

SENTIMENT ANALYSIS SUDUT PANDANG GENERASI Z
TERHADAP KETERWAKILAN KEPEMIMPINAN MUDA
PILPRES DI TWITTER MENGGUNAKAN ID CONVULATION
NEURAL NETWORK

Tri puji Astuti¹, Rozali Toyib ²,

1,2) Teknik iInformatika, Teknik, Universitas Muhammadiyah Bengkulu, Indonesia

Article Info

Article history:

Received: 14 Juli 2024
Revised: 18 Juli 2024
Accepted: 20 Juli 2024

ABSTRACT

Abstrak

Penelitian ini mengeksplorasi penggunaan pemrosesan bahasa alami (NLP) untuk menganalisis sentimen dalam tweet mengenai kepemimpinan muda dalam pemilihan presiden. Data dikumpulkan dari Twitter, kemudian diproses dan diekstraksi. Proses tokenisasi menghasilkan tokenisasi kata yang kemudian diklasifikasikan menggunakan algoritma Convolutional Neural Network (CNN). Penelitian ini bertujuan untuk membandingkan akurasi algoritma CNN dalam klasifikasi sentimen. Data dikategorikan menjadi dua kelas: positif dan negatif, setelah dikumpulkan dari Twitter. Pre-processing dilakukan menggunakan teknik NLP sebelum tokenisasi. Token yang dihasilkan kemudian dimasukkan ke dalam model CNN untuk proses klasifikasi. Model 1D CNN yang digunakan terdiri dari satu lapisan konvolusi dan satu lapisan pooling, dengan aktivasi ReLU pada lapisan konvolusi. Dataset terdiri dari 200 data, dipecah menjadi 160 data pelatihan dan 40 data uji. Hasil penelitian menunjukkan bahwa label prediksi sesuai dengan label sebenarnya dari kalimat baru. Namun, beberapa faktor seperti ejaan yang tidak konsisten, penggunaan bahasa yang tidak formal, kata-kata yang memiliki arti tersirat atau ambigu, dan kesamaan kata dalam data pelatihan dan uji menyebabkan kesalahan klasifikasi. Penemuan ini menunjukkan bahwa meskipun CNN efektif dalam klasifikasi sentimen, ada tantangan yang perlu diatasi untuk meningkatkan akurasi, terutama dalam memahami konteks bahasa yang kompleks.

Kata Kunci: Pemrosesan Bahasa Alami;NLP;Sentiment Analysis;python; 1D Convolutional Neural Network

Abstract

This study explores the use of natural language processing (NLP) to analyze sentiments in tweets about young leadership in presidential elections. Data collected from Twitter is processed and extracted. The tokenization process generates word tokens, which are then classified using the Convolutional Neural Network (CNN) algorithm. The study aims to compare the accuracy of the CNN algorithm in sentiment classification. Data is categorized into two classes: positive and negative, after being collected from Twitter. Pre-processing is performed using NLP techniques before tokenization. The generated tokens are then input into the CNN model for classification. The 1D CNN model used consists of one convolutional layer and one pooling layer, with ReLU activation on the convolutional layer. The dataset consists of 200 samples, divided into 160 training samples and 40 testing samples. The results show that the predicted labels match the actual labels of new sentences. However, factors such as words with implicit or ambiguous meanings, the use of informal language, inconsistent spelling, and the similarity of words in the training and testing data cause classification errors. These findings indicate that while CNN is effective in sentiment classification, there are challenges to be addressed to improve accuracy, especially in understanding complex language contexts.

Keywords: Natural Language Processing; NLP; Sentiment Analysis; python; 1D Artificial Neural Network

Djtechno: Jurnal Teknologi Informasi oleh Universitas Dharmawangsa Artikel ini bersifat open access yang didistribusikan di bawah syarat dan ketentuan dengan Lisensi Internasional Creative Commons Attribution NonCommercial ShareAlike 4.0 ([CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/)).



Corresponding Author:

E-mail : ozalitoyib@umb.ac.id

1. PENDAHULUAN

Sebagai kelompok generasi yang signifikan dalam sejarah demografi manusia, Generasi Z telah menjadi subjek banyak studi yang mencoba memahami cara mereka berpikir dan berperilaku. Kedua memengaruhi keputusan sosial dan politik, termasuk pemilihan umum, di berbagai negara. Sejak awal abad ke-21, keduanya telah menjadi subjek penelitian karena perbedaan nilai-nilai, preferensi, dan perilaku mereka, yang memengaruhi dinamika politik.

Kelompok generasi yang signifikan dalam sejarah demografi manusia, Generasi Z dan Gen Z telah menjadi subjek banyak penelitian yang mencoba memahami pola pikir dan tindakan mereka. Keduanya sangat memengaruhi keputusan sosial dan politik, termasuk pemilihan umum, di banyak negara. Kedua telah menjadi subjek penelitian yang menarik sejak awal abad ke-21 karena perbedaan nilai-nilai, preferensi, dan perilaku mereka, yang berdampak pada dinamika politik (Annisa Fitriyani¹, 2024). Namun, sedikit penelitian yang membandingkan persepsi dan perilaku generasi Z dan milenial dalam konteks politik.

Sangat penting untuk memahami bagaimana perilaku dan persepsi generasi milenial dan Z mempengaruhi keputusan mereka saat memilih di pemilihan umum tahun 2024. Dengan pergeseran demografis yang semakin jelas menuju generasi ini sebagai pemilih utama, sangat penting bagi para kandidat dan partai politik yang bersaing untuk memahami bagaimana preferensi politik mereka dipengaruhi.

Sebagian orang percaya bahwa milenial dan generasi Z lebih idealis dan terbuka terhadap perubahan. Di sisi lain, generasi Z biasanya lebih pragmatis dan menggunakan teknologi untuk membuat keputusan politik (Simamora et al., 2024). Namun, penelitian lebih lanjut diperlukan untuk membuatnya lebih jelas. Perubahan teknologi, perubahan sosial, dan lingkungan politik dan budaya adalah beberapa faktor utama yang memengaruhi persepsi dan perilaku politik generasi milenial dan Z. perlu dipikirkan. Pandangan politik generasi ini sangat dipengaruhi oleh globalisasi, penyebaran media sosial, dan masalah seperti kesenjangan ekonomi dan perubahan iklim (Bulan Oktober & Al Madinah F, 2024). Generasi milenial dan Z tampaknya lebih aktif secara sosial dan politik ketika mereka berpartisipasi dalam kampanye, protes, dan diskusi online (Rehman et al., 2019). Namun, masih belum jelas apakah aktivisme ini berkorelasi dengan keputusan pemilihan umum mereka. secara sosial dan politik dalam hal mengambil bagian dalam kampanye, protes, dan diskusi online (Listyarini & Anggoro, 2021). Namun, masih belum jelas apakah aktivisme ini sejalan dengan keputusan pemilihan umum mereka.

Sentiment Analysis adalah metode untuk mendapatkan Perspektif dan informasi khusus yang terdapat dalam kalimat dapat diperoleh melalui analisis emosi. Ini memiliki kemampuan untuk menentukan apakah penulis memiliki perasaan terhadap suatu objek atau subjek, apakah itu negatif, atau positif (Rajput et al., n.d.). Dalam proses analisis ini, komputer juga digunakan untuk menemukan perasaan yang relevan dalam kalimat. Salah satu cara komputer memahami input bahasa manusia adalah proses pemrosesan bahasa alam, juga dikenal sebagai NLP (Kim & Jeong, 2019). Proses ini memungkinkan komputer untuk menerjemahkan bahasa alami manusia ke dalam bentuk yang dapat dipahami mesin (Zeng et al., 2020). Menurut penelitian yang dilakukan oleh Young-Seob Jeong dan Hannah Kim, algoritma Convolution Neural Network (CNN) dapat menggolongkan kalimat dengan sangat baik, yang membuatnya sangat efektif dalam pengolahan bahasa natural (S. Singh & Sambyal, 2023). Penelitian ini menyarankan menggunakan Natural untuk menganalisis perasaan tweet pilpres keterpimpinan muda. Proses pengolahan bahasa digunakan untuk mengolah dan mengekstrak dataset dari Twitter. Tokenisasi ini menghasilkan token kata yang akan

diklasifikasikan dengan algoritma Convolutional Neural Network (CNN). Hasil penelitian ini dapat digunakan untuk membandingkan akurasi dan dampak dari jumlah epoch yang digunakan. Selain itu, hasilnya dapat digunakan untuk menguji label prediksi kalimat terhadap kalimat baru yang ditambahkan. pilpres keterpimpinan muda menggunakan pengolahan bahasa alami untuk mengolah dan mengekstrak data dari Twitter. Tokenisasi ini menghasilkan token kata yang akan diklasifikasikan dengan algoritma Convolutional Neural Network (CNN).

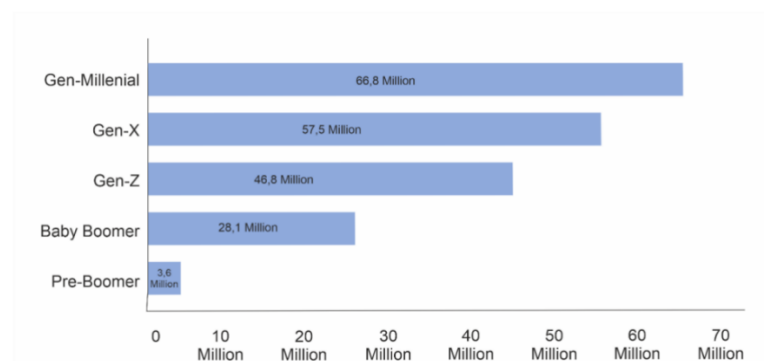
Berdasarkan riset reverensi yang sudah dilangsungkan, mendapatkan hasil bahwa masih minim penelitian yang meneliti jumlah epoch yang dipakai pada model CNN. Sebuah epoch adalah jumlah kali model neural network melihat kumpulan data. Satu epoch adalah jumlah kali model neural network melihat kumpulan data. Karena itu, satu epoch terlalu besar untuk dimasukkan ke dalam komputer dan oleh karena itu perlu dibagi menjadi batch kecil, atau epoch (Yang et al., 2019). Studi reverensi yang sudah dilakukan menunjukkan bahwa masih sedikit penelitian yang mempelajari jumlah epoch yang digunakan pada model neural network. Oleh karena itu, pengujian yang dilakukan dengan data yang berbeda mungkin tidak akurat (Pangestu & Rosyda, 2022).

Penelitian ini mengusulkan penggunaan Natural Language Processing untuk melakukan analisis sentiment tweet pilpres keterpimpinan muda. Data Twitter diolah dan diekstrak. Tokenisasi ini menghasilkan token kata yang akan diklasifikasikan dengan algoritma Convolutional Neural Network (CNN). Hasil penelitian ini dapat digunakan untuk membandingkan akurasi dan dampak dari jumlah epoch yang digunakan. Selain itu, hasilnya dapat digunakan untuk menguji label prediksi kalimat terhadap kalimat baru yang ditambahkan.

Daftar pemilih tetap (DPT) yang dikeluarkan oleh Komisi Pemilihan Umum (KPU) untuk Pemilu 2024 berjumlah 204.807.222 pemilih; ini terdiri dari 46.800.161 pemilih, atau 22,85 persen, dan 66.822.389 pemilih, atau 33,60 persen dari generasi milenial. dari generasi Z, yang masing-masing menyumbang 56,45 persen dari total pemilih. Gen Z dan milenial memiliki kemampuan untuk membentuk perubahan dan berkontribusi

positif pada pengambilan keputusan, menurut angka-angka tersebut. Banyak dari mereka berpendapat bahwa demokrasi hanya digunakan untuk menguntungkan kelompok tertentu. Sebenarnya, setiap orang memiliki hak untuk memilih atau golput, tetapi sebagai warga negara, Gen Z dan milenial yang sudah cukup umur harus menggunakan hak pilihnya untuk memastikan pemilu 2024 berhasil.

Bagan menunjukkan bahwa Gen Z mewakili 46.8 persen pemilih di Indonesia, yang menunjukkan kemungkinan mereka berada di salah satu dari dua kelompok tersebut. jumlah yang sangat besar. Karena mereka aktif mengikuti perkembangan berita melalui media digital, Gen Z mungkin adalah pemilih yang menyadari pilihan mereka, tetapi mereka mungkin juga tidak peduli karena pemistis mereka terhadap berbagai propaganda yang populer. Karena Gen Z memiliki pengaruh dalam pemilihan presiden yang akan datang, mereka harus didukung dengan kesadaran politik.



Gambar 1. Persentase Gen-z 2024

2. METODE PENELITIAN

Cara mengklasifikasikan sentimen dalam teks adalah masalah utama dalam klasifikasi sentimen. Sulit untuk mengetahui apakah sentimen tersebut positif atau negatif. Proses penelitian ini digambarkan pada Gambar 1 (Rhanoui et al., 2019). Pertama, data dikumpulkan (atau diambil dari Twitter), dan kemudian diklasifikasikan menjadi kelas positif dan negatif. Selanjutnya, dalam proses pre-processing, teknik natural language processing (NLP) digunakan. Setelah proses pre-processing selesai, tokenizing dibuat, yang dimasukkan ke dalam proses klasifikasi. Proses klasifikasi menggunakan algoritma Convolutional Neural Network (CNN) atau algoritma deep learning.



Gambar 2. Alur penelitian

Alur penelitian berdasarkan pada Gambar 1. Lebih rinci dijelaskan sebagai berikut:

- 1) *Scrapping Data Twitter* :Dataset yang digunakan dalam penelitian dikumpulkan dari Twitter melalui scraping yang digunakan oleh pemrograman Python. Selain itu, library twint digunakan untuk mengumpulkan data melalui proses ekstrasi informasi yang dilakukan melalui Protokol Transfer Hypertext (HTTP). Hasil scraping data termasuk 193 tweet yang dikumpulkan dengan kata kunci pencarian seperti "genz", "generasi z", "pilpers", dan "pemimpin muda", seperti yang ditunjukkan dalam Tabel 1. Data Twitter yang dikumpulkan secara acak digunakan untuk mengacu pada pemilihan kepala daerah.

Tabel 1.Data tweet yang dikumpulkan secara acak

Tanggal,Waktu	Tweet
ThuDec1403:38:48 +0000 2023	kalo pencalonannya dengan cara yang baik dan beretika bolehlah...tapi ini capres cacat etik.
Fri Dec 15 23:58: +0000 2023	Pertarungan pilpres bukan hanya dimedsos.Polling diberbagai media platform amin menang.
Mon Dec 18 00:57:30 +0000 2023	Semoga pemilu kali ini kembali jujur adil bermartabat.
Mon Dec 18 00:27:32 +0000 2023	JANGAN FANATIK berani memuji kandidat lain setiap orang punya kelebihan.

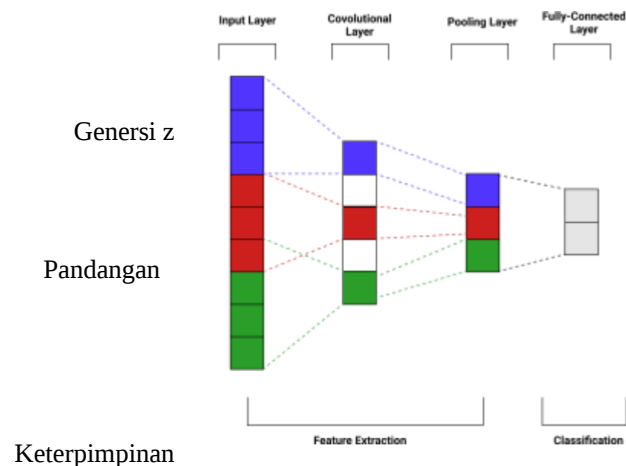
- 2) *Labelling Data* : Dataset dilabeli secara manual dan diberi label positif dan negative. 98 tweet dengan sentimen positif dan 73 dengan sentimen negatif ditemukan oleh kedua belah pihak.
- 3) *Tokenizing* : Token memotong teks dalam proses operasi (T. Singh & Kumari, 2016). Potongan-potongan ini dapat berupa kata, huruf, atau kalimat. Tabel 2 menggambarkan proses tokenisasi.

Tabel 2. Contoh dari Proses Tokenizing

Teks Asli	“sudut pandang genz keterpimpinan muda”
Sesudah Tokenisasi	1. 'sudut' 2. 'pandang' 3. 'genz' 4. 'keterpimpinan' 5. 'muda'

- 4) *Sequences* : suatu prosedur setelah tokenisasi di mana jika terjadi masalah selama proses tokenisasi di mana kata-kata dengan jumlah dan huruf yang sama, seperti "amal" dan "lama", mesin akan membaca bahwa kata-kata tersebut memiliki jumlah dan huruf yang sama, sehingga kita membutuhkan sequences, yang akan menampung semua token yang ada dalam bentuk array.
- 5) *Klasifikasi* : Setelah proses tokenizing dan sequencing, algoritma Convolutional Neural Network—skema deep learning—digunakan untuk mengklasifikasikan data. Salah satu bagian Artificial Neural Network (ANN) adalah Convolutional Neural Network, yang memiliki kemampuan untuk memperhitungkan data dengan ketepatan tinggi (Pambudi & Suprpto, 2021). Model Convolutional Neural Network banyak menyelesaikan masalah pemrosesan gambar, tetapi sekarang telah dikembangkan untuk digunakan dalam pemrosesan bahasa alami, seperti analisis perasaan, kesimpulan teks, klasifikasi polaritas emosional, dan lainnya (Nyoman & Putu Kusuma Negara, 2021). Dalam proses CNN ini, tweet yang mengandung kata-kata "genz", "pemimpin", dan "pilpres" digunakan

sebagai input untuk lapisan convolution. Setelah proses convolutional selesai, hasilnya akan dimasukkan ke dalam aktivasi ReLU. Fungsi aktivasi ReLU mengalami proses down sampling pada Lapisan Penyimpanan, dan hasil dari Pooling Layer. Selanjutnya, Fully Connected Layer menggunakan hasil dari Pooling Layer sebagai input. Gambar 2 menunjukkan arsitektur CNN untuk penelitian ini.



Gambar 3. Arsitektur CNN

Convolutional Neural Network (CNN) adalah teknik pembelajaran mesin yang mengembangkan Perceptron Multilayer. Karena itu, jenis jaringan multilayer CNN biasanya disebut sebagai layer penuh terhubung karena hasil dari satu layer akan digunakan untuk layer berikutnya. (Taradhita & Putra, 2021). Layer Convolutional Neural Network yang sederhana digunakan dalam penelitian ini. Lapisan ini terdiri dari lapisan convolutional, lapisan max pooling, lapisan dropout, dan lapisan output yang benar-benar terhubung. Ditetapkan sejumlah 8 unit convolutional dan ukuran vector 300 pada setting model studi ini. Filter 16, 32, dan 64 digunakan pada delapan unit convolutional layer. Pada penelitian ini, Adam dipilih sebagai Optimizer (Taradhita & Putra, 2021) karena dia dianggap sebagai yang terbaik. Aktivasi ReLU, juga dikenal sebagai Unit Rectified Linear, digunakan untuk aktivasi non-linear karena memiliki kelebihan kerumitan yang rendah.

3. HASIL DAN PEMBAHASAN

Studi ini menggunakan 193 dataset tweet. Dataset yang dikumpulkan dibagi menjadi tiga kategori: satu memiliki sentimen positif, dan dua memiliki sentimen negatif. Labelling atau pembagian data ditunjukkan dalam Tabel 3.

Tabel 3. Hasil Proses Labeling Data

1 (Positif)	2(Negatif)
125	70

Proses tokenizing menghasilkan 1.703 kata atau token dari 200 tweet, dan dataset tweet kalimat dibagi menjadi per kata pada tahap ini, sehingga lebih mudah untuk mengelola teks berikutnya.

```
[522, 4, 9, 25, 2, 1, 24, 523, 158, 1, 173, 1, 195, 293, 33, 2, 105, 43, 292, 2, 113, 147, 109, 293, 334, 145, 170, 2, 225, 69, 233, 8, 33, 793, 2, 105, 41, 293, 213, 89, 2, 235, 8, 41, 23, 163, 26, 1, 177, 29, 654, 192, 2, 113, 147, 55, 6, 93, 22, 7, 174, 29, 78, 2, 22, 7, 164, 89, 3, 32, 123, 23, 163, 113, 45, 189, 2, 173, 25, 2, 1, 158, 1, 17, 3, 17, 19, 207, 617, 250, 143, 1, 46, 11, 2, 7, 3, 253, 21, 120, 44, 2, 7, 60, 42, 37, 346, 175, 2, 9, 1, 3, 4, 5, 620, 2, 66, 1, 42, 15, 104, 50, 4, 6, 612, 2, 113, 254, 651, 74, 15, 364, 625, 4, 5, 6, 646, 64, 42, 15, 27, 257, 9, 91, 352, 105, 27, 47, 106, 88, 144, 35, 335, 88, 89, 90, 356, 357, 89, 38, 90, 50, 4, 5, 6, 623, 4, 5, 6, 624, [254, 625, 105, 106, 89, 48, 47, 69, 358, 359, 360, 1, 10, 3, 253, 21, 46, 11, 2, 0, 211, 626, 4, 5, 6, 627], [15, 60, 42, 32, 2, 7, 73, 23, 1, 3, 638, 639, 4, 5, 6, 640, 26, 641, 180, 59, 1, 109, 3, 72, 20, 40, 36, 4, 5, 6, 646, 64, 7, 26, 256, 4, 5, 6, 648], [60, 42, 32, 2, 7, 8, 31, 1, 649, 4, 5, 650], [254, 651, 74, 15, 364, 652, 231, 60, 42, 15, 257, ], [9, 365, 104, 366, 37, 32, 7, 73, 110, 1, 10, 3], [60, 42, 15, 104, 37, 32, 2, 9, 1, 3, 4, 5, 6, 653, 35, 356, 654], [74, 367, 655, 368, 656, 657, 658, 52, 659, 660, 181, 369, ,661, 35, 258, 214, 209, 16, 662, 56, 663, 21, 53, 3, 664, 75, 121, 259, 214, 22, 4, 5, 6, 665, 2, ,666, 1, 667, 668, 669, 670, 26, 671, 1, 10, 260, 53, 3, 95, 181, 87, 672, 18, 149, 673, 674, 2, 7, 675, 370, 676, 261, 123, 371, 677, 9, 123, 124, 678, 679, 174, 680, 262, 681, 4, 5, 6, 682], [73, 23, 1, 3, 33, 17, 52, 3, 0, 25, 2, 7, 4, 5, 6, 683], [15, 60, 42, 32, 2, 7, 31, 1, 10, 3, 4, 5, 6, 684], [372, 685, 686, 687, 689, 691, 1, 1, 0, 3, 120, 373, 692, 18, 149, 125, 2, 7, 9, 86, 374, 693, 694, 4, 5, 6, 695], [87, 351, 1, 3, 123, 696, 263, 2, 375, 17, 38, ]
```

Gambar 4. Hasil Proses Tokenizing

Setelah tokenizing, setiap bentuk kata diubah menjadi bentuk sequentials yang disimpan dalam bentuk array. Karena jumlah kata dalam setiap kata tidak sama, proses normalisasi digunakan untuk membuat jumlah kalimatnya sesuai, artinya jika ada kata yang lebih banyak akan dipotong dan jika ada kata yang kurang akan ditambahkan padding. Jumlah total kata yang dapat dinormalisasi adalah seratus kata.

```
[522, 4, 9, 25, 2, 1, 24, 523, 158, 1, 173, 1, 195, 293, 33, 2, 105, 43, 292, 2, 113, 147, 109, 293, 334, 145, 170, 2, 225, 69, 233, 8, 33, 793, 2, 105, 41, 293, 213, 89, 2, 235, 8, 41, 23, 163, 26, 1, 177, 29, 654, 192, 2, 113, 147, 55, 6, 93, 22, 7, 174, 29, 78, 2, 22, 7, 164, 89, 3, 32, 123, 23, 163, 113, 45, 189, 2, 173, 25, 2, 1, 158, 1, 173, 17, 19, 207, 617, 250, 143, 1, 46, 11, 2, 7, 3, 253, 21, 120, 44, 2, 7, 60, 42, 37, 346, 175, 2, 9, 1, 3, 4, 5, 620, 2, 66, 1, 42, 15, 104, 50, 4, 6, 612, 2, 113, 254, 651, 74, 15, 364, 625, 4, 5, 6, 646, 64, 42, 15, 27, 257, 9, 91, 352, 105, 27, 47, 106, 88, 144, 35, 335, 88, 89, 90, 356, 357, 89, 38, 90, 50, 4, 5, 6, 623, 4, 5, 6, 624, [254, 625, 105, 106, 88, 48, 47, 69, 358, 359, 360, 1, 10, 3, 253, 21, 46, 11, 2, 7, 211, 626, 4, 5, 6, 627], [15, 60, 42, 32, 2, 7, 73, 23, 1, 3, 638, 639, 4, 5, 6, 640, 26, 641, 180, 59, 1, 109, 3, 72, 20, 40, 36, 4, 5, 6, 646, 64, 7, 26, 256, 4, 5, 6, 648], [60, 42, 32, 2, 7, 8, 31, 1, 649, 4, 5, 650], [254, 651, 74, 15, 364, 652, 231, 60, 42, 15, 257, ], [9, 365, 104, 366, 37, 32, 7, 73, 110, 1, 10, 3], [60, 42, 15, 104, 37, 32, 2, 9, 1, 3, 4, 5, 6, 653, 35, 356, 654], [74, 367, 655, 368, 656, 657, 658, 52, 659, 660, 181, 369, ,661, 35, 258, 214, 209, 16, 662, 56, 663, 21, 53, 3, 664, 75, 121, 259, 214, 22, 4, 5, 6, 665, 2, ,666, 1, 667, 668, 669, 670, 26, 671, 1, 10, 260, 53, 3, 95, 181, 87, 672, 18, 149, 673, 674, 2, 7, 675, 370, 676, 261, 123, 371, 677, 9, 123, 124, 678, 679, 174, 680, 262, 681, 4, 5, 6, 682], [73, 23, 1, 3, 33, 17, 52, 3, 0, 25, 2, 7, 4, 5, 6, 683], [15, 60, 42, 32, 2, 7, 31, 1, 10, 3, 4, 5, 6, 684], [372, 685, 686, 687, 689, 691, 1, 0, 3, 120, 373, 692, 18, 149, 125, 2, 7, 9, 86, 374, 693, 694, 4, 5, 6, 695], [87, 351, 1, 3, 123, 696, 263, 2, 375, 17, 38, ]
```

(a)

```
Array([[ 522, 523, 33, ..., 0, 0, 0],
       [ 595, 78, 44, ..., 0, 0, 0],
       [ 596, 354, 123, ..., 0, 0, 0],
       ...,
       [1089, 145, 78, ..., 0, 0, 0],
       [ 6, 93, 7, ..., 0, 0, 0],
       [ 414, 1, 334, ..., 0, 0, 0],
```

Gambar 5. Contoh Hasil Proses Sequences, Sebelum Sequences

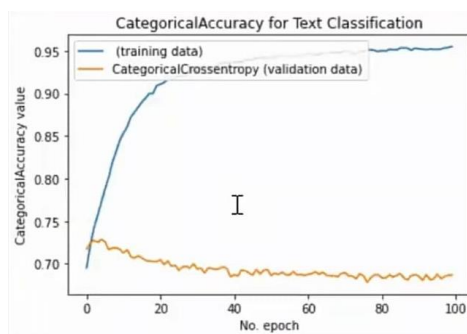
Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 128)	640000
conv1d (Conv1D)	(None, 96, 128)	82048
global_max_pooling1d (GlobalMaxPooling1D)	(None, 128)	0
dense (Dense)	(None, 128)	16512
dense_1 (Dense)	(None, 3)	387

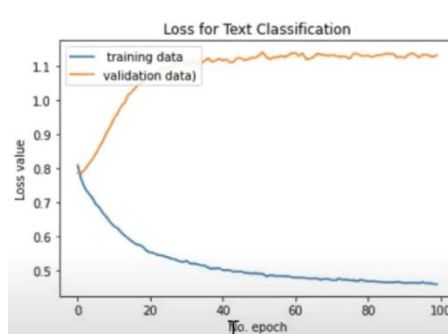
=====
Total params: 738947 (2.82 MB)
Trainable params: 738947 (2.82 MB)
Non-trainable params: 0 (0.00 Byte)
=====

Gambar 6. Model Latih

Data yang dibuat untuk model terdiri dari dua kelas (positif, dan negatif) dan akan digunakan sebagai dataset untuk mengajarkan model 1D CNN yang akan dibangun. Dalam penelitian ini, model 1D CNN menggunakan satu proses lapisan convolution dan satu proses lapisan pooling, di mana convolution dilakukan dengan aktivasi ReLU. Data yang dibuat untuk pelatihan dibagi menjadi 160 data pelatihan dan 40 data tes.



(a)



(b)

Akurasi terbaik adalah 73,11%, dengan kehilangan 0,9187 pada epoch ke-25, seperti yang ditunjukkan dalam tabel 4 dari hasil pengujian data tes. Tabel tersebut menunjukkan bahwa semakin banyak epoch, semakin akurat. Peningkatan akurasi ini berkisar antara 7–12% untuk berbagai epoch. Dalam proses pelatihan, penggunaan jumlah epoch yang berbeda menyebabkan hasil klasifikasi yang berbeda. Dengan epoch yang lebih sedikit, akurasi klasifikasi lebih rendah karena proses adaptasi parameter belum sempurna karena waktu iterasi yang lebih singkat. Tetapi seiring waktu, jumlah epoch meningkat, parameter menjadi lebih baik dan hasil klasifikasi menjadi lebih baik.

Menurut tabel 4, hasil pengujian data tes menunjukkan bahwa akurasi terbaik adalah 73,11% dengan kehilangan 0,9187 pada epoch ke-25. Tabel tersebut menunjukkan bahwa semakin banyak epoch, semakin besar akurasi; peningkatan ini berkisar antara 7–12% untuk masing-masing epoch yang disebutkan sebelumnya.

Penggunaan jumlah epoch yang berbeda dalam proses pelatihan menghasilkan kinerja klasifikasi yang bervariasi. Ketika epoch yang digunakan sedikit, akurasi yang diperoleh cenderung rendah karena parameter model belum teradaptasi dengan sempurna akibat waktu iterasi yang terbatas. Namun, saat jumlah epoch ditingkatkan, parameter menjadi lebih optimal dan kinerja klasifikasi secara bertahap meningkat.

Tabel.4Komparasi Akurasi Konvolusi Pada *Epoch* yang Berbeda

Accuracy Loss	
Epoch 10	0.6410(64%) 0.8447
Epoch 15	0.5927(59,27%)

	0.8839
<i>Epoch 20</i>	0.5572(69.72%)
	0.9092
<i>Epoch 25</i>	0.5411(73,11%)
	0.9189

Uji klasifikasi dilakukan pada kalimat baru yang tidak ada dalam dataset untuk mengetahui apakah label prediksi sesuai dengan label aslinya. Tabel 5 menunjukkan temuan tersebut.

TABEL 5 Perbandingan Pada Length text Dan Label sentiment

NO	Tweet	Length text	Sentiment
1	Capres Cuma jago ngomong,tapi minin prestasi dan gagak dalam bekerja.	345	Negatif
2	Setelah melihat depot capres semalam jadi bingung milih siapa,semua kandidat.	228	Positif
3	Ya Allahkabulkan do'a kami agar beliau menjadi pemimpin negeri.	224	Positif
4		330	Negatif
5	Kita paham dan mengerti,mungkin orang yang susah mengerti bapak,makanya beliau banyak fitnah.	231	Negatif
	Ga ada survey yang indenpenden semua di intervensi oleh kekuasaan.		

Model diuji dengan kalimat baru di Tabel 5. Hasilnya menunjukkan bahwa panjang teks pada kalimat baru sesuai dengan label sentiment. Namun, kalimat baru yang dilabelkan negatif sebenarnya merupakan sentimen positif. Kalimat dengan

makna ambigu atau tersirat, penggunaan bahasa informal, ejaan yang tidak konsisten, dan kemiripan kata dalam data pelatihan dan tes dapat menyebabkan kesalahan klasifikasi ini.

4. SIMPULAN

Dengan epoch yang digunakan sebanyak 25 epoch, penelitian "Sentiment Analysis Sudut Pandang Generasi Babyz Terhadap Keterwakilan kepemimpinan Muda Pada Pilpres 2024 di Twittwer Menggunakan ID Convulation Neural Network" mencapai tingkat akurasi tertinggi sebesar 73,11%. Seiring bertambahnya jumlah epoch yang diterapkan, pengujian menjadi lebih akurat. Akurasi meningkat sebesar 7-12% dari variasi epoch sebelumnya dari epoch 10, 15, 20, dan 25. Hasil uji klasifikasi kalimat baru menunjukkan bahwa label prediksi sesuai dengan label isi kalimat tersebut. Namun, beberapa faktor menyebabkan kesalahan klasifikasi; ini termasuk kata-kata dengan makna tersiratkan atau ambigu, penggunaan bahasa informal, ejaan yang tidak konsisten, dan kemiripan kata dalam data pelatihan dan tes.

UCAPAN TERIMA KASIH

Saya ingin menyampaikan terima kasih yang sebesar-besarnya kepada dosen pembimbing saya atas bimbingan, ilmu, dan dukungan yang tak ternilai selama proses studi ini. Kepada keluarga saya, terima kasih atas cinta, doa, dan dukungan tanpa henti yang selalu menjadi sumber kekuatan bagi saya. Untuk teman-teman saya, terima kasih atas kebersamaan, semangat, dan bantuan yang diberikan sepanjang perjalanan ini. Terakhir, kepada idola saya, terima kasih atas inspirasi dan motivasi yang telah mendorong saya untuk terus berusaha dan menyelesaikan penelitian ini.

PUSTAKA

- Annisa Fitriyani¹, A. D. R. , F. L. , D. P. A. I. (2024). *DETERMINAN PERSEPSI PERILAKU GENERASI MILENIAL DENGAN GENERASI Z MEMPENGARUHI KEPUTUSAN DALAM PEMILIHAN UMUM TAHUN 2024*. Bulan Oktober, P., & Al Madinah F, S. M. (2024). PT. Media Akademik Publisher ANALISIS FRAMING PEMBERITAAN CALON PRESIDEN 2024 DIMASA MENUJU KAMPANYE DALAM PEMILU 2024 PADA MEDIA JAWAPOS.COM DAN KOMPAS.COM. In JMAJ (Vol. 2, Issue 1).
- Kim, H., & Jeong, Y. S. (2019). Sentiment classification using Convolutional Neural Networks. *Applied Sciences (Switzerland)*, 9(11). <https://doi.org/10.3390/app9112347>
- Listyarini, S. N., & Anggoro, D. A. (2021). Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network (CNN). *Jurnal Pendidikan Dan Teknologi Indonesia*, 1 (7), 261–268. <https://doi.org/10.52436/1.jpti.60>

- Nyoman, P., & Putu Kusuma Negara. (2021). Deteksi Masker Pencegahan Covid19 Menggunakan Convolutional Neural Network Berbasis Android. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 576–583. <https://doi.org/10.29207/resti.v5i3.3103>
- Pambudi, A., & Suprpto, S. (2021). Effect of Sentence Length in Sentiment Analysis Using Support Vector Machine and Convolutional Neural Network Method. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(1), 21. <https://doi.org/10.22146/ijccs.61627>
- Pangestu, G. T., & Rosyda, M. (2022). Sentiment Analysis Tweet Pilkada 2020 Saat Pandemi COVID-19 di Media Sosial Twitter Menggunakan Metode 1D Convolutional Neural Network. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(2), 1017. <https://doi.org/10.30865/mib.v6i2.3765>
- Rajput, A., Nazlah Al Yamaniyyah, A., & Arabia, S. (n.d.). *Natural Language Processing, Sentiment Analysis and Clinical Analytics*.
- Rehman, A. U., Malik, A. K., Raza, B., & Ali, W. (2019). A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis. *Multimedia Tools and Applications*, 78(18), 26597–26613. <https://doi.org/10.1007/s11042-019-07788-7>
- Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BiLSTM Model for Document-Level Sentiment Analysis. *Machine Learning and Knowledge Extraction*, 1(3), 832–847. <https://doi.org/10.3390/make1030048>
- Simamora, I. Y., Azra, A., Nasution, M., Novita, D. D., Syahira, Z., Sakila Nazwa, W., & Siregar, R. A. (2024). *Peran Generasi Z dalam Pemilu 2024 di Indonesia*.
- Singh, S., & Sambyal, N. (2023). *Automated Glaucoma Detection Using Deep Convolutional Neural Networks*. <https://doi.org/10.21203/rs.3.rs-2788554/v1>
- Singh, T., & Kumari, M. (2016). Role of Text Pre-processing in Twitter Sentiment Analysis. *Procedia Computer Science*, 89, 549–554. <https://doi.org/10.1016/j.procs.2016.06.095>
- Taradhita, D. A. N., & Putra, I. K. G. D. (2021). Hate speech classification in Indonesian language tweets by using convolutional neural network. *Journal of ICT Research and Applications*, 14(3), 225–239. <https://doi.org/10.5614/itbj.ict.res.appl.2021.14.3.2>
- Yang, Z., Gao, M., Zhang, J., Ma, Y., Chen, W., Yan, Y., & Shen, G. (2019). Unsupervised Neural Network for Modulation Format Discrimination and Identification. *IEEE Access*, 7, 70077–70087. <https://doi.org/10.1109/ACCESS.2019.2916806>
- Zeng, Q., Ma, X., Cheng, B., Zhou, E., & Pang, W. (2020). GANS-based data augmentation for citrus disease severity detection using deep learning. *IEEE Access*, 8, 172882–172891. <https://doi.org/10.1109/ACCESS.2020.3025196>