

Sentiment Analysis of X Application Users on the Conflict Between Israel and Palestine Using Support Vector Machine Algorithm

Fadhillah Meisya Carina, Admi Salma*, Dony Permana, Zamahsary Martha

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia.

*Corresponding author: admisalmaal@fmipa.unp.ac.id

Submitted : 03 Mei 2024

Revised : 21 Mei 2024

Accepted : 30 Mei 2024

ABSTRACT

The conflict between Israel and Palestine is the Middle East's longest-running conflict since 1917 and is still ongoing today. This is one of the international conflicts that involves many Arab countries and Western countries in the dispute. The conflict between Israel and Palestine has caused countries in the world to be divided into two camps, namely the pro Palestinian independence camp and the contra camp. The impact of this conflict also creates polarization among Indonesians and forms diverse public opinions on the social media application X. The purpose of this research is to find out how the classification of sentiment of X application users affects the conflict between Israel and Palestine. An analysis that is utilized to convert text-based public opinion data into information is sentiment analysis. The chosen algorithm to separate data classes is the Support Vector Machines algorithm, which can classify data by determining the best hyperplane to provide a separator between opinions that are pro Israel or pro Palestine. After the preprocessing stage, 1000 tweets data were obtained with 800 training data and 200 testing data. The accuracy rate is 93%, precision is 92.93%, recall is 100%, and f-measure is 96.33%. From the results of testing 200 data points, there were 198 pro Palestine opinions and 2 pro Israel opinions, so that it might be said that more individuals favor or support Palestinian independence in the conflict that occurred between Israel and Palestine.

Keywords: Israel, Palestine, Sentiment Analysis, Support Vector Machine, X.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Menurut Simanjorang dkk (2023), perebutan tanah di wilayah Palestina adalah sumber sebenarnya dari konflik antara Israel dan Palestina, meskipun sebagian individu di seluruh dunia percaya bahwa ini merupakan konflik agama. Menurut data resmi Pemerintah Israel dan data dari Kementerian Kesehatan Gaza serta *United Nations Office for the Coordination of Humanitarian Affairs* (OCHA), secara keseluruhan terdapat 15.767 korban jiwa dari pihak Palestina selama periode 7 Oktober - 3 Desember 2023 dan korban jiwa Israel yang jumlahnya sekitar 1.275 orang. Menurut Putra (2023), terdapat dua faksi yang muncul dari konflik antara Israel dan Palestina ini yaitu faksi yang mendukung kemerdekaan Palestina, termasuk di dalamnya adalah Indonesia, Cina, Rusia, dan Korea Selatan. Serta faksi yang menentangnya yaitu seperti Jerman, Inggris, Prancis, dan Amerika Serikat. Dampak dari konflik Israel dan Palestina ini juga menciptakan polarisasi di antara masyarakat Indonesia dan membentuk opini publik yang beragam khususnya di media sosial X mulai dari tanggal 7 Oktober 2023 saat konflik antara dua negara ini kembali memanas.

Pada tahun 2006, Jack Dorsey, Biz Stone, Evan Williams, dan Noah Glass mendirikan platform media sosial X yang kemudian dibeli oleh Elon Musk pada tahun 2022 lalu. Menurut Mustaqilillah dkk (2023), Indonesia terbilang sangat pesat perkembangannya dalam penggunaan media sosial X, terdapat 500 juta *tweet* atau kicauan oleh pengguna X per harinya yang digunakan untuk berbagi informasi berupa fakta dan opini serta mengunggah informasi pribadi. Konten *tweet* juga dapat menyampaikan emosi (Mahbubah, 2019). *Web scraping* adalah metode yang digunakan untuk mendapatkan materi tekstual, seperti data atau informasi dari situs *web* seperti *tweet*. *Tweet* dan data opini lainnya dapat diasimilasi menjadi sebuah informasi jika dilakukan sebuah analisis untuk menentukan apakah opini tersebut mendukung atau tidak mendukung terhadap suatu fenomena. Analisis sentimen adalah jenis analisis yang dapat dilakukan untuk mengekstrak informasi dari *tweet* tersebut.

Menurut Nomleni (2015), analisis sentimen dilakukan untuk mengetahui sikap emosional masyarakat dalam argumen yang diklasifikasikan sehingga argumen tersebut masuk pada kelompok sentimen yang sama. Terdapat beberapa algoritma klasifikasi yang dapat digunakan seperti *Naïve Bayes Classifier* (NBC), *Decision Tree*, K-NN, dan *Support Vector Machines* (SVM). Algoritma yang akan digunakan untuk pengklasifikasian data adalah algoritma SVM. Ide utama dari klasifikasi data SVM ialah mencari tahu *hyperplane* mana yang akan memberikan dua kelas yang telah ditentukan sebelumnya suatu ruang atau jarak yang ukurannya paling kecil satu sama lain. (Kowalczyk, 2017).

Dalam sebuah penelitian yang berkaitan dengan analisis sentimen oleh Fajri dkk (2021), mengklasifikasikan opini publik mengenai konflik antara Israel dan Palestina tanpa mempertimbangkan keberpihakan sentimen kepada salah satu pihak dengan memanfaatkan algoritma NBC dan SVM. Oleh karena itu pada penelitian ini akan dilakukan analisis sentimen dengan kelas sentimen keberpihakan pengguna X terhadap salah satu pihak. Berdasarkan hasil pengujian, algoritma SVM memberikan nilai akurasi yang lebih tinggi yaitu sebesar 80%, *precision* 79%, serta untuk nilai *recall* juga dihasilkan nilai tertinggi pada algoritma SVM yaitu sebesar 72%. Algoritma SVM lebih unggul daripada algoritma NBC dalam hal klasifikasi dataset, yang dibuktikan dengan nilai akurasi, presisi, dan *recall* tersebut. Penelitian lainnya dilakukan oleh Guaia dkk (2019), mengenai pengklasifikasian sentimen *review* ponsel menggunakan algoritma NBC, SVM, *Decision Tree*, dan *Random Forest*. Pada penelitian tersebut diperoleh bahwa SVM memperoleh nilai tertinggi untuk semua nilai akurasi, *recall*, *precision*, dan *F1-score* yaitu sebesar 89% dibanding dengan nilai akurasi, *recall*, *precision*, dan *F1-score* ketiga algoritma lainnya.

Berdasarkan penelitian terdahulu, dapat diketahui bahwa menggunakan algoritma *Support Vector Machines* pada analisis sentimen dapat memberikan akurasi yang lebih tinggi dalam memodelkan klasifikasi sentimen. Penelitian ini bertujuan untuk mengetahui persepsi pengguna aplikasi X yang memberikan atau mengunggah komentar berbahasa Indonesia mengenai keberpihakan mereka kepada salah satu pihak yaitu Israel ataupun Palestina terkait konflik antara Israel dan Palestina mulai dari tanggal 7 Oktober 2023 sampai dengan 30 November 2023 menggunakan algoritma SVM. Manfaat dari penelitian ini juga dapat membantu pemerintah dalam mengantisipasi kemungkinan terjadinya peristiwa sosial seperti protes atau aksi berskala besar sebagai reaksi atas konflik antara Israel dan Palestina.

II. METODE PENELITIAN

A. Jenis Penelitian dan Sumber Data

Penelitian ini merupakan penelitian terapan dengan menggunakan data sekunder berupa data komentar yang berbentuk opini ataupun pendapat yang biasa disebut sebagai *tweet* berbahasa Indonesia yang terdapat dalam aplikasi media sosial X. *Tweet* yang digunakan adalah *tweet* yang mengandung kata "Israel" ataupun kata "Palestina". Data *tweet* yang digunakan merupakan historis unggahan *tweet* dari tanggal 7 Oktober 2023 sampai dengan 30 November 2023. Berikut variabel penelitian yang digunakan ditampilkan pada Tabel 2.

Tabel 2. Variabel Penelitian

Variabel	Keterangan
Komentar (X)	Unggahan komentar pengguna aplikasi X terkait konflik antara Israel dan Palestina yang telah dilakukan proses pembobotan kata.
Sentimen (Y)	Persepsi emosi keberpihakan dari unggahan komentar pengguna aplikasi X terkait konflik antara Israel dan Palestina.

Penelitian ini dilakukan dengan menggunakan bantuan dari *software* Python yang tersedia pada Google *Colab*. Pemilihan data dilakukan dengan menggunakan metode *web scraping*, dimana *web scraping* dapat diartikan sebagai metode pengumpulan data dan informasi dari sebuah alamat *website* atau media sosial tertentu kemudian disimpan dalam *file* teks untuk digunakan dalam analisis.

B. Teknik Analisis Data

Adapun langkah-langkah yang dilakukan pada penelitian ini ialah sebagai berikut.

1. Pengumpulan data yang dilakukan dengan menggunakan teknik *web scraping* pada aplikasi media sosial X dengan bantuan *software* Python.
2. Melakukan *Data selection* dan *text preprocessing*

Data selection akan dilakukan yang bertujuan untuk menyeleksi *tweet* mana saja yang dapat disebut sebagai data komentar, kemudian dilakukan *text preprocessing* yang digunakan untuk membersihkan dan menyiapkan data *tweet* untuk proses klasifikasi. Langkah pertama dalam pemrosesan data komentar yang diperoleh dari aplikasi X yang disebut sebagai *tweet* ialah *text preprocessing* yang melibatkan pengorganisasian dokumen menjadi data terstruktur berdasarkan

kebutuhannya sehingga *text mining* dapat memproses lebih lanjut. Menurut Rahman dkk (2021), semakin bersih praproses yang dilakukan, maka kemungkinan besar hasil analisis akan semakin akurat. Terdapat beberapa tahapan *text preprocessing*, sebagai berikut ini:

- a. *Cleaning* adalah tahapan untuk membersihkan data komentar dengan cara menyeleksi kata-kata yang tidak diperlukan, tidak memiliki dampak emosional, atau kata yang tidak memengaruhi sentimen.
- b. *Tokenize* adalah penghapusan tanda baca dan pemecahan kalimat ke dalam bentuk token atau potongan kata.
- c. *Transform case* adalah tahapan merubah kalimat data teks menjadi kata yang seragam, tahapan ini dapat berperan dalam penyamarataan penggunaan bentuk standar atau huruf kecil.
- d. *Stopword removal* adalah tahapan untuk menghapus kata-kata yang sering muncul tapi tidak memiliki arti penting dan maknanya tidak berpengaruh pada sentimen.
- e. *Stemming* adalah sebuah tindakan untuk menghilangkan kata imbuhan sehingga kata-kata yang tidak baku menjadi kata baku.

3. Pelabelan data

Pelabelan data dilakukan secara manual untuk memberikan label pada setiap data komentar yang terkait konflik antara Israel dan Palestina dengan kelas sentimen yaitu 'Pro Israel' dan 'Pro Palestina'. Pelabelan data juga meliputi proses untuk memisahkan data menjadi dua bagian yaitu data *training* dan data *testing* dengan perbandingan 80% data *training* dan 20% data *testing*.

4. Pembobotan kata

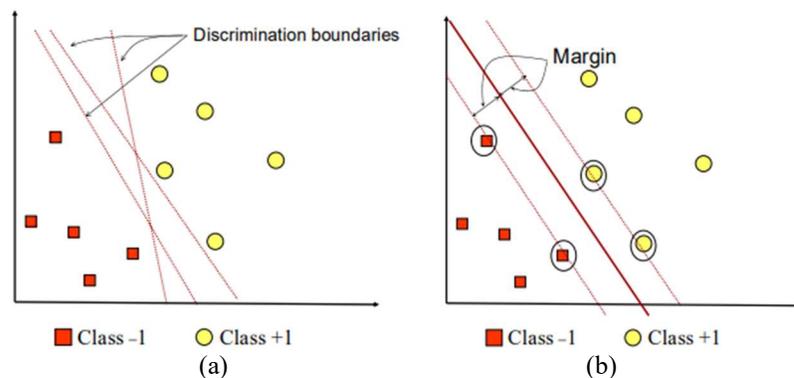
Menurut Robertson (2004), dengan mengalikan nilai *Term Frequency* (TF) dengan *Inverse Document Frequency* (IDF) kata-kata yang terorganisir dalam setiap *tweet* dapat diberi bobot. TF akan menampilkan jumlah kata dalam setiap dokumen, sementara IDF menampilkan kelangkaan kemunculan kata. Proses penghitungan bobot melibatkan pencarian kemunculan setiap kata dalam *tweet* terkait konflik antara Israel dan Palestina sehingga data dapat dikonversi menjadi angka dan bisa diolah oleh komputer. Perhitungan bobot tiap token dengan menggunakan metode TF-IDF dalam dokumen dirumuskan pada Persamaan 1.

$$W_{at} = tf_{at} \cdot IDF_t = tf_{at} \cdot \log \left(\frac{N}{n_t} \right) \quad (1)$$

Dimana W_{at} merupakan bobot dari kata t dalam satu dokumen d , tf_{at} merupakan frekuensi kemunculan kata t dalam dokumen d , IDF_t merupakan *Inversed Document Frequency* kata t , N sebagai jumlah semua dokumen d , dan n_t adalah jumlah dokumen d yang mengandung kata t .

5. Penerapan algoritma SVM untuk mengklasifikasikan data dengan metode optimasi parameter *grid search*.

Support Vector Machines (SVM) pertama kali diperkenalkan pada tahun 1992 pada saat *Annual Workshop on Computational Learning Theory* oleh Vladimir Vapnik dan rekan-rekannya Bernhard Boser dan Isabelle Guyon. Secara teoritis, SVM dirancang untuk menemukan *hyperplane* yang optimal dan memecahkan masalah klasifikasi antar dua kelas. (Vapnik & Cortes, 1995). *Hyperplane* adalah sebuah fungsi yang membagi ruang *input* menjadi dua kelas sehingga data komentar yang beragam dapat dikategorikan. Ilustrasi untuk penerapan algoritma SVM ditampilkan pada Gambar 1.



Gambar 1. Ilustrasi Penerapan SVM (Sumber: Nugroho dkk, 2003)

Garis pemisah alternatif digambarkan pada Gambar 1 (a) dan *hyperplane* yang membentang tepat di antara dua kelas ditunjukkan pada Gambar 1 (b). Margin mengukur jarak antara pola terdekat setiap kelas dengan *hyperplane*.

Istilah *support vector* mengacu pada sekumpulan data yang terletak di margin. Selain memisahkan data non linear secara linear, kernel juga digunakan dalam SVM dengan tujuan untuk mengubah data menjadi ruang berdimensi tinggi. *Radial Basis Function* (RBF), polinomial, *sigmoid*, dan kernel linier adalah beberapa pilihan fungsi kernel yang tersedia. Menurut Nugroho dkk (2020), Persamaan 2 menyatakan *hyperplane* pemisah dengan w sebagai bobot vektor, x_i merupakan data ke- i , dan b sebagai bias.

$$w \cdot x_i + b = 0 \quad (2)$$

Nilai x yang sesuai dengan kelas sentimen -1 (sampel negatif) dapat ditemukan dengan menggunakan persamaan 3 berikut.

$$w \cdot x_i + b \leq 0 \quad (3)$$

Nilai x yang sesuai dalam kelas sentimen +1 (sampel positif) dapat ditemukan dengan menggunakan persamaan 4 berikut.

$$w \cdot x_i + b \geq 0 \quad (4)$$

Adapun langkah-langkah klasifikasi SVM ialah sebagai berikut:

- Menginisialisasi dan mengoptimalkan parameter yang akan diaplikasikan yaitu Λ (λ), α (α_x), γ (γ), C (C), ϵ (ϵ), dan fungsi kernel (K).
- Menghitung nilai Matriks Hessien (D_{ij}) dengan menggunakan Persamaan 5.

$$D_{ij} = y_i y_j K(x_i, x_j) + \lambda^2 \quad (5)$$

dengan y sebagai kelas data ke- i dan K merupakan fungsi kernel.

- Perhitungan nilai *error* (E_i) dengan menggunakan persamaan 6.

$$E_i = \sum_{j=1}^n \alpha_x D_{ij} \quad (6)$$

- Perhitungan nilai *delta alpha* ($\delta\alpha_i$) dengan menggunakan Persamaan 7.

$$\delta\alpha_i = \min \{ \max [\gamma(1 - E_i), -\alpha_x], C - \alpha_x \} \quad (7)$$

- Perhitungan nilai *alpha* baru (α_i) dengan menggunakan Persamaan 8.

$$\alpha_i = \alpha_x + \delta\alpha_i \quad (8)$$

- Menghitung nilai w^+ (bobot nilai *alpha* terbesar pada kelas pro Palestina) dan w^- (bobot nilai *alpha* terbesar pada kelas pro Israel) menggunakan Persamaan 9 dan Persamaan 10.

$$w^+ = \sum_{i=1}^n \alpha_i y_i (K(x_i, x^+)) \quad (9)$$

$$w^- = \sum_{i=1}^n \alpha_i y_i (K(x_i, x^-)) \quad (10)$$

- Perhitungan nilai bias (b) dengan menggunakan Persamaan 11.

$$b = -\frac{1}{2}(w^+ + w^-) \quad (11)$$

- Pengujian terhadap data *testing* kemudian perhitungan keputusan $h(x)$ menggunakan Persamaan 12.

$$h(x) = \begin{cases} +1, & \text{if } wx + b \geq 0 \\ -1 & \text{if } wx + b < 0 \end{cases} \quad (12)$$

Apabila hasil perhitungan keputusan memiliki nilai 0 atau lebih besar, maka tanda *sign* $h(x)$ adalah +1 yang artinya menunjukkan kelas pro Palestina. Sebaliknya, jika hasil perhitungan keputusan memiliki nilai kurang dari 0, tanda *sign* $h(x)$ adalah -1. Perhitungan *sign* $h(x)$ mengikuti Persamaan 13.

$$h(x) = w \cdot x + b = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \quad (13)$$

Menurut Zakrani dkk (2018), peningkatan akurasi secara signifikan dipengaruhi oleh pemilihan parameter. Oleh karena itu, pemilihan nilai terbaik untuk parameter pembelajaran diperlukan untuk membentuk model yang akurat. Untuk menemukan parameter yang ideal, *grid search* membentuk kisi-kisi dari rentang parameter yang harus dioptimalkan dan melintasi setiap titik tersebut. Menemukan seperangkat *hyperparameter* yang optimal adalah hal yang diperlukan untuk memungkinkan pengklasifikasi memprediksi data komentar yang tidak diketahui sebelumnya dengan akurat. Penerapan *grid search* pada *dataset* kanker payudara dengan model SVM memberikan hasil yang jauh lebih baik dibandingkan tanpa menerapkan *grid search*, menurut penelitian Deshwal & Sharma (2019).

6. Evaluasi hasil

Menurut Dutt dkk (2018), suatu klasifikasi dikatakan benar apabila sesuatu yang diprediksi benar dan pada kenyataannya memang benar-benar terjadi atau sesuai dengan hasil prediksi. Salah satu cara paling komprehensif untuk merepresentasikan hasil evaluasi klasifikasi khususnya data komentar adalah dengan menggunakan *confusion matrix* atau matriks konfusi. Pada intinya, menentukan apakah suatu objek tersebut benar atau salah merupakan pengertian dari evaluasi model klasifikasi (Dewi, 2019). Di dalam matriks konfusi terdapat empat kasus seperti yang disajikan pada Tabel 1.

Tabel 1. Confusion Matrix

		Prediksi	
		True	False
Aktual	True	True Positive (TP)	False Negative (FN)
	False	False Positive (FP)	True Negative (TN)

Terdapat empat keluaran dalam perhitungan rumus untuk merangkum informasi di dalam matriks konfusi, yaitu akurasi, presisi, *recall*, dan *F-measure*. Nilai persentase yang ditemukan dan ditentukan sebagai benar berdasarkan total data disebut akurasi. Presisi adalah tingkat kesesuaian antara prediksi model dan data yang dimiliki. Nilai yang menunjukkan kemampuan model untuk mengambil kembali sebuah informasi disebut *recall*. *F-measure* adalah ukuran lain dari kinerja model yang menggabungkan presisi dan *recall*, dimana kinerjanya baik pada kumpulan data yang tidak seimbang. Nilai akurasi, presisi, *recall*, dan *F-measure* dihitung berturut-turut dengan menggunakan persamaan 14, 15, 16, dan 17 berikut.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$\text{Presisi} = \frac{TP}{TP + FP} \tag{15}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{16}$$

$$F \text{ measure} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \tag{17}$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan teknik *web scraping* pada aplikasi media sosial X dengan bantuan *software* Python. Python yang telah dihubungkan dengan aplikasi sosial media X akan mengumpulkan *tweet* yang terkait dengan kata kunci ‘Israel’ ataupun ‘Palestina’ berbahasa Indonesia. Hasilnya data yang terkumpul ialah sebanyak 3.406 *tweet* yang tersimpan dalam format CSV.

B. Text Preprocessing

Data selection dilakukan untuk menyeleksi *tweet* mana saja yang dapat disebut sebagai data komentar terkait konflik antara Israel dan Palestina. Komentar diartikan sebagai suatu ulasan atau tanggapan atas sebuah fenomena. Setelah dilakukan tahapan *data selection* diperoleh hasil sebanyak 1000 data komentar terkait konflik ini. Kemudian

tahapan *text preprocessing* akan dilakukan, berikut beberapa data hasil *text preprocessing* menggunakan *software* Python disajikan pada Tabel 3.

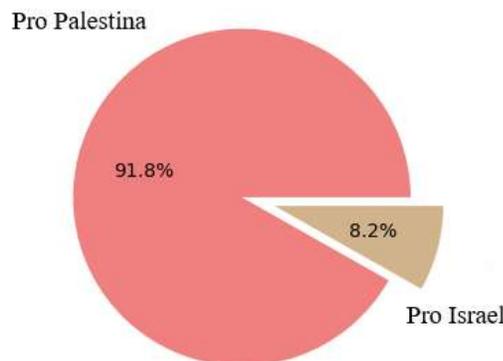
Tabel 3. Contoh *Text Preprocessing*

Teks Awal	Teks Hasil <i>Preprocessing</i>
Ayuk dukung Palestina, warga dan mahasiswa sudah memulai untuk boikot produk yang terafiliasi dan pro Israel. - https://t.co/qexefi208k @CNNIndonesia boikot malah meningkatkan usaha dalam negeri!!! Hello produk2 pendukung zionis bagaimana? Rugi dikit ribut!!! Palestina dibantai bisu. @Zayn gue tetap dukung Israel, mereka punya hak untuk mengambil kembali tanahnya.	dukung palestina warga mahasiswa sudah boikot produk afiliasi israel boikot tingkat usaha negeri produk dukung zionis rugi rebut palestina bantai bisu tetap dukung israel punya hak ambil kembali tanah

Tabel 3 menunjukkan hasil berupa data komentar berbentuk *tweet* yang telah melalui tahapan *text preprocessing* dengan tujuan untuk membersihkan dan menyiapkan data komentar tersebut menjadi data yang lebih baik serta lebih bersih untuk dilakukan analisis sentimen menggunakan algoritma klasifikasi SVM.

C. Pelabelan Data

Pelabelan data dilakukan secara manual untuk memberikan label pada setiap data komentar yang terkait konflik antara Israel dan Palestina dengan kelas sentimen yaitu ‘Pro Israel’ dan ‘Pro Palestina’. Komentar ‘Pro Israel’ merupakan opini dari pengguna aplikasi X yang mendukung atau setuju dengan tindakan yang dilakukan oleh Israel sekaligus menganggap bahwa Israel merupakan korban dalam konflik ini, begitupula sebaliknya. Gambar 2 menunjukkan visualisasi frekuensi antar kelas sentimen Pro Israel dan Pro Palestina pada data yang diperoleh.



Gambar 2. Diagram Lingkaran Kelas Sentimen

Pada Gambar 2 ditunjukkan bahwa dari total 1000 data komentar yang diperoleh didapatkan hasil yaitu sebanyak 918 data komentar diklasifikasikan sebagai kelas sentimen Pro Palestina dan 82 sisanya diidentifikasi sebagai kelas sentimen Pro Israel.

D. Pembobotan Kata

Metode ekstraksi fitur yang digunakan untuk pembobotan kata yaitu metode *Term Frequency - Inverse Document Frequency* (TF-IDF) menggunakan Persamaan 1. Semakin sering sebuah kata muncul dalam dokumen, maka semakin besar bobot TF yang diperoleh. Sementara itu, IDF membantu dalam membedakan dokumen. Tingkat kesamaan antara dokumen dengan kata kunci yang sama akan meningkat seiring dengan semakin besarnya nilai bobot, atau sebaliknya. Untuk menggambarkan kata yang sering muncul pada keseluruhan data dapat divisualisasikan dengan menggunakan grafik *wordcloud* yang disajikan pada Gambar 3.



Gambar 3. Wordcloud

Berdasarkan *wordcloud* pada Gambar 3 dapat dilihat kata-kata yang paling sering muncul secara keseluruhan, semakin besar ukuran teks yang ditampilkan pada *wordcloud* menunjukkan semakin sering pula kata tersebut digunakan atau ditemukan pada data komentar. Pada Gambar 3 dapat dilihat bahwa kata-kata Israel, Palestina, dan Hamas merupakan kata yang mendominasi pada *tweet* karena pada *wordcloud* kata tersebut memiliki ukuran teks yang lebih besar dari pada kata lainnya.

E. Support Vector Machine

Setelah melalui tahap pelabelan data dan pembobotan kata, klasifikasi algoritma SVM akan dilakukan dengan menggunakan *grid search* untuk optimasi parameter pada data pelabelan yang telah dibagi menjadi data *training* dan data *testing*. Diperoleh jumlah data *training* sebanyak 800 data dan jumlah data *testing* sebanyak 200 data. Tabel 4 menyajikan perbandingan nilai akurasi yang diperoleh dari penggunaan beberapa fungsi kernel.

Tabel 4. Perbandingan Nilai Akurasi Fungsi Kernel

Kernel	Nilai Akurasi
Linear	93%
Polinomial	92%
RBF	92%
Sigmoid	92%

Terlihat pada Tabel 4 bahwasanya nilai akurasi untuk keempat fungsi kernel yang diperoleh sama-sama menghasilkan nilai akurasi yang tinggi, dimana berdasarkan pemilihan fungsi kernel terbaik pada SVM menggunakan algoritma *grid search* dimiliki oleh kernel linear karena menghasilkan nilai akurasi tertinggi dibanding dengan fungsi kernel lainnya. Evaluasi hasil klasifikasi akan dilakukan dengan tujuan untuk menentukan nilai kegunaan dan keakuratan dari model SVM yang diperoleh. Hasil pengujian dari data *testing* menggunakan algoritma SVM dengan menerapkan metode *grid search* yang dilakukan pada data *tweet* terkait konflik antara Israel dan Palestina diperoleh *confusion matrix* yang ditampilkan pada Tabel 5.

Tabel 5. Hasil Confusion Matrix

		Prediksi	
		Pro Palestina	Pro Israel
Aktual	Pro Palestina	184	0
	Pro Israel	14	2

Tabel 5 menyajikan *confusion matrix* yang mewakili setiap kelas klasifikasi Pro Palestina dan klasifikasi Pro Israel. Dari *confusion matrix* dapat dijelaskan bahwa model SVM dapat mengklasifikasikan secara benar sebesar 184 *tweet* sebagai Pro Palestina dan 2 *tweet* sebagai Pro Israel. Sebaliknya model salah dalam memprediksi 14 data kedalam kelas Pro Palestina yang seharusnya Pro Israel. Berdasarkan nilai *confusion matrix* pada Tabel 5 diperoleh nilai performa hasil klasifikasi dengan menggunakan Persamaan 14, 15, 16, dan 17 yang disajikan pada Tabel 6.

Tabel 6. Performa Hasil Klasifikasi

Akurasi	Presisi	Recall	F-measure
93%	92,93%	100%	96,33%

Tabel 6 menyajikan nilai presisi sebesar 92,93% yang artinya jumlah perbandingan data yang diprediksi Pro Palestina ternyata benar-benar Pro Palestina, nilai *recall* sebesar 100% yang artinya perbandingan jumlah data yang benar Pro Palestina dalam pengklasifikasian terhadap jumlah data yang diklasifikasikan ke dalam kelas Pro Palestina, nilai *F-measure* sebesar 96,33% yang menunjukkan perbandingan nilai rata-rata dari presisi dan *recall* yang telah dibobotkan, dan nilai akurasi sebesar 93% yang artinya tingkat keberhasilan model SVM dalam mengklasifikasi data komentar secara benar dalam penelitian ini ialah sangat baik.

IV. KESIMPULAN

Pengumpulan data *tweet* pada aplikasi media sosial X terkait konflik antara Israel dan Palestina mengumpulkan sebanyak 3.406 *tweet* pada periode 7 Oktober 2023 sampai dengan 30 November 2023. Setelah tahapan *data selection* dilakukan diperoleh hasil sebanyak 1000 data komentar terkait konflik ini. Kemudian didapatkan jumlah data *training* sebanyak 800 data komentar dan jumlah data *testing* sebanyak 200 data komentar. Berdasarkan hasil yang diperoleh dari analisis sentimen pengguna aplikasi X menggunakan algoritma SVM didapatkan kesimpulan bahwa algoritma SVM dapat melakukan pengklasifikasian opini mengenai konflik antara Israel dan Palestina yang ditunjukkan oleh hasil evaluasi klasifikasi dengan nilai akurasi sebesar 93%, presisi sebesar 92,93%, *recall* sebesar 100% dan *F-measure* sebesar 96,33%.

Berdasarkan nilai yang dihasilkan oleh *confusion matrix* dengan mengacu kepada sentimen terhadap konflik antara Israel dan Palestina menggunakan data *testing*, dengan menerapkan algoritma SVM dan *grid search* sebagai optimasi parameter diperoleh kelas sentimen Pro Palestina sebanyak 198 dari 200 data *testing* dan kelas sentimen Pro Israel sebanyak 2 dari 200 data *testing*. Berdasarkan hasil analisis tersebut dapat diperoleh kesimpulan bahwa sebagian besar pengguna aplikasi X memihak atau mendukung kemerdekaan Palestina. Penelitian selanjutnya diharapkan dapat menggunakan algoritma klasifikasi selain *Support Vector Machine*, sehingga kinerja klasifikasinya dapat dibandingkan dengan cara melihat keakuratan model yang diperoleh.

DAFTAR PUSTAKA

- Deshwal, V., & Sharma, M. (2019). Breast Cancer Detection Using SVM Classifier with Grid Search Technique. *International Journal of Computer Application*, 178(31), 18-23.
- Dewi, S. (2019). Komparasi Metode Algoritma Data Mining pada Prediksi Uji Kelayakan *Credit Approval* pada Calon Nasabah Kredit Perbankan. *Jurnal Khatulistiwa Informatika*, 7(1), 59-65.
- Dutt, S., Chandramouli, S., & Das, A. K. (2018). *Machine Learning*. India: Pearson Education.
- Fajri, H. R. A., Sinaga, R. B., Mubarak, H., Pangestu, A. D., & Prasvita, D. S. (2021). Analisis Sentimen Pengguna Twitter terhadap Konflik antara Palestina dan Israel Menggunakan Metode *Naïve Bayesian Classification* dan *Support Vector Machine*. *SENAMIKA, Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya*, 166-175.
- Guaia, M., Silva, R. R., & Bernardino, J. (2019). Comparison of Naïve Bayes, Support Vector Machine, Decision Trees, and Random Forest on Sentiment Analysis. *11th International Conference on Knowledge Discovery and Information Retrieval*, 525-531.
- Kowalczyk, A. (2017). *Support Vector Machine Succinctly*. Synconfusion Inc.
- Mahbubah, L. D., & Zuliarso, E. (2019). Analisa Sentimen Twitter pada Pilpres 2019 Menggunakan Algoritma *Naïve Bayes*. *Prosiding SINTAK 2019*, ISBN: 978-602-8557-20-7, (p): 193-199, 2019.
- Mustaqililah, R., Widyaningtyas, O., & Wantoro, T. (2023). Efektivitas Penggunaan Twitter Sebagai Sarana Peningkatan Berpikir Kritis Mahasiswa Ilmu Komunikasi. *MUKASI: Jurnal Ilmu Komunikasi*, 2(1), 18-28.
- Natasuwarna, A. P. (2020). Seleksi Fitur *Support Vector Machine* pada Analisis Sentimen Keberlanjutan Pembelajaran Daring. *Techno.Com*, 19(4), 437-448.
- Nomleni, P. (2015). *Sentiment Analysis Menggunakan Support Vector Machine (SVM)*. *Seminar Nasional Teknologi & Komunikasi 2015 (Sentika)*, 1-8.
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika. Kuliah Umum IlmuKomputer.Com*.

- Putra, R. M., Yunisa, L., Putra, F. K., Rizki, M. A., & Antoni, H. (2023). Pertanggungjawaban Dewan Keamanan PBB terhadap Penggunaan Bom Fosfor oleh Israel kepada Palestina Sebagai Suatu Kejahatan Perang. *Jurnal Pendidikan Tambusai*, 7(3), 25031-25040.
- Rahman, I. A., Indra, A., Alita, D., & Satya, M. N. (2021). Sentimen Analisis Publik terhadap Kebijakan *Lockdown* Pemerintah Jakarta Menggunakan Algoritma SVM. *JDMSI*, 2(1), 31-37.
- Robertson, S. (2004). Understanding Inverse Document Frequency: on Theoretical Argumentr for IDF. *Journal of Documentation*, 60(5), 503-520.
- Simanjorang, B. M., Putra, B. A. S., Husin, M. H., Bangun, I. B., Zawani, N., Siburian, T. D. N., Perangin-Angin, Z. G., & Prayetno. (2023). Pengaruh Konflik Palestina dengan Israel terhadap Gerakan Perlawanan Hamas dan Dampaknya Bagi Indonesia. *Journal of Law*, 2(2), 24-30.
- Vapnik, V., & Cortes, C. (1995). Support Vector Networks. *Machine Learning*, 20, 273-297.
- Zakrani, A., Najm, A., & Marzak, A. (2018). Support Vector Regression Based on Grid-Search Method for Agile Software Effort Prediction. *Colloquium in Information Science and Technology*, 8(2), 26-32.