

Analisis Sentimen *Crawling* Data dari Sosial Media X tentang Gaza Menggunakan Metode SVM dan *Decision Tree*

Nicola Van Robert Jhosefhin ¹, Christine Dewi ^{2*}

^{1,2*} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Kota Salatiga, Provinsi Jawa Tengah, Indonesia.

Email: 672020297@student.uksw.edu ¹, christine.dewi@uksw.edu ^{2*}

Histori Artikel:

Dikirim 1 November 2024; *Diterima dalam bentuk revisi* 15 November 2024; *Diterima* 30 Desember 2024; *Diterbitkan* 10 Januari 2025. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Konflik Palestina dengan Israel telah menyita banyak perhatian mata dunia terkhusus dijalar Gaza setelah kelompok Hamas mengambil ahli kepemimpinan, sehingga masyarakat sering membuat tweet (postingan) dengan berbagai pandangan, ada yang memiliki pandangan netral, positif, negatif. Analisis sentiment digunakan untuk menganalisa data yang didapatkan dari postingan yang ada di twitter. Tujuan dari penelitian ini untuk mengetahui pandangan masyarakat terhadap jalur Gaza dengan menerapkan algoritma Decision Tree dan Support Vector Machine untuk mengukur Tingkat akurasi. Hasil dari analisis sentimen menunjukkan data sebanyak 805 netral, 415 positif, dan 308 negatif lalu algoritma Support Vector Machine menghasilkan Tingkat akurasi yang tertinggi dengan nilai 73.86 yang sudah dapat dikategorikan baik.

Kata Kunci: Decision Tree; Support Vector Machine; Analisis Sentimen; Gaza; Hamas; Palestina; Israel; Twitter.

Abstract

The Palestinian conflict with Israel has attracted a lot of attention from the world, especially in the Gaza Strip after the Hamas group took over the leadership, so that people often make tweets (posts) with various views, some have neutral, positive, negative views. Sentiment analysis is used to analyze data obtained from posts on Twitter. The purpose of this study is to determine the public's views on the Gaza Strip by applying the Decision Tree and Support Vector Machine algorithms to measure their level of accuracy. The results of the sentiment analysis show 805 neutral data, 415 positive, and 308 negative, then the Support Vector Machine algorithm produces the highest level of accuracy with a value of 73.86 which can be categorized as good.

Keyword: Decision Tree; Support Vector Machine; Sentiment Analyst; Gaza; Hamas; Palestine; Israel; Twitter.

1. Pendahuluan

Konflik Palestina dengan Israel telah menyita banyak perhatian mata dunia, terutama karena begitu banyak korban yang bergelimpangan jatuh dalam sebuah konflik tersebut. Berdasarkan informasi yang ada tentang awal penyerangan tentara Israel, di mana penyerangan tentara Israel terhadap Jalur Gaza menjadi sorotan penting (Harian, 2018). Gaza sendiri merupakan wilayah yang terletak di pantai timur laut Mediterania yang berbatasan dengan Israel di timur dan utara, serta Mesir di barat daya. Wilayah ini memiliki sejarah panjang yang penuh dengan konflik dan ketegangan, terutama terkait dengan perjuangan Palestina untuk kemerdekaan dan hak-hak mereka (Hadi, 2021; BBC News Indonesia, 2023). Secara de facto, Hamas menjadi penguasa di Jalur Gaza sejak 2007, di mana Jalur Gaza ini terus diperebutkan dan menjadi medan tempur dalam konflik antara Israel dan Palestina (Adryamarthanino, 2021). Hamas sendiri merupakan gerakan nasionalis dan Islamis Palestina di mana tujuannya untuk mendirikan negara Islam independen di Palestina. Gerakan ini juga dikenal karena menggunakan kekerasan, termasuk tindakan teroris, untuk mencapai tujuan. Walau demikian, Hamas juga terlibat dengan kegiatan sosial dan amal pada wilayah yang didudukinya (Marpaung *et al.*, 2024). Setelah Hamas mengambil alih di Jalur Gaza, kelompok ini menjadi kekuatan yang dominan dan menyebabkan ketegangan dengan Israel (Memantik *et al.*, 2024). Apa yang kita ketahui sekarang terkait kejadian pada Gaza tidak terlepas dari perkembangan zaman yang semakin modern sehingga memudahkan kita untuk mengetahui informasi dengan mudah, di mana perkembangan teknologi informasi dan komunikasi telah membuka akses yang lebih luas bagi masyarakat untuk menyuarakan opini dan pandangannya melalui berbagai platform media sosial, salah satunya adalah *Twitter* (Studi *et al.*, 2024).

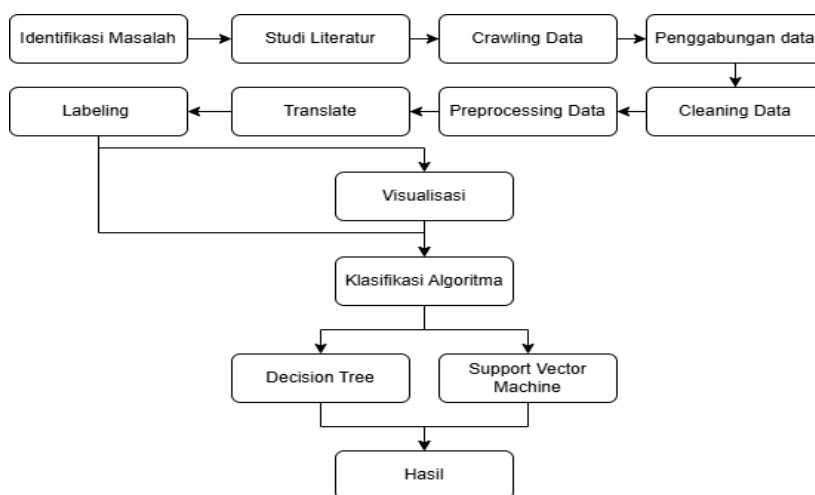
Twitter atau yang sekarang sudah dikenal sebagai *X* merupakan salah satu platform media sosial yang memungkinkan pengguna untuk mengirim dan membaca pesan singkat dan merupakan salah satu jejaring sosial yang paling signifikan. *Twitter* digunakan oleh berbagai kalangan, termasuk individu, perusahaan, organisasi, selebriti, dan politisi. *Twitter* juga mengumumkan bahwa mereka memiliki lebih dari 140 juta pengguna aktif yang membuat 340 juta *tweet* tiap hari sehingga tidak heran jika *Twitter* sering digunakan untuk berbagai berita terkini, berkomunikasi, mempromosikan produk atau layanan. Kemudian, *Twitter* ini juga dikenal karena kecepatannya dalam menyebarkan informasi sehingga membuatnya menjadi alat yang penting dalam situasi darurat dan saat terjadi peristiwa penting (Dewi & Chen, 2019). Salah satu contohnya ialah kejadian di Gaza. Selain itu, *Twitter* ini juga banyak digunakan untuk menyampaikan pandangan politik, sosial, ekonomi, dan masih banyak lagi (Mailoa, 2024). Konflik yang terjadi di Gaza sering menjadi perbincangan hangat di *Twitter*. Dengan adanya platform ini, dapat digunakan untuk mengetahui pandangan atau pendapat masyarakat terkait kejadian yang ada di Gaza dengan memanfaatkan analisis sentimen. Analisis sentimen merupakan sebuah metode yang digunakan untuk mengekstrak dan menganalisis data secara kuantitatif maupun kualitatif untuk memahami opini, perasaan, dan emosi yang terkandung di dalamnya. Analisis sentimen dapat dilakukan dengan mengidentifikasi kata dan frasa positif, negatif, atau netral (Sejati *et al.*, 2023). Berdasarkan latar belakang yang ada, maka akan dilakukan penelitian dengan melakukan analisis sentimen yang berkaitan dengan kejadian yang berada di Jalur Gaza, di mana pada penelitian ini juga akan menggunakan metode algoritma *Machine Learning*, yaitu *Support Vector Machine (SVM)* dan *Decision Tree* dengan menggunakan *dataset* dari *Twitter* tentang kejadian Gaza. Penelitian ini diharapkan dapat memberikan informasi mengenai pandangan masyarakat terhadap kejadian yang berada pada Jalur Gaza. Berbagai penelitian sebelumnya telah mengkaji penerapan algoritma *Decision Tree* dan *Support Vector Machine (SVM)* dalam analisis sentimen, menunjukkan kinerja yang efektif dalam berbagai situasi. Singgalen (2023) menggunakan pendekatan *CRISP-DM* untuk menganalisis sentimen wisatawan terhadap layanan akomodasi hotel di Makassar berdasarkan ulasan di *Tripadvisor*. Tahapan yang dilalui meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi. Algoritma *SVM* menunjukkan performa tinggi dengan tingkat akurasi 98,98%, presisi 100%, *recall* 97,96%, *f-measure* 98,97%, dan *AUC* 100%, dengan dominasi sentimen positif. Tuhuteru (2020) menggunakan algoritma *SVM* untuk mengevaluasi pandangan masyarakat

terhadap kebijakan PSBB di Kota Ambon. *Dataset Twitter* yang terdiri dari 1.075 *tweet* diolah dengan pembagian data latih sebanyak 350 dan data uji sebanyak 725. Hasil analisis menunjukkan distribusi sentimen sebesar 28% positif, 27% negatif, dan 45% netral, mencerminkan pandangan masyarakat yang cukup berimbang terhadap kebijakan tersebut. Zy *et al.* (2023) membandingkan kinerja algoritma *Naïve Bayes*, *SVM*, dan *Decision Tree* dalam mendeteksi ancaman keamanan jaringan seperti malware. *Decision Tree* menunjukkan presisi tertinggi sebesar 99%, diikuti oleh *SVM* sebesar 98%, dan *Naïve Bayes* sebesar 86%.

Penelitian ini memperlihatkan kemampuan masing-masing algoritma dalam menangani data yang kompleks dan memerlukan tingkat presisi tinggi. Oktavia *et al.* (2023) menggunakan algoritma *SVM* untuk menganalisis sentimen terhadap penerapan sistem *e-Tilang* di *Twitter*. Algoritma *SVM* menghasilkan presisi sebesar 83,33% dan *recall* sebesar 5,28%, dengan dominasi sentimen netral. Yasir dan Suraji (2023) membandingkan algoritma *Naïve Bayes*, *Decision Tree*, dan *Random Forest* dalam analisis sentimen terhadap kenaikan biaya haji. *Naïve Bayes* mencapai akurasi tertinggi sebesar 90%, diikuti oleh *Random Forest* sebesar 87%, dan *Decision Tree* sebesar 83%. Ramadhan dan Wahyudin (2022) menganalisis sentimen terhadap keberhasilan Indonesia di ajang *Thomas Cup 2020* menggunakan algoritma *Naïve Bayes* dan *Decision Tree*. Algoritma *Naïve Bayes* memiliki akurasi 95,16%, sedangkan *Decision Tree* mencapai 84,68%. Penelitian ini mengonfirmasi efektivitas algoritma *Naïve Bayes* dalam pengolahan data sentimen. Hasil dari berbagai penelitian tersebut menunjukkan efektivitas algoritma *SVM* dan *Decision Tree* dalam analisis sentimen pada berbagai isu, termasuk sektor pariwisata, kebijakan sosial, dan isu publik lainnya. Berdasarkan hasil tersebut, kedua algoritma tersebut memiliki dasar metodologis yang kuat untuk digunakan dalam analisis sentimen publik terkait situasi di Gaza menggunakan data yang diambil dari media sosial *Twitter*. Hal ini mendukung keandalan metode dalam menganalisis opini masyarakat secara luas.

2. Metode Penelitian

Tahapan penelitian yang dilakukan, ditunjukkan seperti pada gambar 1.



Gambar 1. Tahapan penelitian

Alur penelitian pada gambar 1 dilakukan secara bertahap untuk menghasilkan keluaran yang dapat dipahami dalam bentuk akurasi. Tahap pertama adalah Identifikasi Masalah, di mana dilakukan identifikasi terhadap permasalahan terkait sentimen masyarakat terhadap kejadian di Jalur Gaza. Langkah ini mencakup definisi masalah, kendala, tujuan penelitian, serta analisis manfaat. Tahap berikutnya adalah Studi Literatur, yang bertujuan mencari teori-teori dan metode relevan dari

penelitian sebelumnya untuk menghasilkan solusi yang sesuai dengan permasalahan. Setelah itu, dilakukan *Crawling Data*, yaitu pengumpulan data secara online menggunakan aplikasi tambahan atau bahasa pemrograman *Python*. Peneliti menggunakan *Python* untuk melakukan *crawling* data dari media sosial *Twitter* sesuai dengan kebutuhan penelitian. Data yang telah diperoleh kemudian melewati tahap Penggabungan Data, di mana data dari 8 file CSV digabungkan menjadi satu untuk melanjutkan ke proses selanjutnya. Pada tahap *Cleaning Data*, data yang telah digabungkan dibersihkan dari duplikasi, tanda baca, dan elemen lain yang tidak diperlukan untuk memastikan data siap digunakan. Tahap ini diikuti oleh *Preprocessing*, yaitu proses mengubah data mentah menjadi lebih terstruktur melalui langkah normalisasi, *stopword removal*, *tokenization*, dan *stemming*. Normalisasi memperbaiki struktur data agar terorganisasi, sedangkan *stopword removal* menghapus kata-kata umum yang tidak bermakna. Selanjutnya, *tokenization* memecah teks menjadi unit-unit kecil, seperti kata, dan *stemming* mengubah kata-kata menjadi bentuk dasarnya. Setelah *preprocessing* selesai, data diterjemahkan dari bahasa Indonesia ke bahasa Inggris menggunakan *Google Spreadsheet* pada tahap *Translate*. Tahap berikutnya adalah *Labeling*, di mana data dianalisis sentimennya menggunakan *Natural Language Toolkit (NLTK)*, sebuah pustaka dalam *Python*. Sentimen diklasifikasikan menjadi kategori positif, netral, dan negatif. Pada tahap Visualisasi, dilakukan visualisasi data untuk menampilkan kata-kata yang paling sering muncul dan hasil analisis sentimen dalam bentuk grafik. Visualisasi ini membantu memahami distribusi sentimen dan kata-kata utama yang mendominasi dataset. Selanjutnya, analisis dilakukan dengan dua algoritma, yaitu *Decision Tree* dan *Support Vector Machine (SVM)*. Pada *Decision Tree*, data dibagi dengan rasio 80:20 (80% untuk data latih dan 20% untuk data uji). Algoritma ini mengklasifikasikan data berdasarkan fitur dalam dataset menggunakan vektor TF-IDF. Rumus untuk menghitung *entropy* dan *gain* digunakan untuk mengukur ketidakpastian dan perbedaan setelah pembagian data. Pada algoritma SVM, data juga dibagi dengan rasio 80:20. Algoritma ini digunakan untuk klasifikasi dan regresi pada data linear maupun non-linear. Rumus *kernel similarity* digunakan untuk mengukur kesamaan dalam ruang fitur dan mengontrol margin *hyperplane*.

$$Entropy (S) = \sum_{i=0}^n -p_i * \log^2 p_i$$

Keterangan:

S = Himpunan kasus

A = Fitur

n = Jumlah Partisi S

Pi = proporsi dari Si terhadap S

Pada rumus (1) yang bisa dilihat diatas dapat digunakan untuk Mengukur ketidakpastian atau keacakan dalam suatu sistem.

$$Gian (S, A) = Entropy (S) = \sum_i^n \frac{|s_i|}{|S|} * Entropy (S_i)$$

Keterangan:

S = Himpunan

A = Atribut

n = Jumlah Partisi atribut A

|Si| = Jumlah kasus pada partisi ke-i

|S| = Jumlah kasus dalam S

Pada rumus (2) diatas digunakan untuk Mengukur perbedaan antara Entropy sebelum dan sesudah dilakukan pemisahan data pada suatu node.

$$D_{ij} = Y_i Y_j (\vec{X}_i \cdot \vec{X}_j) + \lambda^2$$

Keterangan

D_{ij} = Nilai Kesamaan kernel antara sampel ke-i dan ke-j

Y_i = label kelas untuk sampel data ke-i

Y_j = label kelas untuk sampel data ke-j

X_i = Vektor fitur untuk sampel data ke-i

X_j = Vektor fitur untuk sampel data ke-j

λ = Lamda.

Pada rumus (3) diatas digunakan untuk Mengukur kesamaan dalam ruang fitur dan parameter regulasi untuk mengontrol *margin hyperplane*.

3. Hasil dan Pembahasan

3.1 Hasil

Penelitian memanfaatkan dua algoritma, yaitu *Support Vector Machine (SVM)* dan *Decision Tree*, yang masing-masing memiliki kelebihan serta kekurangan. Algoritma SVM unggul dalam menangani data berdimensi tinggi dan efisien dalam penggunaan memori. Namun, SVM membutuhkan waktu pelatihan yang lama serta memerlukan komputasi yang intensif. Sebaliknya, *Decision Tree* memiliki keunggulan pada kemudahan interpretasi melalui visualisasi, tetapi kinerjanya sangat bergantung pada kualitas serta keteraturan data pelatihan. Hal ini menjadikannya kurang efektif ketika data pelatihan tidak optimal (Dewi & Chen, 2020; Irawan, 2021). Pengumpulan data dilakukan melalui *Twitter* menggunakan *auth_token*, yang memungkinkan program untuk terhubung dengan platform tersebut. Proses pengumpulan data menggunakan kata kunci seperti "Hamis", "Gaza", dan "Kejadian Gaza". Sebanyak 1.924 *tweet* berhasil dikumpulkan dan selanjutnya diproses sesuai tahapan penelitian. Tabel 1 menampilkan contoh data mentah yang belum melalui proses pembersihan dan pengolahan.

Tabel 1. Sampel Data Hasil *Crawling*

| Full_Text |
|---|
| SEBARKANâ€™â€Œ Katanya Hamas pembela warga sipil Gaza. Faktanya Hamas menyiksa warga Gaza. Video Fakta Israel membuka KEBOHONGAN Hamas yg di di rilis oleh IDF. Silahkan di simak https://t.co/CxTkg0jvmO |
| Aslinya warga palestina benci sekali sama Hamas sayangnya mereka dbawah tekanan hamas sehingga tidak berani melawan. Sekarang mereka bebas mengutuk dan memaki hamas. https://t.co/jHU8aK0kuf |
| @erlanishere KENYATAANNYA GAZA MAKIN HANCYUURRR gara2 HAMAs. ISRAEL bukan MELAWAN PAKESTINA tetapi MELAWAN TERORIST HAMAs. Palestina boleh Punya Otoritas SENDIRI tapi INGAT JANGAN LAGI MINTA2 kek PEMULUNG PENGEMIS dr Iang pajak kami yah..! |

Dari data yang terdapat pada tabel 1 ini akan dilakukan *cleaning data* untuk membersihkan *username*, tandabaca, yang bisa dilihat pada tabel 2 dibawah ini.

Tabel 2. Sampel Data hasil *Cleaning*

| Full_Text |
|--|
| sebarkan katanya hamas pembela warga sipil gaza faktanya hamas menyiksa warga gaza video fakta israel membuka kebohongan hamas yg di di rilis oleh idf silahkan di simak |

aslinya warga palestina benci sekali sama hamas sayangnya mereka dbawah tekanan hamas sehingga tidak berani melawan sekarang mereka bebas mengutuk dan memaki hamas
kenyataannya gaza makin hancuuurrr gara2 hamas israel bukan melawan pakestina tetapi melawan terorist hamas palestina boleh punya otoritas sendiri tapi ingat jangan lagi minta2 kek pemulung pengemis dr iang pajak kami yah

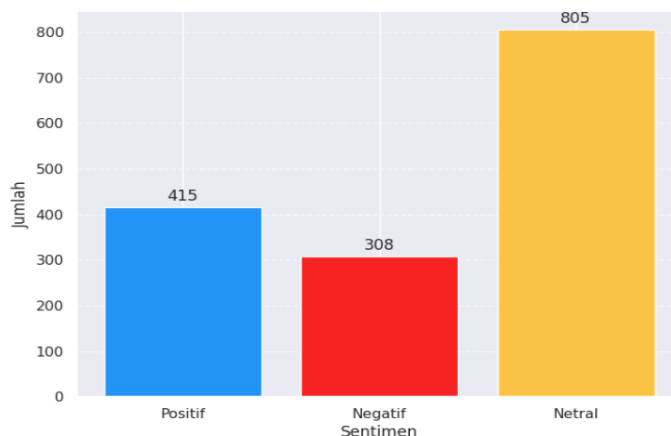
Tabel 2 merupakan sampel data yang sudah dilakukan pembersihan, dan dapat dilihat perbedaannya antara dengan tabel 1, berikutnya dilakukan *Preprocessing*, yang dapat dilihat pada tabel 3.

Tabel 3. Sampel Data Hasil *Preprocessing*

| Normalisasi |
|---|
| sebar kan katanya hamas pembela warga sipil gaza faktanya hamas menyiksa warga gaza video fakta israel membuka kebohongan hamas yang di di rilis oleh idf silahkan di simak |
| aslinya warga palestina benci sekali sama hamas sayangnya mereka dbawah tekanan hamas sehingga tidak berani melawan sekarang mereka bebas mengutuk dan memaki hamas |
| kenyataannya gaza makin hancuuurrr gara2 hamas israel bukan melawan pakestina tetapi melawan terorist hamas palestina boleh punya otoritas sendiri tapi ingat jangan lagi minta2 kek pemulung pengemis dr iang pajak kami yah |
| Stopwords |
| sebar kan katanya hamas pembela warga sipil gaza faktanya hamas menyiksa warga gaza video fakta israel membuka kebohongan hamas rilis idf silahkan di simak |
| aslinya warga palestina benci sekali sama hamas sayangnya dbawah tekanan hamas tidak berani melawan sekarang bebas mengutuk memaki hamas |
| kenyataannya gaza makin hancur gara2 hamas israel bukan melawan pakestina melawan terorist hamas palestina punya otoritas sendiri ingat jangan minta2 kek pemulung pengemis dr iang pajak yah |
| Tokenize |
| ['sebar kan', 'katanya', 'hamas', 'pembela', 'warga', 'sipil', 'gaza', 'faktanya', 'hamas', 'menyiksa', 'warga', 'gaza', 'video', 'fakta', 'israel', 'membuka', 'kebohongan', 'hamas', 'rilis', 'idf', 'silahkan', 'di', 'simak'] |
| ['aslinya', 'warga', 'palestina', 'benci', 'sekali', 'sama', 'hamas', 'sayangnya', 'dbawah', 'tekanan', 'hamas', 'tidak', 'berani', 'melawan', 'sekarang', 'bebas', 'mengutuk', 'memaki', 'hamas'] |
| ['kenyataannya', 'gaza', 'makin', 'hancur', 'gara2', 'hamas', 'israel', 'bukan', 'melawan', 'pakestina', 'melawan', 'terorist', 'hamas', 'palestina', 'punya', 'otoritas', 'sendiri', 'ingat', 'jangan', 'minta2', 'kek', 'pemulung', 'pengemis', 'dr', 'iang', 'pajak', 'yah'] |
| Stemming |
| sebar kata hamas bela warga sipil gaza fakta hamas siksa warga gaza video fakta israel buka bohong hamas rilis idf silah di simak |
| asli warga palestina benci sekali sama hamas sayang dbawah tekan hamas tidak berani lawan sekarang bebas kutuk maki hamas |
| nyata gaza makin hancur gara2 hamas israel bukan lawan pakestina lawan terorist hamas palestina punya otoritas sendiri ingat jangan minta2 kek pulung emis dr iang pajak yah |

Tabel 3 merupakan sampel data yang sudah melewati tahap preprocessing, mulai dari tahap normalisasi, stopword, tokenize, stemming. Pada tahap normalisasi, kata kata yang ada didalam data akan diperbaiki, contohnya “wqrga” akan dinormalisasikan menjadi “warga”. Kemudian pada tahap *stopword* ada beberapa kata yang akan dihapus didalam data contohnya kata “di” jadi dapat diartikan ditahap ini akan membuang kata yang tidak memiliki makna yang berrarti. Pada tahap tokenize ini akan mengubah bentuk kalimat menjadi tiap kata. Pada tahap stemming ini akan akan merubah atau menghapus kata yang memiliki himbuan. Dari proses *cleaning* higgs pada tahap preprocessing data

Gambar 2 menunjukkan bahwa kata yang paling sering muncul didalam data yang ada ialah kata Hamas, kemudian disusul kata Israel dan seterusnya, jadi jika makin besar kata pada gambar 2 maka makin sering muncul.



Gambar 3. Visualisasi hasil sentiment keseluruhan data

Gambar 3 menunjukkan visualisasi hasil sentiment keseluruhan data, yang mana dapat dilihat dari total 1528 data, di dapatkan bahwa 415 data berlabel positif, 308 berlabel negative, dan 805 data berlabel netral. Setelah melalui tahapan-tahapan yang ada sehingga menghasilkan sentiment NLTK, berikutnya akan dibagi menjadi dua dataset yaitu data *train* dan data *test* dengan perbandingan 80% untuk data *train* lalu 20% untuk data *test*, dimana jumlah data yang dilatih 1222 lalu data yang di uji 306. Kemudian dilakukan pelatihan model algoritma *Support Vector Machine* dan *Decision Tree* kedua model yang sudah dilatih menggunakan data tai akan dilakukan prediksi untuk menentukan hasil *accuracy*, *precision*, *recall*, *f1-score*.

| | precision | recall | f1-score | support |
|-------------------|-----------|--------|----------|---------|
| Negatif | 0.56 | 0.39 | 0.46 | 75 |
| Netral | 0.73 | 0.86 | 0.79 | 152 |
| Positif | 0.51 | 0.48 | 0.50 | 79 |
| accuracy | | | 0.65 | 306 |
| macro avg | 0.60 | 0.58 | 0.58 | 306 |
| weighted avg | 0.63 | 0.65 | 0.63 | 306 |
| Accuracy: 64.71% | | | | |
| Precision: 63.08% | | | | |
| Recall: 64.71% | | | | |
| F1-score: 63.22% | | | | |

Gambar 4. Hasil *Decision Tree*

Gambar 4 merupakan hasil dari klasifikasi sentimen dengan menggunakan algoritma *Decision Tree* dengan perbandingan data 80:20 yang dimana menunjukkan bahwa hasil *accuracy* 64.71% *precision* 63.71% *Recall* 64.71% *F1-score* 63.22%.

| | precision | recall | f1-score | support |
|-------------------|-----------|--------|----------|---------|
| Negatif | 0.75 | 0.52 | 0.61 | 75 |
| Netral | 0.79 | 0.88 | 0.83 | 152 |
| Positif | 0.63 | 0.68 | 0.65 | 79 |
| accuracy | | | 0.74 | 306 |
| macro avg | 0.72 | 0.69 | 0.70 | 306 |
| weighted avg | 0.74 | 0.74 | 0.73 | 306 |
| Accuracy: 73.86% | | | | |
| Precision: 73.92% | | | | |
| Recall: 73.86% | | | | |
| F1-score: 73.24% | | | | |

Gambar 5. Hasil *Support Vector Machine*

Gambar 5 merupakan hasil klasifikasi sentimen dengan menggunakan algoritma *Support Vector Machine* dengan perbandingan data yang sama dengan algoritma sebelumnya dan algoritma ini menunjukkan hasil *accuracy* 73.86% *precision* 73.92% *Recall* 73.86% *F1-score* 73.24%.

Tabel 6. Hasil Perbandingan Algoritma

| Algoritma | Accuracy (%) |
|-------------------------------|--------------|
| <i>Decision Tree</i> | 64.71% |
| <i>Support Vector Machine</i> | 73.86% |

Tabel 6 menunjukkan hasil perbandingan antara dua algoritma dalam melakukan analisis sentimen dengan menggunakan dataset tentang kejadian yang berada pada jalur Gaza. Berdasarkan tabel diatas algoritma *Support Vector Machine* memiliki akurasi yang lebih tinggi dengan Tingkat akurasi sebesar 73.86% dibanding algoritma *Decision Tree* yang memiliki Tingkat akurasi sebesar 64.71%, dengan Tingkat akurasi 73.86% dapat dikategorikan baik dalam melakukan analisis sentimen, dengan hasil akurasi yang sudah didapatkan, dapat dikatakan bahwa pandangan masyarakat mengenai Gaza ialah kebanyakan netral, lalu disusul pandangan yang positif, kemudian yang terakhir negative, sesuai dengan Tingkat akurasi yang sudah digunakan pada penelitian ini. Dari hasil penelitian ini ada beberapa batasan masalah yang didapatkan dimana dalam keterbatasan data, karena data yang digunakan hanya sebanyak 1528 *tweet*, dan penelitian ini juga mempunyai keterbatasan algoritma dimana dalam penelitian ini hanya menggunakan dua algoritma saja yaitu algoritma *Support Vector Machine* dan *Decision Tree*.

3.2 Pembahasan

Penelitian menggunakan algoritma *Support Vector Machine* (SVM) dan *Decision Tree*, masing-masing dengan kelebihan serta keterbatasan. SVM unggul dalam menangani data berdimensi tinggi dan efisien dalam memori, tetapi membutuhkan waktu pelatihan yang lama serta memerlukan komputasi yang tinggi (Dewi & Chen, 2020). *Decision Tree* menawarkan kemudahan interpretasi melalui struktur visual, meskipun sangat bergantung pada kualitas data pelatihan. Ketidakteraturan data dapat memengaruhi kinerjanya secara signifikan (Irawan, 2021). Data diperoleh dari *Twitter* dengan kata kunci seperti "Hamas" dan "Gaza." Sebanyak 1.924 *tweet* dikumpulkan melalui *crawling* menggunakan *auth_token* untuk menghubungkan program dengan platform tersebut. Setelah proses pembersihan, pra-pemrosesan, dan pelabelan, data yang digunakan untuk analisis berjumlah 1.528 *tweet*. Tahap pra-pemrosesan mencakup normalisasi, penghapusan *stopword*, tokenisasi, dan stemming (Adam, 2022). Langkah ini memastikan data terorganisir dan siap digunakan untuk analisis lebih lanjut (ivosights, 2023). Pelabelan dilakukan dengan pustaka *Natural Language Toolkit* (NLTK), menghasilkan tiga kategori sentimen: positif, netral, dan negatif. Sebagian besar data termasuk kategori netral, diikuti oleh sentimen positif, lalu negatif.

Pengujian menunjukkan bahwa SVM memiliki akurasi lebih tinggi dibandingkan *Decision Tree*, dengan nilai 73,86% berbanding 64,71%. Selain itu, SVM juga unggul pada precision, recall, dan F1-score, yang menunjukkan performanya dalam mengolah data yang lebih kompleks. Dominasi sentimen netral menunjukkan bahwa sebagian besar masyarakat memiliki pandangan yang seimbang terhadap isu di Gaza, sementara sentimen positif dan negatif memiliki proporsi lebih kecil. Hasil ini sejalan dengan penelitian lain. Singgalen (2023) menunjukkan keunggulan SVM pada analisis sentimen ulasan wisata dengan akurasi 98,98%. Zy *et al.* (2023) menunjukkan bahwa *Decision Tree* lebih efektif untuk mendeteksi ancaman keamanan jaringan dengan presisi 99%. Hal ini mengindikasikan bahwa performa algoritma sangat bergantung pada jenis dataset dan tujuan penelitian. Pada penelitian ini, dataset yang digunakan fokus pada opini publik terkait isu sosial-politik, berbeda dengan dataset teknis seperti ulasan wisata atau ancaman jaringan. Penelitian oleh Tuhuteru (2020) juga menunjukkan bahwa SVM menghasilkan distribusi sentimen yang seimbang dalam analisis kebijakan sosial. Penelitian ini konsisten dengan Oktavia *et al.* (2023), yang menemukan dominasi sentimen netral dalam analisis isu kebijakan publik. Batasan penelitian meliputi jumlah dataset yang relatif kecil dan hanya menggunakan dua algoritma. Untuk meningkatkan hasil, dataset yang lebih besar dan algoritma berbasis *deep learning* seperti LSTM dapat digunakan pada penelitian mendatang. Selain itu, fokus pada konflik Gaza relevan dengan kajian sejarah konflik Palestina-Israel (Hadi, 2021; BBC News Indonesia, 2023), yang menekankan pentingnya pemahaman opini publik terhadap isu tersebut. SVM menunjukkan performa lebih baik dibandingkan *Decision Tree* pada analisis sentimen terhadap data *Twitter* terkait konflik Gaza, dengan akurasi mencapai 73,86% dibandingkan 64,71%. Sebagian besar data menunjukkan sentimen netral, yang mencerminkan pandangan publik yang cenderung tidak ekstrem. Proses pembersihan dan pra-pemrosesan berhasil meningkatkan kualitas data, sehingga analisis dapat dilakukan secara lebih efektif. Penelitian ini membuka peluang penggunaan algoritma lain untuk analisis lanjutan di masa depan.

4. Kesimpulan dan Saran

Penelitian ini menghasilkan analisis sentiment dengan menunjukkan data sebanyak 805 netral, 415 positif, dan 308 negatif, jadi dapat disimpulkan bahwa pandangan masyarakat terhadap kejadian dijalar gaza lebih banyak yang netral, kemudian positif lalu disusul negative dan sudah diterapkan juga dua algoritma untuk mengukur tingkat akurasinya, yang dimana algoritma *Support Vector Machine* dengan akurasi 73.86 yang kemudian disusul dengan algoritma *Decision Tree* dengan akurasi 64.71%. Saran untuk penelitian ini selanjutnya ialah;

- 1) Menggunakan dataset yang lebih banyak dan melakukan *preprocessing* dengan lebih akurat.
- 2) Melakukan perbandingan antara algoritma *machine learning* dan *deep learning*.

5. Daftar Pustaka

- Adryamarthanino, V. (2021). Kenapa Jalur Gaza diperebutkan Israel dan Palestina?. *Kompas.com*.
https://www.kompas.com/stori/read/2021/12/19/110000279/kenapa-jalur-gaza-diperebutkan-israel-dan-palestina-?page=all#google_vignette.
- BBC News Indonesia. (2023). Sejarah konflik Palestina-Israel, pertikaian berkepanjangan yang berlangsung puluhan tahun. *BBC.com*.
<https://www.bbc.com/indonesia/articles/cjr0pz20z7po>.
- Dewi, C., & Chen, R. C. (2019). Integrating real-time weather forecasts data using OpenWeatherMap and Twitter. *International Journal of Information Technology and Business*, 1(2), 48-52.

- Dewi, C., & Chen, R. C. (2020). Decision making based on IoT data collection for precision agriculture. *Intelligent Information and Database Systems: Recent Developments 11*, 31-42. https://doi.org/10.1007/978-3-030-14132-5_3.
- Dewi, C., Zandrato, J., & Christanto, H. J. (2024). Original Research Article Improvement of support vector machine for predicting diabetes mellitus with machine learning approach. *Journal of Autonomous Intelligence*, 7(2). <https://doi.org/10.32629/jai.v7i2.888>.
- Dharma, R. (2022). Data Preprocessing: Pengertian, Manfaat, dan Tahapan Kerjanya.
- Irawan, Y. (2021). Penerapan Algoritma Decision Tree C4. 5 Untuk Memprediksi Kelayakan Calon Pendoron Melakukan Donor Darah Dengan Klasifikasi Data Mining. *JTIM: Jurnal Teknologi Informasi dan Multimedia*, 2(4), 181-189. <https://doi.org/10.35746/jtim.v2i4.75>.
- Marpaung, V. P., Sihombing, G. A., Maulida, H., Ridho, A., & Ardianto, B. (2024). Serangan Militer Israel di Jalur Gaza: Pertanggung jawaban Pidana Berdasarkan Statuta Roma Mahkamah Pidana Internasional. *Aliansi: Jurnal Hukum, Pendidikan dan Sosial Humaniora*, 1(5), 18-28. <https://doi.org/10.62383/aliansi.v1i5.371>.
- Ramadhan, M. A., & Wahyudin, M. I. (2022). Analisis Sentimen Mengenai Keberhasilan Indonesia di Ajang Thomas Cup 2020 (Studi Kasus Media Sosial Twitter) Menggunakan Metode Naïve Bayes dan Decision Tree. *Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi)*, 6(4), 505-511. <https://doi.org/10.35870/jtik.v6i4.560>.
- Ramlan, R., Satyahadewi, N., & Andani, W. (2023). Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM. *Jambura Journal of Mathematics*, 5(2), 431-445. <https://doi.org/10.34312/jjom.v5i2.20860>.
- Sejati, W., Bist, A. S., & Tambunan, A. (2023). Pengembangan analisis sentimen dalam rekayasa software engineering menggunakan tinjauan literatur sistematis. *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, 2(1), 95-103. <https://doi.org/10.33050/mentari.v2i1.377>
- Singgalen, Y. A. (2023). Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 4(1), 323-332. <https://doi.org/10.30865/klik.v4i1.1153>.
- Tuhuteru, H. (2020). Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine. *Journal Information System Development (ISD)*, 5(2).
- Zy, A., Sasongko, A. T., & Kamalia, A. Z. (2023). Penerapan Naïve Bayes Classifier, Support Vector Machine, dan Decision Tree untuk Meningkatkan Deteksi Ancaman Keamanan Jaringan. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(1), 610-617. <https://doi.org/10.30865/klik.v4i1.1134>.