

SENTIMENT ANALYSIS OF INDONESIAN ELECTION 2024 USING THE K-NEAREST NEIGHBOR METHOD

Rido Dwi Kurniawan^{*1}, Joshua Muliawan²

^{1,2}Business Information System Program, Faculty of Science and Technology, Universitas Pradita, Indonesia
Email: ¹rido.Dwi@pradita.ac.id, ²joshua.muliawan@student.pradita.ac.id

(Article received: March 15, 2024; Revision: April 08, 2024; published: May 24, 2024)

Abstract

The abstract of this research discusses the analysis of Indonesian public sentiment regarding the 2024 election as observed via Twitter. Sentiment graph Research uses the Natural Language Processing method and the K-Nearest Neighbor algorithm to classify sentiment as positive, neutral, or negative. The current era of globalization influences the rapid progress of information technology circulating in society, one of the intermediaries is through the social media Twitter. Twitter can be used as a means of conveying opinions regarding suggestions, criticism and public opinion. Currently social media has a big impact on building public political sentiment and preferences. The social media I took is Twitter so that people's Tweets related to elections can be used to see a picture of public opinion. There are various opinions of Twitter users with positive, neutral and negative sentiments. However, classifying sentiment from Twitter users requires quite a lot of time and effort due to the large number of tweets found. The aim of this research is to conduct a public sentiment analysis of public opinion regarding the 2024 election. Data was collected in October and December 2023. The results show that positive sentiment dominates with 76%, followed by neutral sentiment at 16%, and negative 6%. This analysis helps understand public opinion regarding the 2024 election on social media, especially Twitter.

Keywords: Sentiment Analysis, K-Nearest Neighbor, 2024 Election, Presidential Election, Twitter.

ANALISIS SENTIMEN PEMILU INDONESIA 2024 MENGGUNAKAN METODE K-NEAREST NEIGHBOR

Abstrak

Abstrak penelitian ini membahas analisis sentimen publik Indonesia mengenai pemilu 2024 yang diamati melalui Twitter. Grafik sentimen Penelitian menggunakan metode Natural Language Processing dan algoritma K-Nearest Neighbor untuk mengklasifikasikan sentimen sebagai positif, netral, atau negatif. Era globalisasi saat ini mempengaruhi pesatnya kemajuan teknologi informasi yang beredar di masyarakat, salah satu perantaranya melalui media sosial Twitter. Twitter dapat digunakan sebagai sarana penyampaian pendapat mengenai saran, kritik dan opini masyarakat. Saat ini media social sangat berdampak besar membangun sentiment dan preferensi politik publik, media social yang saya ambil Twitter sehingga Tweet masyarakat yang berkaitan dengan pemilu dapat digunakan untuk melihat gambaran opini masyarakat. Terdapat berbagai opini pengguna twitter yang bersentimen positif, netral dan negatif. Namun dalam mengklasifikasi sentimen dari pengguna twitter membutuhkan waktu dan usaha yang cukup banyak dikarenakan banyaknya jumlah tweet yang ditemukan. Adapun tujuan dari penelitian ini adalah untuk melakukan analisis sentimen publik terhadap opini masyarakat terkait pemilu 2024. Data dikumpulkan pada Oktober dan Desember 2023. Hasil menunjukkan bahwa sentimen positif mendominasi dengan 76%, diikuti sentimen netral 16%, dan negatif 6%. Analisis ini membantu memahami opini publik terkait pemilu 2024 di media sosial, khususnya Twitter.

Kata kunci: Analisis Sentimen, K-Nearest Neighbor, Pemilu 2024, Pilpres, Twitter.

1. PENDAHULUAN

Analisis sentimen merupakan sesuatu proses komputasi dengan menguasai, mengekstrak serta mencerna informasi tekstual secara otomatis buat memperoleh suatu data sentimen yang ada dalam sesuatu kalimat opini ataupun komentar, sikap serta

emosi seorang terhadap entitas (menggambarkan orang, peristiwa ataupun topik), tujuannya sediakan data berharga untuk seorang yang tercantum dari suatu dataset yang tidak terstruktur [1].

Analisis sentimen merupakan salah satu teknik untuk mengekstrak sebuah informasi berupa sikap

seseorang terhadap isu atau kejadian dengan mengelompokkan polaritas dari sebuah teks [2]. Analisis sentimen berguna dalam mengembangkan sistem untuk menganalisis, mengidentifikasi, dan mengungkapkan pendapat dalam bentuk teks dan juga merupakan proses yang berfungsi untuk mengidentifikasi berupa opini atau sentimen dari isi dataset, berupa teks tentang topik atau peristiwa positif dan negatif. Informasi berbentuk teks kini tersebar luas di internet dalam bentuk blog, forum, sosial media, dan situs berisi ulasan. Sentimen positif atau negatifnya sebuah opini bisa diproses dengan cara manual, namun tentunya semakin besar jumlah sumber opini akan bertambah banyak pula waktu dan tenaga yang diperlukan dalam pengklasifikasian polaritas opini tersebut. Maka dari itu, diusulkan penerapan metode pembelajaran mesin untuk mengklasifikasikan polaritas opini dari sumber data yang berjumlah banyak dengan memanfaatkan fungsi dari text mining [3].

Pemilu merupakan sebuah proses demokrasi yang diselenggarakan setiap lima tahun sekali oleh Negara Indonesia. Demokrasi ditandai dengan diadakannya pemilihan calon legislatif, kepala daerah, hingga kepala negara yang langsung dipilih oleh rakyat. Indonesia telah mengadakan pemilihan umum sejak tahun 1955 yang diikuti lebih dari tiga puluh partai politik yang berbeda [4].

Pemilu dimaknai sebagai prosedur untuk mencapai demokrasi atau merupakan prosedur untuk memindahkan kedaulatan rakyat kepada kandidat tertentu untuk menduduki jabatan-jabatan politik. Pemilu hanyalah instrumen dan dapat dijamin berdasarkan asas konstitusi dan arah kebijakan negara yang dimaksud. Oleh karena itu, metode dapat dipertahankan atau diubah jika dipandang sebagai jalan demokrasi yang benar dalam kondisi tertentu. Penentuan suatu pilihan harus melewati banyak penelaahan pemikiran dan pengalaman untuk memastikan baik buruknya pilihan itu sendiri [5]. Indonesia menjunjung tinggi sistem demokrasi. Praktik demokrasi adalah penyelenggaraan pemilihan umum di Indonesia untuk memilih calon wakil rakyat. Sudah ada beberapa nama calon presiden yang telah diusung oleh partai pendukung dan pendukungnya masing-masing jelang pemilu 2024, antara lain Bapak H. Anies Rasyid Baswedan, S.E., M.P.P., Ph.D., Bapak H. Ganjar Pranowo, S.H, MIP, dan Bapak Letnan Jenderal (Purn)H. Prabowo Subianto [6]. Namun, dengan berkembangnya teknologi komunikasi dan digital, cara masyarakat berpartisipasi dalam pemilihan umum telah mengalami perubahan yang signifikan [7].

Media sosial memiliki karakteristik-karakteristik yang membuat partai politik terdorong untuk menerapkan strategi kampanye online [8]. Media sosial merupakan salah satu platform utama yang memfasilitasi interaksi dan ekspresi publik. Di era digital ini, Twitter telah menjadi salah satu media sosial paling populer yang memainkan peran penting

dalam mendukung komunikasi dan berbagi informasi antara pengguna [9].

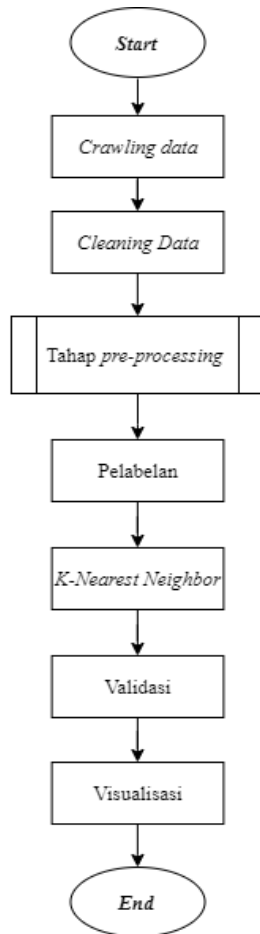
K-Nearest Neighbor merupakan salah satu metode untuk mengambil keputusan menggunakan pembelajaran terawasi dimana hasil dari data masukan yang baru diklasifikasi berdasarkan terdekat dalam data nilai. Algoritma K-Nearest Neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek yang berdasarkan dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN merupakan algoritma supervised learning dimana hasil dari query instance yang baru diklasifikasi berdasarkan mayoritas dari kategori pada algoritma KNN. Dimana kelas yang paling banyak muncul yang nantinya akan menjadi kelas hasil dari klasifikasi [10].

Pada penelitian ini, sumber data yang diperoleh melalui Twitter yang di-scraped sebagai dataset penelitian, kicauan yang dikirim masyarakat akan terkumpul saat proses scraper lalu akan tersimpan pada file excel kemudian di teliti dan diproses dengan algoritma K-Nearest Neighbor untuk mengetahui elektabilitas tokoh politik [11].

Dalam data mining terdapat beberapa tahapan, diantaranya tahapan preprocessing, tahap processing, dan tahap post-processing. Pada tahap preprocessing terdiri dari beberapa tahapan diantaranya adalah data cleaning, data integration, data reduction, dan data transformation [12]. Pada penelitian ini akan dilakukan analisis sentimen dengan menggunakan pendekatan dalam machine learning yang dikenal dengan metode K- Nearest Neighbor (K-NN) [13]. Penelitian mengenai implementasi K-Nearest Neighbor telah banyak dilakukan oleh peneliti-peneliti sebelumnya [14]. Algoritma K-Nearest Neighbor (K-NN) merupakan metode lazy learning dimana tidak ada model yang dipelajari dari data testing, jadi hanya belajar dari contoh uji harus diklasifikasikan [15].

2. METODE PENELITIAN

Tahapan-tahapan yang dijalankan meliputi pengumpulan data atau crawling data dari platform media sosial Twitter (X), dilanjutkan dengan cleaning data, tahap ketiga dilakukan tahap pre-processing data. Kemudian dilakukan pelabelan data latih yang dilanjutkan dengan pembobotan kata (weighting) menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Data yang telah diproses tersebut, akan dilakukan pelabelan dan proses klasifikasi menggunakan Algoritma KNN (K-Nearest Neighbor).



Gambar 1. Flowchart Metode Penelitian

2.1. Identifikasi Masalah

Penelitian dilakukan didasarkan oleh permasalahan yang telah dijelaskan sebelumnya yaitu:

1. Untuk mengetahui sentimen publik terhadap Pemilu Indonesia 2024. Analisis sentimen akan menghasilkan persentase seberapa besar tanggapan Masyarakat yang mendukung diadakannya pemilu dengan pelabelan positif, negatif, atau netral.
2. Mengetahui kata-kata yang sering muncul pada setiap sentimen positif, negatif, dan netral.

2.2. Crawling data

Pengumpulan data merupakan tahapan awal untuk dilakukan analisis sentimen pada sebuah topik. Dalam mengumpulkan data Twitter (X), digunakan tools dari Google Colab serta twitter API untuk memperoleh token yang dapat digunakan untuk menarik data tweet. Setelah berhasil melakukan penarikan, maka data akan disimpan ke dalam format csv.

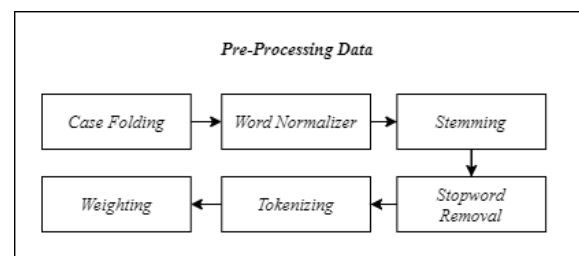
2.3. Cleaning Data

Data mentah yang telah didapat melalui *crawling* masih terdapat banyak sekali kotoran seperti data duplikat, *mention*, *hashtag*, *emoji*, alamat

website URL, hingga simbol-simbol yang tidak relevan untuk digunakan pada penelitian ini. Oleh karena itu, *cleaning data* dilakukan untuk memastikan bahwa data sudah benar-benar bersih dan dapat dibaca dengan baik untuk dapat dilakukan proses lebih lanjut.

2.4. Pre-Processing Data

Tahapan *pre-processing* dilakukan untuk mengolah data menjadi bahasa yang lebih standar, umum, dan dasar. Berikut gambaran tahapan *pre-processing data*:



Gambar 2. Flowchart pre-processing data

1. *Case Folding*: Setelah data melewati tahap sebelumnya, maka dilanjutkan dengan *case folding* yang dilakukan untuk menyeragamkan karakter pada data. Penyeragaman kata pada penelitian ini dilakukan dengan mengubah seluruh huruf menjadi huruf kecil.
2. *Word Normalizer*: Untuk memperbaiki kata-kata dalam ulasan sehingga menghasilkan kalimat yang baik dan benar sesuai dengan aturan tata bahasa Indonesia. Contoh *word normalizer* dilakukan seperti pada kata “mksh” diubah menjadi “terima kasih”.
3. *Stemming*: Proses penguraian suatu kata menjadi kata dasarnya tanpa imbuhan di awal maupun di akhir kata.
4. *Stopword Removal*: Kata-kata yang tidak memiliki makna arti yang jelas seperti imbuhan akan dihilangkan dari data yang digunakan pada penelitian.
5. *Tokenizing*: berfungsi untuk membagi teks input menjadi *array token*
6. *Weighting*: Merupakan proses pemberian bobot pada setiap kata yang ada dalam sebuah data [24]. Pembobotan kata menggunakan metode TF-IDF (*Term Frequency - Inversed Document Frequency*).

2.5. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) adalah algoritma yang digunakan untuk melakukan klasifikasi terhadap suatu objek, berdasarkan k-buah data latih yang jaraknya paling dekat dengan objek tersebut. Syarat nilai k adalah tidak boleh lebih besar dari jumlah data. Perhitungan jarak pada KNN dapat menggunakan persamaan [18]:

$$d_i = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \tag{1}$$

3. HASIL DAN PEMBAHASAN

Penelitian ini melakukan pengumpulan data *tweets* pada Bulan Oktober 2023. Adapun tanggal pengambilan data pada setiap bulannya dilakukan secara acak sehingga didapati data yang terkumpul sebanyak 250 data *tweets*. Kata kunci yang dipakai untuk penarikan data adalah “Pemilu”. Untuk dapat mengetahui bahwa yang beropini adalah orang Indonesia, maka dilakukan *filter* penarikan data *tweets* yang diatur hanya yang berbahasa Indonesia.

3.1. Data Mentah

Untuk mendapatkan data mentah, dilakukan Teknik *crawling* menggunakan tools dari Google Colab dengan bahasa pemrograman Python dan *library* *pandas* sebagai dasar perintah untuk penarikan datanya. Data awal memiliki banyak sekali tulisan yang tidak diperlukan seperti data duplikat, *@mention*, *hashtag*, *emoji*, simbol-simbol, dan URL seperti pada Tabel 1 berikut:

Nomor	Teks
1	Netralitas Polisi Dipertanyakan: Apakah pemilu masih berlangsung adil? Dimana Netralitas Aparat Netralitas Polri Dipertanyakan https://t.co/3G5a9nhz8i
2	Waspada kecurangan penguasa pada pemilu 2024 Netralitas Polri Dipertanyakan Dimana Netralitas Aparat https://t.co/2XFyjDZXAc
3	Ribuan WNA China di New York diberikan KTP utk Pemilu 2024 ?!* https://t.co/7frFwS1TEQ
4	Urgently Pemilu sudah di depan mata! 2024 Rakyat Indonesia harus sehat dan sejahtera setuju Sis dan bro sudah waktunya kita mempunyai wakil rakyat yang berkomitmen dan memperjuangkan hak-hak kita! Khususnya dalam memperoleh kesehatan dan kesejahteraan Ayo gasken! https://t.co/0gwn6ldvKr
5	Beginilah kalau Pemilu kebobolan Peserta Dunggu. Debat ngga berani. Kampanye tertutup. Diwawancara kabur. Betul-betul Pemilu terKOPLAK dalam sejarah. https://t.co/0BZt3rPQA1

3.2. Cleaning Data

Data mentah yang telah diperoleh melalui proses *crawling* masih mengandung banyak elemen yang perlu dibersihkan, seperti data duplikat, *@mention*, *hashtag*, *emoji*, URL, dan simbol-simbol yang tidak relevan untuk keperluan penelitian ini. Oleh sebab itu dilakukan *cleaning data* menggunakan tools *rapidminer*. Dari total 291 data *tweets* yang berhasil di *cleaning*, tersisa sebanyak 200 data yang dilanjutkan ke tahap berikutnya. Hasil dari data yang telah melalui proses *cleaning* tertera pada Tabel 2 berikut:

Nomor	Teks
1	Netralitas Polisi Dipertanyakan: Apakah pemilu masih berlangsung adil Dimana Netralitas Aparat Netralitas Polri Dipertanyakan
2	Waspada kecurangan penguasa pada pemilu 2024 Netralitas Polri Dipertanyakan Dimana Netralitas Aparat
3	Ribuan WNA China di New York diberikan KTP utk Pemilu 2024
4	Urgently Pemilu sudah di depan mata 2024 Rakyat Indonesia harus sehat dan sejahtera setuju Sis dan bro sudah waktunya kita mempunyai wakil rakyat yang berkomitmen dan memperjuangkan hak-hak kita! Khususnya dalam memperoleh kesehatan dan kesejahteraan Ayo gasken

3.3. Pre-Processing Data

Setelah melakukan *cleaning data* dilakukan tahap *pre-processing data*. Tahap ini dilakukan untuk mengolah data *tweets* menjadi bahasa yang lebih standar, umum, dasar, dan bebas dari kata-kata yang tidak memiliki arti yang jelas. Dilakukan beberapa tahapan dalam menjalankan *pre-processing data*, yakni *Case Folding*, *Word Normalizer*, *Stemming*, *Stopword Removal*, *Tokenizing*, dan *Weighting*. Berikut contoh beberapa kalimat *tweets* yang melalui tahapan *Case Folding* pada Tabel 3 berikut:

Nomor	Case Folding
1	netralitas polisi dipertanyakan: apakah pemilu masih berlangsung adil dimana netralitas aparat netralitas polri dipertanyakan
2	waspada kecurangan penguasa pada pemilu 2024 netralitas polri dipertanyakan dimana netralitas aparat
3	ribuan wna china di new york diberikan ktp utk pemilu 2024
4	urgently pemilu sudah di depan mata 2024 rakyat indonesia harus sehat dan sejahtera setuju sis dan bro sudah waktunya kita mempunyai wakil rakyat yang berkomitmen dan memperjuangkan hak-hak kita! khususnya dalam memperoleh kesehatan dan kesejahteraan ayo gasken
5	beginilah kalau pemilu kebobolan peserta dunggu. debat ngga berani. kampanye tertutup. diwawancara kabur. betul-betul pemilu terkoplak dalam sejarah.

3.4. Pelabelan Data

Dataset yang telah dilakukan *pre-processing* belum memiliki label klasifikasi, oleh sebab itu dilakukan pelabelan secara manual. Dataset dibagi ke dalam persentase 50% data latih dan 50% data uji. Data latih merupakan data yang dilakukan pelabelan secara manual, sedangkan data uji merupakan data yang dilakukan pengujian menggunakan Algoritma klasifikasi *k-nearest neighbor*. Terdapat 100 data *tweets* data latih dan 100 data *tweets* data uji yang

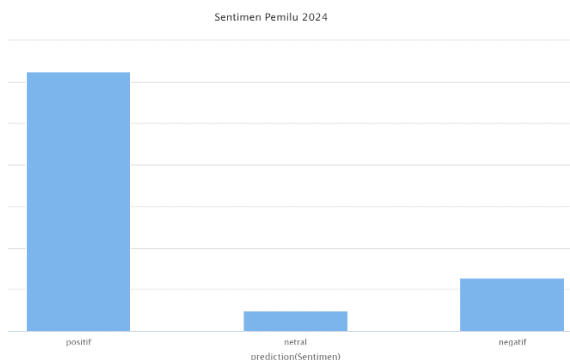
labelnya akan diprediksi oleh algoritma. Total keseluruhan dari data sebanyak 200 data.

Dalam proses pelabelan, dilakukan tiga klasifikasi apakah komentar *tweets* memiliki makna positif yang berarti setuju atau mendukung penuh pemindahan ibu kota. Netral yang berarti komentar tidak menyatakan dukungan atau ambigu akan topik pemindahan ibu kota. Negatif yang berarti menolak, menyindir, menghina, ataupun tidak setuju akan pemindahan ibu kota Hasil pelabelan tertera pada Tabel 4 berikut:

Tabel 4. Pelabelan Data

Nomor	Data Tweets	Sentimen
1	netralitas polisi dipertanyakan: apakah pemilu masih berlangsung adil dimana netralitas aparat netralitas polri dipertanyakan	Netral
2	waspada kecurangan penguasa pada pemilu 2024 netralitas polri dipertanyakan dimana netralitas aparat	netral
3	ribuan wna china di new york diberikan ktp utk pemilu 2024	negatif
4	urgently pemilu sudah di depan mata 2024 rakyat indonesia harus sehat dan sejahtera setuju sis dan bro sudah waktunya kita mempunyai wakil rakyat yang berkomitmen dan memperjuangkan hak-hak kita! khususnya dalam memperoleh kesehatan dan kesejahteraan ayo gasken	Positif
5	beginilah kalau pemilu kebobolan peserta dungu. debat ngga berani. kampanye tertutup. diwawancara kabur. betul-betul pemilu terkoplak dalam sejarah.	negatif

Persentase hasil pelabelan sentimen digambarkan pada grafik gambar 3 berikut:



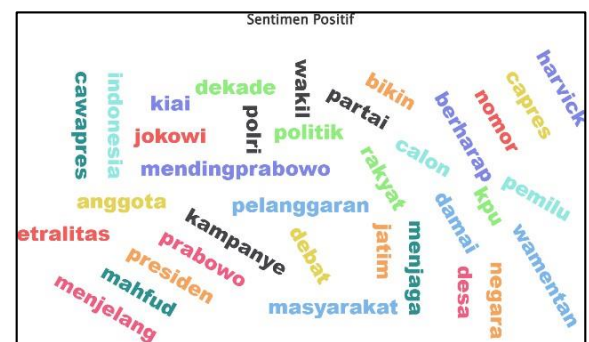
Gambar 1. Grafik Persentase Sentimen

Dari proses pelabelan data *tweets* yang telah dilakukan, sebanyak 156 *tweets* dikategorikan sebagai sentimen positif sebesar 78%, 32 *tweets* dikategorikan sebagai sentimen netral sebesar 16%, dan 12 *tweets* dikategorikan sebagai sentimen negatif sebesar 6% dari total keseluruhan data penelitian.

3.5. Visualisasi Data

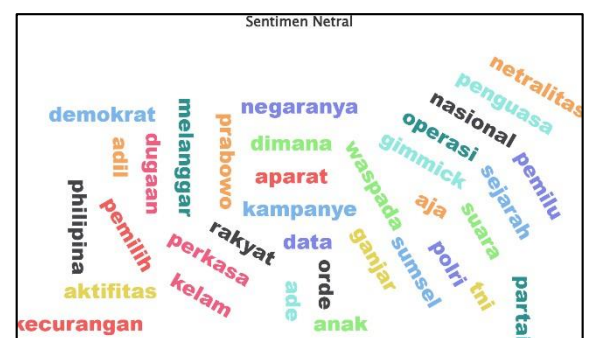
Pada setiap sentimen positif, negatif, dan netral, dilakukan visualisasi kata-kata kunci melalui media *word cloud* untuk mengetahui kata-kata apa saja yang sering muncul pada setiap sentimen tersebut sehingga didapat kesimpulan kata kunci yang sering dibahas oleh Masyarakat terkait topik pemindahan ibu kota negara. Hasil visualisasi ini dibagi menjadi tiga bagian yaitu visualisasi data positif, visualisasi data negatif, dan visualisasi data netral. Setiap visualisasi menggambarkan 35 kata yang paling sering muncul dalam masing-masing sentiment pada sentiment positif, negatif, dan netral.

Berdasarkan visualisasi sentimen positif pada gambar 5, frekuensi kata-kata yang sering muncul pada sentimen positif diantaranya adalah “damai” sebanyak 13 kali, “indonesia” sebanyak 10 kali, dan “netralitas” sebanyak 8 kali. Penggunaan kata “damai”, “indonesia”, “netralitas” merujuk pada harapan masyarakat agar proses pemilu 2024 tetap berjalan dengan lancar.



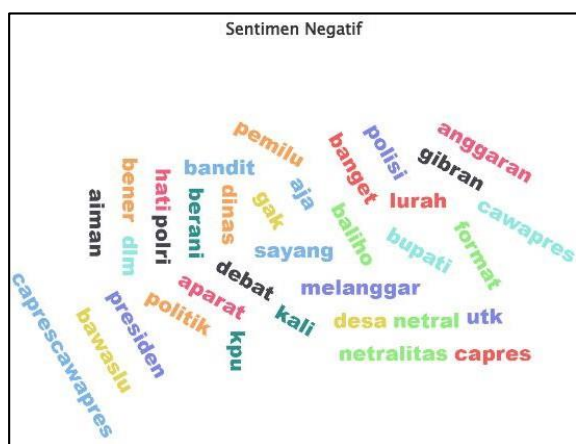
Gambar 4. Word Cloud Sentimen Positif

Pada visualisasi sentimen netral, frekuensi kata-kata yang sering muncul pada diantaranya “suara” sebanyak 5 kali, “partai” sebanyak 6 kali, dan “data” sebanyak 2 kali. Penggunaan kata-kata tersebut menunjukkan ketidakberpihakan atau skeptis akan pemilu 2024.



Gambar 5. Word Cloud Sentimen Netral

Berdasarkan visualisasi sentimen negatif pada gambar 7, frekuensi kata-kata yang sering muncul pada diantaranya “gibran” sebanyak 8 kali, “melanggar” sebanyak 3 kali. Penggunaan kata-kata tersebut menunjukkan ketidakpercayaan masyarakat akan pemilu berlangsung dengan baik dan jujur.



Gambar 6. Word Cloud Sentimen Negatif

4. DISKUSI

Penelitian terdahulu menghasilkan banyak sumber untuk penelitian ini. Hasil penelitian dahulu yang bertemakan analisis sentimen terhadap pemindahan ibu kota negara dengan membandingkan tiga algoritma: Naive Bayes, KNN, dan Random Forest menunjukkan hasil bahwa nilai positif sebesar 36.8%, netral sebesar 25%, dan negatif sebesar 38.1% [16]. Dari penelitian sebelumnya yang membahas terkait Klasifikasi Sentimen Masyarakat di Media Sosial Twitter terhadap Calon Presiden 2024 Prabowo Subianto dengan Metode K-NN yang bertujuan Penelitian ini memiliki tujuan untuk mengklasifikasikan sentimen masyarakat terhadap Calon Presiden (capres) Prabowo Subianto di Twitter. Jumlah data yang digunakan adalah 2100 tweet yang dikumpulkan berdasarkan kata kunci yaitu “Calon Presiden” dan “Prabowo Subianto”. Penerapan metode K-Nearest Neighbor (K-NN) dengan pembobotan berupa TF-IDF dan seleksi fitur berupa Threshold dilakukan implementasi dengan menggunakan Google Colab. Maka dari penelitian sebelumnya penelitian ini bertujuan untuk lebih general netral terkait pemilu 2024 sentimen terkait calon dan penyelenggaraan dari pemilu tersebut.

5. KESIMPULAN

Penelitian ini pada awalnya mengumpulkan sebanyak 291 tweets sebagai dataset awal. Setelah dilakukan tahap *cleaning*, didapati bahwa data yang bersih dan siap untuk dijalankan analisis sentimen menjadi sebanyak 200 jumlah data. Hasil analisis sentimen yang diperoleh terhadap topik pemilu 2024 dengan metode *k-nearest neighbor* dengan sentimen positif masyarakat sebesar 78% yang berjumlah 156 data, sentimen netral sebesar 16% yang berjumlah 32 data, dan sentiment negatif dengan nilai 6% yang berjumlah 12 data. Adapun frekuensi kata-kata kunci yang sering muncul dalam sentimen positif adalah “indonesia”, “netralitas”, dan “damai”, pada sentimen netral terdapat “suara” dan “data”, sedangkan kata

“gibran” dan “melanggar” merupakan kata yang sering muncul pada sentimen negatif. Hal ini mengindikasikan bahwa kata-kata tersebut merupakan ekspresi yang umum digunakan oleh pengguna Twitter dalam beropini dalam konteks sentimen yang berbeda.

DAFTAR PUSTAKA

- [1] P. D. Tommy, U. Ema, P. M. Kurniawan. “Analisis Sentimen Pemilu 2024 dengan Naive Bayes Berbasis Particle Swarm Optimization (PSO)”. *EXPLORE – Volume 13 No 1 Tahun 2023* p-ISSN : 2087- 894X e-ISSN : 2656-615X.
- [2] F. P. Midhat, “Perbandingan Algoritma K-Nearest Neighbor (KNN) dan Naive Bayes Classifier (NBC) dengan pelabelan Transformers serta Ekstraksi Fitur TF-IDF dan N-Gram untuk Analisis Sentimen Terhadap Penundaan Pemilu 2024”.
- [3] N. P. Aprilia, H. L. April, H. S. Shofa, Tukino, P. Bayu. “Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter”. *JOINTECS*, Vol. 8 No. 1, pp. 11 - 18, 2023, e-ISSN:2541-6448 p- ISSN:2541-3619, <http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>. <https://kab-minahasaselat.kpu.go.id/page/read/34/pemilu-tahun-1955#:~:text=Pemilu%20ini%20diikuti%20oleh%20lebih,daftar%20kumpulan%20dan%20calon%20perorangan>
- [4] D. A. Dwi, K. L. Kristoforus. “EFEKTIVITAS PEMILIHAN UMUM SERENTAK TAHUN 2024”. *JHP 17 (Jurnal Hasil Penelitian)*, Vol. 8, No. 2, Juli 2023, E-ISSN: 2502-8308 P-ISSN: 2579-7980, <http://jurnal.untag-sby.ac.id/index.php/jhp17>
- [5] A. Muhammad, S. Mohammad. “ANALISIS SENTIMEN ANIES BASWEDAN MENJADI CALON PRESIDEN 2024 MENGGUNAKAN KSTRASI FITUR COUNTVECTORIZER DAN ALGORITMA KNN”, *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, Vol. 2, No. 2, September 2023, ISSN 2962-8628 (online)
- [6] Z. H. Haekal, A. M. Rolly, R. I. Woro. “Perbandingan Model Svm, Knn Dan Naive Bayes Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024”. *JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, Vol. 8, No. 3, pp. 2083-2093, 2023. P-ISSN: 2964-7231, E-ISSN 2614-3658 DOI: <https://doi.org/10.24815/jimps.v8i3.25342>

- [7] A. A. Firdaus, A. Yudhana, I. Riadi, "ANALISIS SENTIMEN PADA PROYEKSI PEMILIHAN PRESIDEN 2024 MENGGUNAKAN METODE SUPPORT VECTOR MACHINE", *DECODE: Jurnal Pendidikan Teknologi Informasi*, Vol. 3, No. 2, pp. 236-245, September 2023, ISSN Cetak: 2775-2984 ISSN Online: 2775-1813, Doi: <http://dx.doi.org/10.51454/decode.v3i2.172> <http://journal.umkendari.ac.id/index.php/decode>.
- [8] M. Oktaviani, P. Cahyo, H. H. N. Harani. "Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based dan Random Forest". *JURNAL ILMIAH INFORMATIK*, Vol. 11 No.02, 2023, ISSN (Print) 2337-8379 | ISSN (Online) 2615-1049.
- [9] T. J. Asahar, Y. Delfi, A. Kurnia. "IMPLEMENTASI METODE K-NEAREST NEIGHBOR (KNN) DAN SIMPLE ADDITIVE WEIGHTING (SAW) DALAM PENGAMBILAN KEPUTUSAN SELEKSI PENERIMAAN ANGGOTA PASKIBRAKA", *Jurnal Pseudocode*, Vol. 3, No. 2, September 2016, ISSN 2355-5920.
- [10] P. A. Rizki, P. Wahyu, "PENERAPAN ALGORITMA K- NEAREST NEIGHBOR (KNN) ANALISIS SENTIMEN TWEET PADA ELEKTABILITAS BAKAL CALON PRESIDEN 2024". *3rdSeminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, Vol. 2, No. 2, September 2023, ISSN 2962-8628 (online)
- [11] I. A. Muhammad, "PENINGKATAN AKURASI ALGORITMA K-NEAREST NEIGHBOR MENGGUNAKAN NORMALISASI Z-SCORE DAN PARTICLE SWARM OPTIMIZATION UNTUK PREDIKSI CUSTOMER CHURN".
- [12] S. R. Aris. "Analisis Sentimen Pada Media Sosial Menggunakan Metode K- Nearest Neighbor (K-NN)", *Teknologipintar.org*, Vol. 2, No. 8, 2022.
- [13] Z. H. Farhan, "IMPLEMENTASI ALGORITMA KNN DALAM MENGUKUR KETEPATAN KELULUSAN MAHASISWA UIN SYARIF HIDAYATULLAH JAKARTA".
- [14] S. Retno, "Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi menggunakan Algoritma K-Nearest Neighbor (K-NN)". *Evolusi: Jurnal Sains dan Manajemen*, Vol 8 No. 1, Maret 2020, ISSN:2338-8161 E-ISSN: 2657-0793.
- [15] J. Muliawan and E. Dazki, "SENTIMENT ANALYSIS OF INDONESIA'S CAPITAL CITY RELOCATION USING THREE ALGORITHMS: NAÏVE BAYES, KNN, AND RANDOM FOREST," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 5, pp. 1227–1236, 2023, doi: 10.52436/1.jutif.2023.4.5.347.