

Klasifikasi Sentimen Opini *Metaverse* dari Twitter Menggunakan Algoritma *Support Vector Machine*

Herlawati^{*1}, Adi Muhajirin², Zalfa Izdihar³

^{1,2,3}Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya
Jl. Raya Perjuangan, Margamulya, Bekasi, Jawa Barat, Indonesia, Telp. (021)7231948
¹herlawati@ubharajaya.ac.id, ²adi.muhajirin@dsn.ubharajaya.ac.id,
³zalfa.izdihar19@mhs.ubharajaya.ac.id

*) Corresponding Author

(received: 29-08-23, revised: 08-11-23, accepted: 25-01-24)

Abstract

With the increasing use of Twitter, a real-time social media platform, it has become one of the places or spaces for people to express their opinions about the metaverse. Therefore, the development of a program capable of classifying tweets based on their opinions into positive, negative, and neutral categories is necessary. In conducting sentiment analysis, the Support Vector Machine (SVM) algorithm is used for classification. The results of this research, through testing using a confusion matrix, yield an accuracy rate of 0.83 or 83%, indicating the level of agreement between the model's predictions and the actual outcomes. Additionally, a precision of 0.93 or 93% is obtained, which shows the model's ability to accurately identify positive, negative, and neutral sentiments in tweets, and a recall of 0.83 or 83%, which describes the model's capability to find and classify accurately.

Keyword: *Metaverse, Twitter, Sentiment Analysis, Real-Time, Support Vector Machine (SVM).*

Abstrak

Seiring dengan peningkatan penggunaan Twitter, platform media sosial yang beroperasi secara real-time, telah menjadi salah satu tempat atau ruang bagi masyarakat untuk menyampaikan opini mereka mengenai metaverse. Karena itu, diperlukan pengembangan program yang mampu mengklasifikasikan tweet berdasarkan opini ke dalam kategori positif, negatif, dan netral. Dalam melakukan analisis sentimen, algoritma *Support Vector Machine (SVM)* digunakan untuk melakukan klasifikasi. Hasil dari penelitian ini, melalui pengujian menggunakan matrik konfusi, menghasilkan akurasi yang menunjukkan tingkat kecocokan antara prediksi dan hasil sebenarnya dari model, sebesar 0.83 atau 83%. Selain itu, diperoleh presisi yang berfungsi untuk menunjukkan kemampuan model dalam mengidentifikasi dengan tepat sentimen positif, negatif, dan netral dalam tweet, sebesar 0.93 atau 93%, dan *recall* yang berfungsi untuk menggambarkan kemampuan model untuk menemukan dan mengklasifikasikan secara akurat, sebesar 0.83 atau 83%.

Kata Kunci: *Metaverse, Twitter, Analisis Sentimen, Real-Time, Support Vector Machine (SVM).*

I. Pendahuluan

Metaverse adalah kombinasi dari awalan kata “*meta*” dan digabungkan oleh kata “*universe*” yang dapat dideskripsikan dengan suatu paralel atau lingkungan virtual yang terhubung dengan dunia nyata. *Metaverse* pertama kali dikemukakan oleh seorang novelis yang berasal dari Amerika Neal Stephenson, pada sebuah novel Sci-fi nya yang berjudul pada tahun 1992. Di dalam novel tersebut, Stephenson menginterpretasikan *metaverse* sebagai tempat atau wadah untuk manusia melarikan diri dari rasa sakit yang disebabkan oleh dunia nyata dengan menjelajahi sebuah dunia digital lewat beberapa avatar digital [1].

Metaverse pada akhirnya akan menjadi paradigma baru dalam dunia sosial, ekonomi maupun kebudayaan karena *metaverse* menyediakan platform yang mudah diakses, interaktif, dan tidak terbatas. Konsep *metaverse*

ini tentu adanya manfaat dan juga resiko. Dalam perspektif sosial, ide *metaverse* ini memberikan kesempatan untuk manusia dalam berinteraksi dan bersosialisasi dengan kerabat di seluruh dunia tanpa keterbatasan ekonomi dan protokol travel dan itu akan mendapatkan dampak positif. Dan dalam perspektif kritis, dimaksudkan untuk mengatasi travelling dan interaksi fisik manusia dan kebutuhan manusia lainnya yang memperkuat ikatan dan hubungan sosial [2].

Topik *metaverse* ini menuai banyak perbincangan di beberapa media sosial, salah satu contohnya adalah pada Twitter. Twitter menjadi salah satu tempat atau wadah masyarakat untuk memberikan opininya terhadap *metaverse* tersebut. Opini-opini masyarakat tersebut akan diproses menjadi suatu informasi dengan cara yaitu analisis sentimen. Twitter adalah layanan *micro-blogging* yang dibangun untuk menemukan apa yang terjadi pada setiap saat, di mana saja di dunia. Pesan-pesan di Twitter singkat dan terus menerus dihasilkan, serta cocok digunakan untuk penemuan pengetahuan menggunakan teknik data stream mining [3]. Orang menggunakan situs jaringan sosial seperti Twitter, Facebook, dan Google+ untuk berbagi dan mengungkapkan pendapat mereka tentang berbagai topik, berpartisipasi dalam diskusi dengan berbagai komunitas, dan mengirim pesan ke seluruh dunia. Bidang analisis sentimen terhadap data Twitter telah mengalami banyak kemajuan [4].

Analisis sentimen adalah proses yang bertujuan untuk menentukan apakah teks yang dapat berupa dokumen, kalimat, paragraf, maupun dataset tersebut positif, negatif atau netral [5]. Pada saat ini, opini publik menjadi sumber informasi penting untuk dalam mengambil keputusan pada suatu produk. Algoritma klasifikasi seperti *Naïve Bayes* (NB), *Support Vector Machine* (SVM), dan *K-nearest Neighbour* (K-NN) diusulkan oleh banyak peneliti untuk digunakan pada analisis sentimen [6]. Oleh karena itu, penelitian ini bertujuan untuk mengetahui opini masyarakat terhadap topik *metaverse* dan mengklasifikasikan opini tersebut ke dalam kategori positif, negatif, ataupun netral.

Salah satu penelitian analisis sentimen yang memanfaatkan data media sosial twitter dengan menggunakan algoritma *Support Vector Machine* (SVM) dalam penelitiannya digunakan dataset twitter mengenai topik pemindahan ibu kota Indonesia, hasil yang diperoleh dengan tweets sebanyak 1.236 mendapatkan akurasi =96,68%, precision =95.82%, dan recall =94.04% [7].

Beberapa penelitian terkait tentang analisis sentimen menggunakan algoritma *Support Vector Machine* (SVM) sebagai berikut: Dalam penelitian ini, data yang digunakan mengacu pada kata kunci "bipolar". Terdapat 2177 data cuitan yang berhasil diperoleh melalui proses pengambilan data menggunakan kunci API akses dari pengembang Twitter. Setelah itu, data tersebut akan mengalami tahap pra-pemrosesan. Perbandingan presentase yang dihasilkan menunjukkan bahwa sekitar 70,92% dari data mengungkapkan pendapat negatif, sementara 29,08% mengungkapkan pendapat positif. Hasil analisis yang dilakukan dalam penelitian ini menggunakan algoritma *Naïve Bayes*, dengan tingkat akurasi mencapai sekitar 92,110092%. [8].

Platform-platform media sosial (seperti Facebook, Instagram, Twitter, dan sejenisnya) telah diadopsi secara luas. Mereka memiliki berbagai keunggulan, terutama dalam konteks bisnis. Meskipun begitu, kadang-kadang media-media sosial tersebut juga bisa berdampak negatif, seperti menurunkan produktivitas karyawan, memicu konflik dalam hubungan, atau bahkan mendorong tindakan kriminal. Karena itulah, penelitian ini mengajukan sebuah metode untuk mengumpulkan data dari salah satu platform media sosial, yaitu Twitter, guna membuat profil atau gambaran. Dalam proses analisis jaringan, aplikasi Gephi diaplikasikan setelah data diambil melalui Twecoll, sebuah aplikasi pengambilan data yang dibangun dengan Python. Selanjutnya, sebuah aplikasi berbasis web juga dirancang dan diimplementasikan, melibatkan penggunaan server berbasis Apache serta skrip yang berbasis Python. Hasil penelitian menunjukkan bahwa akun-akun yang telah berhasil dikumpulkan ini membentuk beberapa kelompok atau komunitas, dengan tingkat koneksi yang berbeda di antara mereka. [9]. Platform-platform media sosial kini bisa dimanfaatkan untuk mengkaji berbagai objek, seperti pandangan politik, produk, layanan, dan sebagainya. Dalam upaya memahami kinerja suatu objek, analisis sentimen telah menjadi alat yang umum digunakan untuk mendapatkan ulasan dari konsumen atau pengguna (baik respons yang positif maupun negatif). [10]. Seiring dengan pertumbuhan jumlah individu yang menggunakan media sosial dalam rutinitas harian mereka, data yang berasal dari media sosial sedang dianalisis dalam berbagai bidang ilmu. Proses analisis media sosial melibatkan empat tahapan yang berbeda, termasuk penemuan data, pengumpulan, persiapan, dan analisis. [11].

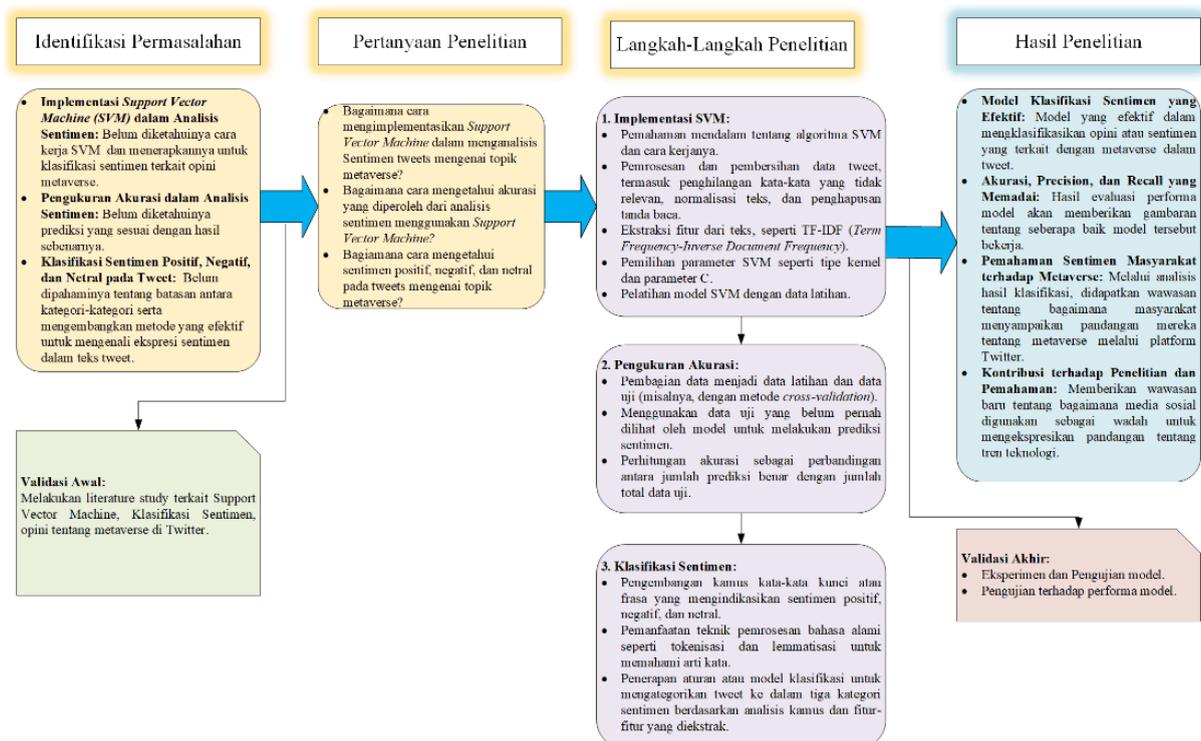
Analisis sentimen diterapkan untuk membedakan apakah komentar tersebut memiliki nuansa positif atau negatif. Dalam penelitian ini, kita memeriksa efektivitas metode *Naïve Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen pada ulasan Summarecon Mal Bekasi. Data yang digunakan terdiri dari 2.143 komentar. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* mencapai akurasi sebesar 80,95%,

sementara Support Vector Machine mencapai akurasi 100%. Untuk mengimplementasikan konsep ini, sebuah prototipe yang berbasis Jason diimplementasikan menggunakan kerangka kerja Flask. [12].

Pandangan terhadap aplikasi KAI Access, baik yang bersifat positif maupun negatif, dapat diartikan sebagai materi penelitian untuk mendapatkan informasi. Hasil penelitian dengan menerapkan klasifikasi *Support Vector Machine* (SVM) menunjukkan tingkat akurasi sekitar 93%, sedangkan dengan metode *Naïve Bayes*, akurasinya sekitar 89%. Oleh karena itu, metode *Support Vector Machine* menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan metode *Naïve Bayes*. [13].

II. Metodologi Penelitian

Berikut kerangka pemikiran dari penelitian ini yang dapat dilihat pada Gambar 1:



Gambar 1. Kerangka Pemikiran

Gambar 1 menjelaskan tentang kerangka pemikiran dimulai dari identifikasi permasalahan, pertanyaan penelitian, langkah-langkah penelitian, dan hasil penelitian.

Pengumpulan data diambil dari twitter dengan teknik *crawling*. Proses *crawling* data ini adalah proses awal untuk pengumpulan data yang dilakukan dengan menggunakan python dan teks editor Google Colab.

Setelah mengumpulkan data, data akan memasuki tahapan *preprocessing*, tahapan *preprocessing* ini merupakan tahapan pembersihan kata pada data atau tweets (*Cleansing*) yang berfungsi untuk membuang kata atau tweets yang tidak diperlukan seperti menghilangkan karakter-karakter selain huruf, menyamakan bentuk kata, dan lain-lain sehingga menghasilkan kata-kata yang lebih singkat.

Tahapan pembobotan ini dilakukan setelah melalui tahap *preprocessing* yang bertujuan untuk menghitung seberapa banyak kata yang muncul pada setiap dokumen. Teknik yang digunakan adalah TF-IDF (*Term Frequency-Inverse Document Frequency*). TIF-IDF gabungan dari *Term Frequency* (TF) dan *Invers Documents Frequency* (IDF). Teknik TF-IDF ini telah terbukti cukup efektif untuk klasifikasi sentimen [14]. Term dapat berupa kata, frase, atau unit pengindeksan lain dalam dokumen yang dapat digunakan untuk menentukan konteks dokumen [13].

Setelah melalui proses *case folding*, tokenisasi, penghapusan kata penghubung, dan stemming, data kemudian diberi bobot. Setelah diberi bobot, data ini diklasifikasikan. Jika nilai kata atau istilah tersebut kurang dari 0, maka kata tersebut akan dianggap sebagai komentar negatif. Jika nilainya adalah 0, maka kata tersebut akan dianggap sebagai komentar netral. Jika nilai kata tersebut lebih dari 0, maka kata tersebut akan dianggap sebagai komentar positif. Konsep dari *Support Vector Machine* (SVM) dapat dijelaskan dengan sederhana sebagai upaya untuk menemukan bidang pemisah terbaik, yang berfungsi sebagai pembatas antara dua kelas di dalam ruang input. Pola yang termasuk dalam dua kelas, yaitu +1 dan -1, memiliki alternatif garis pemisah (batas diskriminasi). Margin merujuk pada jarak antara bidang pemisah tersebut dengan pola terdekat dari kedua kelas. Pola yang paling dekat ini disebut sebagai vektor pendukung (*support vector*). [15].

III. Hasil dan Pembahasan

Penelitian analisis sentimen opini publik mengenai topik *metaverse* pada twitter menggunakan *Support Vector Machine* (SVM) adalah penelitian tentang menganalisis opini atau pendapat masyarakat pada topik *metaverse* di Twitter dari pendapat positif, negatif, maupun netral. Data yang digunakan adalah tweets berbahasa Indonesia, data ditarik dengan menggunakan teknik *crawling* dan media twitter sebagai alat mengambil data. Tweets diambil sebanyak 350 tweets yang kemudian dilakukan proses *preprocessing* dengan menghasilkan data akhir sebanyak 105 tweets. Data tersebut dianalisis dengan menggunakan metode SVM melalui pembagian data menjadi data latih dan data uji secara terstruktur, dengan tujuan untuk mencapai hasil yang optimal.

3.1. Perancangan Dataset

Perancangan ini melibatkan proses pengumpulan data yang diperlukan untuk analisis sentimen dan memasukkannya ke dalam dataset. Hal ini memungkinkan data untuk diproses oleh program dan menghasilkan output yang diinginkan. Pelabelan data dilakukan dengan cara manual atau secara subyektif.

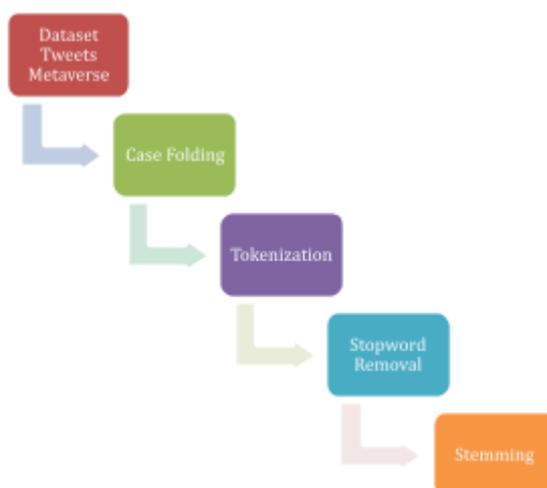
Pengolahan ini melibatkan proses pengumpulan data yang diperlukan untuk analisis sentimen dan memasukkannya ke dalam dataset. Hal ini memungkinkan data untuk diproses oleh program dan menghasilkan output yang diinginkan. Berikut Gambar 2 adalah contoh bagian dari data mentah yang akan dimasukkan ke dalam program. Data mentah yang sudah didapat akan dilakukan proses *cleansing* melalui *preprocessing*.

	ULASAN	LABEL
11	tak mau kalah dari korea selatan jepang bakal ...	Netral
10	ini gak setuju kalian kenapa sih bilangnye ban...	Negatif
27	gue tau kok soal pubj ngatain sama ² metaverse ...	Netral
0	sok sok an buka metaverse gaada yg tertarik cu...	Positif
29	gajelas sih metaverse menurut gua	Negatif
40	lagipula dan dari kapan metaverse muncul	Netral
2	kalo pencet tombol power itu bisa masuk langsu...	Netral
64	tunggu aja nnti zaman nya metaverse	Netral
58	uniknya batik carnival konsep metaverse rayaka...	Netral
23	mengenal nusameta platform metaverse yang akan...	Netral

Gambar 2. Data Mentah *Metaverse*

3.2 Preprocessing

Data awal yang telah diperoleh akan melalui serangkaian langkah pembersihan melalui tahap *preprocessing* yang meliputi proses *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, sehingga data tersebut bisa diolah dalam program. Tahapan *preprocessing* ini dapat dilihat pada Gambar 3.



Gambar 3. Tahapan Preprocessing

- Pada tahap *Case Folding* merupakan tahapan proses yang berfungsi untuk menyamakan bentuk huruf-huruf kapital menjadi huruf kecil dan menghilangkan karakter selain a-z.
- Pada tahap *Tokenization* adalah tahapan untuk memecahkan kalimat menjadi satuan kata. Dan memecahkan kalimat menjadi kata yang mempunyai Batasan tanda baca dan spasi. Berikut adalah hasil dari data yang sudah ditokenizing.
- Pada tahap *Stopword Removal* adalah tahap yang berfungsi untuk menghilangkan kata yang tidak memiliki arti atau makna. Stopword removal melibatkan penghilangan kata-kata dengan konten informasi rendah dari sebuah teks. Contoh stopwords yaitu “dari”, “dan”, “yang”, “di”, dan seterusnya. Berikut adalah hasil stopword removal.
- Pada tahap Stemming adalah Tahap selanjutnya data yang sebelumnya sudah melalui tahap filtering atau stopword removal akan diubah menjadi ke bentuk dasarnya. Stemming adalah tahap yang mentransformasi kata-kata pada suatu kalimat menjadi kata-kata akarnya misalnya kata “menyelesaikan”, “selesaikan”, ‘diselesaikan’ akan menjadi kata “selesai”.

Pembobotan merupakan metode pengambilan keputusan dalam sebuah proses yang melibatkan sejumlah faktor secara bersamaan, dengan memberikan nilai penting pada setiap faktor tersebut. Proses pembobotan bisa dilakukan secara objektif melalui perhitungan statistik, atau secara subyektif melalui penentuan berdasarkan pertimbangan khusus. Dalam konteks penelitian ini, setelah data melewati tahap preprocessing, data tersebut akan diubah menjadi model dan representasi ruang vektor dengan menggunakan pembobotan melalui algoritma TF-IDF. *Term Frequency* (TF) menghitung jumlah kemunculan kata pada sebuah dokumen, *Dokumen Frequency* (DF) menghitung jumlah dokumen yang mengandung term, dan *Inverse Dokumen Frequency* (IDF), jumlah term yang ada pada dokumen.

3.3 Pelabelan Data

Analisis sentimen diawali dengan pelabelan data yang dilakukan secara manual dengan subyektif. Pada tahapan pelabelan akan dilakukan kedalam tiga kelas sentimen yaitu sentimen positif, sentimen negatif, dan sentimen netral. Berikut adalah hasil pelabelan kelas sentimen pada topik *metaverse*.

Tabel 1. Hasil Label Kelas Sentimen

Sentimen	Metaverse
Positif	14
Negatif	11
Netral	80

3.4 Model

Setelah berhasil menghitung pembobotan, selanjutnya data masuk ke tahap klasifikasi SVM. Data akan diklasifikasikan dengan 3 (tiga) kategori yaitu positif, negatif, dan netral dengan menggunakan algoritma *Support Vector Machine* (SVM).

3.5 Uji Coba Validasi

Uji coba validasi merupakan proses evaluasi yang dilakukan dengan membandingkan hasil keluaran dari data uji dalam program dengan hasil perhitungan, bertujuan untuk memverifikasi korelasi antara keduanya. Dalam penelitian ini, 105 cuitan digunakan sebagai data uji, yang dibagi menjadi beberapa bagian atau partisi. Data training terdiri dari 52 data, dan data uji terdiri dari 53 data. Untuk mendapatkan nilai akurasi rata-rata, metode *k-fold cross-validation* digunakan dengan melakukan pengujian sebanyak 5 kali per lipatan. Dalam setiap pengulangan, satu subset digunakan sebagai data uji, sedangkan yang lainnya digunakan sebagai data training.

Hasil uji coba dengan metode *k-fold cross-validation* pada data cuitan terkait *Metaverse* menghasilkan rata-rata akurasi sekitar 0,976 atau 97,6% dalam pembagian antara data latih dan data uji. Akurasi terendah ditemukan pada lipatan kelima dengan nilai sekitar 0,96 atau 96%, sementara akurasi tertinggi terjadi pada lipatan kedua dengan nilai sekitar 0,99 atau 99%.

Matriks Konfusi pada permasalahan klasifikasi dua kelas dan klasifikasi banyak kelas ditunjukkan dalam Gambar 4(a) dan 4(b) masing-masingnya. Tiap kolom dalam matriks menggambarkan contoh-contoh yang telah diprediksi ke dalam suatu kelas, sementara tiap baris menggambarkan contoh-contoh yang sebenarnya termasuk dalam suatu kelas tertentu. Setiap elemen dalam matriks konfusi pada baris *i* dan kolom *j* mengindikasikan jumlah contoh di mana kelas yang diprediksi adalah *j*, namun kenyataannya kelas yang sebenarnya adalah *i*.

Secara kontekstual, matriks konfusi memberikan gambaran tentang bagaimana model klasifikasi "terbingung" saat membuat prediksi. Lebih dari sekadar menunjukkan kesalahan yang terjadi pada pengklasifikasi, matriks ini juga memberikan pemahaman mengenai jenis kesalahan yang dihasilkan. Lihatlah Gambar 4. sebagai contoh matriks konfusi, di mana (a) menggambarkan matriks konfusi untuk permasalahan klasifikasi dua kelas, dan (b) menggambarkan matriks konfusi untuk permasalahan klasifikasi banyak kelas.

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

(a)

		Predicted Class			
		C ₁	C ₂	...	C _N
Actual Class	C ₁	C _{1,1}	FP	...	C _{1,N}
	C ₂	FN	TP	...	FN

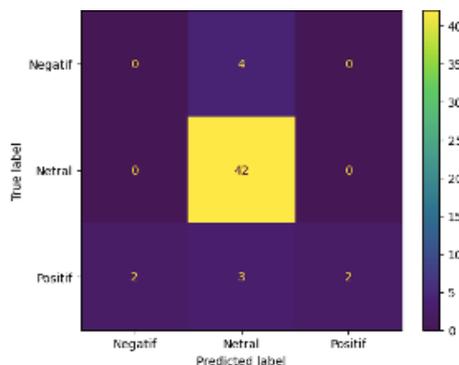
	C _N	C _{N,1}	FP	...	C _{N,N}

(b)

Sumber: [16]–[18]

Gambar 4. Contoh *Confusion matrix*. (a) *Confusion matrix* untuk klasifikasi 2 kelas. (b) *Confusion matrix* untuk *multiclass*

Selanjutnya, dilakukan uji coba untuk mengevaluasi kinerja metode SVM pada setiap dataset dalam melakukan klasifikasi data. Uji coba ini menggunakan *confusion matrix* dengan membagi data secara terstruktur untuk mendapatkan tingkat akurasi yang tinggi. Berikut hasil uji coba *confusion matrix* dapat dilihat pada Gambar 5.



Gambar 5. Confusion Matrix

Berikut adalah hasil matrik konfusi menggunakan *classification reports* yang sudah dihitung oleh program.

	precision	recall	f1-score	support
Negatif	0.00	0.00	0.00	4
Netral	0.86	1.00	0.92	42
Positif	1.00	0.29	0.44	7
accuracy			0.83	53
macro avg	0.62	0.43	0.46	53
weighted avg	0.81	0.83	0.79	53

Gambar 6. Classification Reports

Hasil dari pelabelan data menghasilkan tweet negatif memiliki 4 tweets, tweet netral memiliki 42 tweets, dan tweet positif memiliki 7 tweets. Tweet positif mengenai topik *metaverse* pada twitter dapat divisualisasikan dengan word cloud yang dapat dilihat pada Gambar 7, Gambar 8 dan Gambar 9.



Gambar 7. Word Cloud *Metaverse* Positif

Kata-kata dengan ukuran yang paling besar dalam Gambar 8 melambangkan kata-kata yang paling sering muncul, sementara kata-kata yang lebih kecil melambangkan kata-kata yang jarang muncul. Dalam visualisasi word cloud ini, kata-kata yang paling sering muncul dalam tweet positif tentang topik *metaverse* di Twitter termasuk '*metaverse*', '*keren*', '*bagus*', dan '*suka*'. Dominasi kata-kata ini mencerminkan pandangan positif terhadap konsep *metaverse* yang diungkapkan dalam tweet. Di sisi lain, visualisasi word cloud juga diterapkan pada tweet negatif tentang topik *metaverse* pada Twitter, dan hasilnya bisa dilihat pada Gambar 8.

mengenali dan mengklasifikasikan secara akurat seluruh sentimen yang secara aktual positif, negatif, dan netral dalam kumpulan data. Hal ini menandakan bahwa model memiliki kemampuan yang baik dalam mengenali dan mengklasifikasikan sentimen dengan benar. Hasil penelitian ini menunjukkan bahwa model analisis sentimen yang dikembangkan memiliki kinerja yang baik dalam mengklasifikasikan sentimen pada tweet terkait *metaverse*. *Confusion Matrix* sebagai alat evaluasi telah membantu mengukur performa model dengan akurasi, precision, dan recall yang cukup memadai. Dengan demikian, model ini memiliki potensi untuk memberikan wawasan yang berharga tentang pandangan masyarakat terhadap topik *metaverse* melalui media sosial Twitter.

Daftar Pustaka

- [1] N. Stephenson, "Snow crash Neal Stephenson, London, RoC(Pengiu), 1993, 440 pages," *Futures*, vol. 26, no. 7, pp. 798–800, 1994.
- [2] Z. Allam, A. Sharifi, S. E. Bibri, D. S. Jones, and J. Krogstie, "The Metaverse as a Virtual Form of Smart Cities: Opportunities and Challenges for Environmental, Economic, and Social Sustainability in Urban Futures," *Smart Cities*, vol. 5, no. 3, pp. 771–801, 2022, doi: 10.3390/smartcities5030040.
- [3] A. Bifet and E. Frank, "Sentiment knowledge discovery in Twitter streaming data," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6332 LNAI, pp. 1–15, 2010, doi: 10.1007/978-3-642-16184-1_1.
- [4] Priyanka Takalkar, Prajjawal Neware, Shravya Shetty, Bilal Shaikh, and Renuka Jetthy, "Sentiment Classification for Social Media Posts using Machine Learning," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 2, no. 5, pp. 20–23, 2022, doi: 10.48175/ijarsct-4005.
- [5] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4065–4074, 2013, doi: 10.1016/j.eswa.2013.01.001.
- [6] V. Chandani and R. S. Wahono, "Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film," *Journal of Intelligent Systems*, vol. 1, no. 1, pp. 55–59, 2015.
- [7] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
- [8] O. S. D. Silaen, H. Herlawati, and R. Rasim, "Analisis Sentimen Mengenai Gangguan Bipolar Pada Twitter Menggunakan Algoritma Naïve Bayes," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 6, no. 2, pp. 63–73, 2022, doi: 10.31603/komtika.v6i2.8198.
- [9] H. Herlawati, R. Trias Handayanto, I. Ekawati, K. I. Meutia, J. Asian, and U. Aditiawarman, "Twitter scrapping for profiling education staff," *2020 5th International Conference on Informatics and Computing, ICIC 2020*, no. November, 2020, doi: 10.1109/ICIC50835.2020.9288607.
- [10] Herlawati, R. T. Handayanto, D. Setiyadi, and E. Retnoningsih, "Corpus Usage for Sentiment Analysis of a Hashtag Twitter," *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*, no. May 2021, 2019, doi: 10.1109/ICIC47613.2019.8985772.
- [11] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger, "Social media analytics – Challenges in topic discovery, data collection, and data preparation," *International Journal of Information Management*, vol. 39, no. October 2017, pp. 156–168, 2018, doi: 10.1016/j.ijinfomgt.2017.12.002.
- [12] R. T. Handayanto, H. Herlawati, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, and D. Y. Septia, "Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 2, pp. 153–163, 2021, doi: 10.31603/komtika.v5i2.6280.
- [13] M. Riky Sudrajat, P. D. Atika, and . H., "Implementasi Support Vector Machine (SVM) dan Naïve Bayes untuk Analisis Sentimen Aplikasi KAI Access," *Jurnal ICT : Information Communication & Technology*, vol. 20, no. 2, pp. 254–259, 2021, doi: 10.36054/jict-ikmi.v20i2.403.
- [14] P. E. BLATZ, *the Formation of Long Wavelength Absorbing Species From Short Wavelength Absorbing Linear Conjugated Polyenes*, vol. 15, no. 1. 1972.
- [15] R. Munawarah, O. Soesanto, and M. R. Faisal, "Penerapan Metode Support Vector Machine Pada Diagnosa Hepatitis," *Kumpulan jurnal Ilmu Komputer (KLIK)*, vol. 04, no. 01, pp. 103–113, 2016, doi: 10.20527/klik.v3i1.39.
- [16] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, "Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem," *Technologies*, vol. 9, no. 4, 2021, doi: 10.3390/technologies9040081.
- [17] A. Theissler, M. Thomas, M. Burch, and F. Gerschner, "ConfusionVis: Comparative evaluation and

selection of multi-class classifiers based on confusion matrices,” *Knowledge-Based Systems*, vol. 247, p. 108651, 2022, doi: 10.1016/j.knosys.2022.108651.

[18] M. Grandini, E. Bagli, and G. Visani, “Metrics for Multi-Class Classification: an Overview,” pp. 1–17, 2020, [Online]. Available: <http://arxiv.org/abs/2008.05756>.