (TP) dengan 48 *tweet*, menunjukkan jumlah prediksi yang benar ketika model mengidentifikasi data dengan label asli positif sebagai positif.

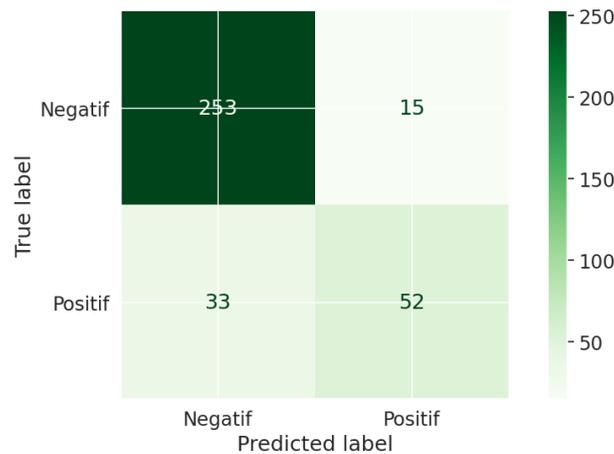
4.5. Bidirectional Encoder Representations from Transformers (BERT)

BERT menggunakan arsitektur *Transformer* untuk memahami konteks kata dalam teks. Model ini dilatih menggunakan dataset yang sama dan diuji dengan parameter *batch size* 32, *learning rate* $2e-5$, *epsilon* $1e-8$ dan *epochs* 50. Menghasilkan garafik *training loss*



Gambar 4.4: *Training Loss Epochs 50*

Setelah pelatihan selesai, performa model dievaluasi menggunakan data uji untuk mengukur tingkat akurasi serta kemampuan model dalam melakukan klasifikasi sentiment. Hasil evaluasi model BERT menunjukkan akurasi 86%, dengan *precision* 78%, *recall* 61%, dan *F1-score* 68%. Analisis *confusion matrix* yang dihasilkan oleh model BERT



Gambar 4.5: *Confusion Matrix BERT*

Dari hasil *confusion matrix* dapat disimpulkan *False Negative* (FN) dengan 15 *tweet*, menunjukkan jumlah kesalahan prediksi ketika model mengidentifikasi data dengan label asli positif sebagai negatif, *False Positive* (FP) dengan 33 *tweet*, menunjukkan jumlah kesalahan prediksi ketika model mengidentifikasi data dengan label asli negatif sebagai positif, *True Negative* (TN) dengan 253 *tweet*, menunjukkan jumlah prediksi yang benar ketika model mengidentifikasi data dengan label asli negatif sebagai negatif, Ini menunjukkan kemampuan model mengenali sentimen

negatif dengan tepat. *True Positive* (TP) dengan 52 *tweet*, menunjukkan jumlah prediksi yang benar ketika model mengidentifikasi data dengan label asli positif sebagai positif.

4.6. Perbandingan Hasil dan Performa antara CNN dan BERT

Berikut perbandingan performa kedua model ditunjukkan dalam tabel berikut :

Tabel 4.1: Hasil *Accuracy*, *Precision*, *Recall*, *F1-score* dari Kedua Model

| Metode | <i>Accuracy</i> | <i>Precision</i> | <i>Recall</i> | <i>F1-score</i> |
|--------|-----------------|------------------|---------------|-----------------|
| CNN | 0.85 | 0.74 | 0.56 | 0.64 |
| BERT | 0.86 | 0.78 | 0.61 | 0.68 |

- *Accuracy*: CNN memperoleh akurasi sebesar 85%, sementara BERT memperoleh 86%. Perbandingan: BERT memiliki akurasi yang lebih tinggi dibandingkan CNN.
- *Precision*: CNN mencapai 74%, sementara BERT memperoleh 78% untuk nilai presisi. Perbandingan: BERT menunjukkan nilai *precision* yang lebih tinggi, yang mengindikasikan kecenderungan untuk mengidentifikasi dengan lebih akurat kasus positif yang sebenarnya.
- *Recall*: CNN mencatatkan nilai *recall* sebesar 56%, sedangkan BERT mencapai 61%. Perbandingan: BERT menunjukkan nilai *recall* yang cenderung lebih tinggi, yang mengindikasikan kemampuannya untuk mengenali lebih banyak instance positif secara keseluruhan.
- *F1-Score*: CNN mencatatkan 64%, sedangkan BERT mencapai 68%. Perbandingan: BERT menunjukkan nilai *F1-Score* yang lebih tinggi, menandakan keseimbangan yang lebih optimal antara *Precision* dan *Recall* dibandingkan dengan CNN.

5. Kesimpulan

Hasil penelitian menunjukkan bahwa metode BERT lebih unggul dalam analisis sentimen terhadap kebijakan TAPERA dibandingkan metode CNN, dengan akurasi, *recall*, *precision*, dan *F1-score* yang lebih tinggi. Oleh karena itu, BERT lebih direkomendasikan untuk memahami opini publik secara lebih akurat.

Referensi

- [1] N. Adawiyah, J. A. Sihombing, S. Mar'ah, and P. Kemala, "Analisis kebijakan perumahan subsidi di indonesia policy analysis Of subsidized housing in indonesia," *El-Mal: Jurnal Kajian Ekonomi & Bisnis Islam*, vol. 5, no. 6, pp. 3064–3075, 2024.
- [2] Y. M. De, "Imam dan dialog sosialnya (analisis kritis program tapera "tabungan perumahan rakyat" bagi kehidupan umat di paroki riam batang kalimantan tengah)," *Jurnal Pendidikan Agama dan Teologi*, vol. 2, no. 3, pp. 57–73, 2024.

- [3] A. H. Alfrida, E. M. Sasea, and A. Suryana, "Legal challenges of tapera implementation in indonesia: Toward an equitable and effective housing financing scheme," *Legalis: Journal of Law Review*, vol. 2, no. 2, pp. 62–71, 2024.
- [4] L. D. Mahbubah and E. Zuliarso, "Analisa sentimen twitter pada pilpres 2019 menggunakan algoritma naive bayes," 2019.
- [5] G. Y. Sitio, S. A. Rumapea, and P. Lumbanraja, "Analisis sentimen pemindahan ibu kota negara di media sosial twitter menggunakan metode convolutional neural network (cnn)," *METHOTIKA: Jurnal Ilmiah Teknik Informatika*, vol. 3, no. 2, pp. 97–104, 2023.
- [6] R. M. R. W. P. Kusuma, W. Yustanti, *et al.*, "Analisis sentimen customer review aplikasi ruang guru dengan metode bert (bidirectional encoder representations from transformers)," *Journal of Emerging Information System and Business Intelligence (JEISBI)*, vol. 2, no. 3, 2021.
- [7] B. Liu, *Sentiment analysis and opinion mining*. Springer Nature, 2022.
- [8] P. W. Ratiasasadara, S. Sudarno, and T. Tarno, "Analisis sentimen penerapan ppkm pada twitter menggunakan naive bayes classifier dengan seleksi fitur chi-square," *Jurnal Gaussian*, vol. 11, no. 4, pp. 580–590, 2023.
- [9] M. N. Humam, "Perbandingan kinerja cnn dan naive bayes pada analisis sentimen performa manchester united di twitter," *JIEET (Journal of Information Engineering and Educational Technology)*, vol. 7, no. 2, pp. 83–91, 2023.
- [10] C. Ariningdyah, D. Lasonda, and F. R. D. Miarsa, "Analisis yuridis penerapan tabungan perumahan rakyat (tapera) dalam perspektif asas keadilan," *Innovative: Journal Of Social Science Research*, vol. 4, no. 3, pp. 18410–18424, 2024.
- [11] O. Oueslati, E. Cambria, M. B. HajHmida, and H. Ounelli, "A review of sentiment analysis research in arabic language," *Future Generation Computer Systems*, vol. 112, pp. 408–430, 2020.
- [12] S. I. Nurhafida and F. Sembiring, "Analisis text clustering masyarakat di twiter mengenai mcdonald'sxbts menggunakan orange data mining," in *Prosiding Seminar Nasional Sistem Informasi Dan Manajemen Informatika Universitas Nusa Putra*, vol. 1, pp. 28–35, 2021.
- [13] S. Imron, E. I. Setiawan, J. Santoso, M. H. Purnomo, *et al.*, "Aspect based sentiment analysis marketplace product reviews using bert, lstm, and cnn," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 3, pp. 586–591, 2023.
- [14] H. Tsaniya, R. Rosadi, and A. Abdullah, "Sentiment analysis towards jokowis government using twitter data with convolutional neural network method," in *Journal of Physics: Conference Series*, vol. 1722, p. 012017, IOP Publishing, 2021.
- [15] S. Setyani and Y. Sibaroni, "Multi aspect sentiment analysis of mutual funds investment app bibit using bert method," *International Journal on Information and Communication Technology (IJoICT)*, vol. 9, no. 1, pp. 44–56, 2023.
- [16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [17] D. Putra and A. Wibowo, "Prediksi keputusan minat penjurusan siswa sma yadika 5 menggunakan algoritma naive bayes," in *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 2, pp. 84–92, 2020.