Jurnal Pendidikan dan Teknologi Indonesia (JPTI)
Vol. 5, No. 3, Maret 2025, Hal. 845-857
DOI: https://doi.org/10.52436/1.jpti.717
p-ISSN: 2775-4227
e-ISSN: 2775-4219

Analisis Sentimen Petani Milenial Pada Media Sosial X Menggunakan Algortitma *Support Vector Machine (SVM)*

Ma'rufudin[1], Aditia Yudhistira[*2]

[1,2]Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia
Email: [1]ma'rufudin@teknokrat.ac.id, [2]aditiayudhistira@teknokrat.ac.id

Abstrak

Media sosial, termasuk aplikasi X, kini menjadi *platform* utama untuk berbagai diskusi, termasuk topik terkait pertanian milenial. Meski demikian, masih terdapat perbedaan pandangan mengenai penerapan teknologi di sektor pertanian oleh petani milenial. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap petani milenial dengan menggunakan algoritma *Support Vector Machine* (SVM). Sebanyak 2.430 tweet dikumpulkan melalui teknik *crawling* dan diproses melalui tahapan *preprocessing* data, seperti tokenisasi, normalisasi, penghapusan *stopword*, dan *stemming*, serta diberikan bobot menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Model SVM yang dikembangkan dalam penelitian ini mengklasifikasikan sentimen ke dalam tiga kategori: positif, netral, dan negatif. Hasil eksperimen menunjukkan bahwa model SVM mencapai akurasi 70% dengan rata-rata *F1-score* sebesar 0,69. Model ini memiliki *precision* tertinggi sebesar 0,72 untuk sentimen negatif dan recall tertinggi sebesar 0,84 untuk sentimen positif. Dibandingkan dengan algoritma Naïve Bayes yang hanya memperoleh akurasi 65%, SVM terbukti lebih efektif dalam analisis sentimen berbasis teks. Temuan ini mengindikasikan bahwa SVM dapat digunakan untuk mengidentifikasi sentimen publik terhadap petani milenial dengan lebih akurat.

Kata kunci: Analisis Sentimen, Petani Milenial, *Support Vector Machine*, TF-IDF, X Platform

Sentiment Analysis of Millennial Farmers on Social Media X Using the Support Vector Machine (SVM) Algorithm

Abstract

Social media, including application X, has now become a primary platform for various discussions, including topics related to millennial farming. However, there are still differing opinions regarding the implementation of technology in the agricultural sector by millennial farmers. This study aims to analyze user sentiment toward millennial farmers using the Support Vector Machine (SVM) algorithm. A total of 2,430 tweets were collected through crawling techniques and processed through data preprocessing stages such as tokenization, normalization, stopword removal, and stemming, with weighting applied using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The SVM model developed in this study classifies sentiment into three categories: positive, neutral, and negative. Experimental results show that the SVM model achieved an accuracy of 70% with an average F1-score of 0.69. This model had the highest precision of 0.72 for negative sentiment and the highest recall of 0.84 for positive sentiment. Compared to the Naïve Bayes algorithm, which only achieved 65% accuracy, SVM proved to be more effective in text-based sentiment analysis. These findings indicate that SVM can be used to more accurately identify public sentiment toward millennial farmers.

Keywords: *Millennial Farmers, Sentiment Analysis, Support Vector Machine, TF-IDF, X Platform*

1. PENDAHULUAN

Perkembangan teknologi dalam beberapa dekade terakhir, khususnya di bidang Kecerdasan Buatan (AI), telah merombak cara manusia hidup dan berinteraksi dalam berbagai aspek kehidupan, salah satunya di bidang pertanian [1]. *Artificial Intelligence (AI)* adalah istilah dari *Industrial Society* 4.0 dan *Society* 5.0 yang merupakan sebuah "program komputer, pembelajaran mesin, perangkat keras dan perangkat lunak". Ilmu yang digunakan untuk membangun kecerdasan menggunakan solusi perangkat keras dan perangkat lunak yang terinspirasi oleh rekayasa terbalik dari pola neotron yang bekerja di otak manusia [2]. Kecerdasan Buatan (*Artificial Intelligence* atau *AI*) mengacu pada cabang ilmu yang berfokus pada pengembangan sistem komputer dengan karakteristik kecerdasan. Dalam hal ini, AI mencakup perkembangan teknologi dan kemampuan sistem untuk menjalankan

p-ISSN: 2775-4227
e-ISSN: 2775-4219

tugas-tugas yang sebelumnya hanya dapat dilakukan oleh manusia [3]. Peran petani muda milenial saat ini sangat strategis dalam mendukung terwujudnya kedaulatan pangan di Indonesia. Namun, cita-cita tersebut menghadapi tantangan besar, seperti keterbatasan lahan pertanian, rendahnya tingkat produksi, dan kurangnya kemampuan petani dalam memanfaatkan teknologi secara optimal [4]. Bersamaan dengan dampak positif perkembangan AI di sektor pertanian yang dilakukan oleh para petani milenial, tentunya terdapat dampak negatif yang perlu diperhatikan salah satunya ketergantungan berlebih pada teknologi.

Di era digital saat ini, *platform* media sosial seperti *Twitter*, yang kini dikenal sebagai X, telah menjadi *platform* yang populer. Dengan fitur *hashtag*, pengguna X dapat memantau topik terkini secara *real-time* [5]. Pengguna X memiliki pengaruh besar dalam membentuk pandangan masyarakat terkait berbagai isu, termasuk perkembangan AI di sektor pertanian. Diskusi mengenai topik ini sering kali menimbulkan kontroversi antara mereka yang mendukung dan menentang penerapan *Artificial Intelligence (AI)* di sektor pertanian[6]. Sentimen tersebut dapat dimanfaatkan untuk pengumpulan data melalui teknik *web crawling* pada *platform* X guna mengumpulkan *dataset* yang relevan [7].

Opinion mining, atau yang sering disebut analisis sentimen, adalah proses mengidentifikasi dan mengekstraksi opini atau perasaan dari sebuah teks. Tujuannya adalah untuk mengenali persepsi atau pandangan publik terhadap suatu topik, peristiwa, atau isu tertentu. Dalam prosesnya, analisis sentimen akan mengklasifikasikan teks ke dalam kategori positif atau negatif berdasarkan opini yang terkandung di dalamnya [8]. Analisis sentimen masyarakat terhadap kemajuan *Artificial Intelligence (AI)* dalam sektor pertanian menjadi semakin penting untuk memahami pandangan, harapan, dan kekhawatiran yang mungkin muncul terkait penerapan teknologi ini di bidang pertanian. Dengan melakukan analisis sentimen, kita dapat mengidentifikasi tanggapan masyarakat terhadap inovasi dan perubahan yang dibawa oleh perkembangan AI di sektor pertanian oleh para petani milenial. Hal ini tidak hanya membantu para pemangku kepentingan dalam pengambilan keputusan, tetapi juga memberikan wawasan mendalam mengenai aspek-aspek yang perlu diperhatikan dalam implementasi AI di sektor pertanian.

Algoritma *SVM* merupakan salah satu algoritma yang populer dalam analisis sentimen. Dalam *machine learning*, *SVM* digunakan sebagai metode klasifikasi dengan memanfaatkan pola data pelatihan untuk memprediksi kelas [9]. Dikembangkan oleh Vladimir Vapnik, *SVM* memisahkan data ke dalam kelas yang berbeda menggunakan *hyperplane*, yang berfungsi sebagai pembatas untuk menciptakan margin maksimum antara dua kelas data, yaitu positif dan negatif. Namun, *SVM* dengan kernel linear kurang efektif dalam menangani data yang tidak terpisah secara linear atau memiliki hubungan yang kompleks [10]. Oleh karena itu, digunakan teknik *non-linear* untuk mengatasi keterbatasan ini. Biasanya, *SVM* dengan teknik *non-linear* menggunakan berbagai jenis kernel untuk meningkatkan kemampuannya dalam memisahkan data yang sebelumnya tidak dapat dipisahkan secara linear [11].

Dalam konteks klasifikasi analisis sentimen, sudah terdapat beberapa penelitian sebelumnya diantaranya penelitian yang dilakukan oleh [12] dengan judul "Implementasi Algoritma *SVM Non-Linear* Pada Klasifikasi Analisis Sentimen Perkembangan AI di Sektor Pendidikan". Penelitian tersebut sama dengan penelitian ini dengan menggunakan metode *SVM*. Pada penelitian tersebut, dari 3.000 data *tweet* yang dianalisis, sebanyak 52,9% menunjukkan sentimen positif, sedangkan 47,1% bersentimen negatif terhadap perkembangan AI di sektor pendidikan. Hal ini menunjukkan kecenderungan publik yang cukup seimbang namun lebih positif terhadap topik tersebut.

Lalu pada penelitian [13] oleh Inda Sari Tomagola, Asep Id Hidiana, dan Puspita Nurul Sabrina membandingkan metode pembobotan TF-IDF dan TF-RF dalam analisis sentimen menggunakan algoritma Naïve Bayes. *Dataset* diambil dari *Twitter* melalui proses *crawling* dan diolah melalui *preprocessing* (*Case Folding, filtering, Tokenizing*, dan *stemming*). Hasil eksperimen dengan pembagian data 80:20 menunjukkan bahwa TF-IDF + Naïve Bayes menghasilkan akurasi 73%, sementara TF-RF + Naïve Bayes menghasilkan akurasi 72%. Eksperimen dengan pembagian data 70:30 dan 60:40 menunjukkan hasil yang konsisten bahwa TF-IDF lebih unggul dibandingkan TF-RF. *Confusion matrix* digunakan untuk evaluasi metrik seperti *Precision, recall*, dan *F1-score*. Hal tersebut menunjukkan metode TF-IDF lebih efektif dibandingkan TF-RF dalam analisis sentimen berdasarkan hasil akurasi.

Pada tahun 2023, Dedy Atmajaya, Anissa Febrianti, dan Herdianti Darwis [14] melakukan studi tentang analisis sentimen terkait *ChatGPT* di *Twitter* dengan menggunakan algoritma *SVM* dan Naive Bayes. Tujuan dari penelitian ini adalah untuk mengklasifikasikan sentimen pengguna *Twitter* tentang *ChatGPT* ke dalam kategori positif dan negatif. Dalam penelitian tersebut, algoritma *SVM* menunjukkan hasil yang lebih baik dibandingkan algoritma Naive Bayes, dengan tingkat akurasi mencapai 59%.

Berdasarkan konteks masalah dan analisis yang relevan, masyarakat telah berhasil memahami kecenderungan sentimen yang muncul. Dengan semakin pentingnya peran Kecerdasan Buatan *(AI)* dalam sektor pertanian, penelitian ini bertujuan untuk mengimplementasikan model *SVM Non-Linear* untuk menganalisis sentimen petani milenial di *platform* X mengenai perkembangan AI di sektor pertanian, menggunakan *dataset* yang diperoleh

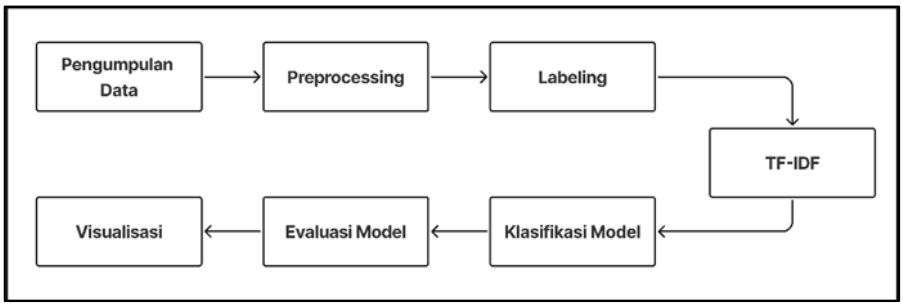
p-ISSN: 2775-4227
e-ISSN: 2775-4219

melalui teknik *crawling* dari *platform* X. Algoritma *SVM Non-Linear* dipilih karena kemampuannya dalam menangani data yang tidak dapat dipisahkan secara *linear*, yang memungkinkan untuk mencapai akurasi yang lebih tinggi. Selain itu, penelitian ini juga bertujuan untuk mengembangkan dan mengevaluasi model *SVM Non-Linear* dengan menggunakan empat jenis kernel untuk menentukan kombinasi terbaik dalam hal akurasi, presisi, *recall*, dan *F1-score*.

Berdasarkan penelitian sebelumnya [12-14], analisis sentimen dalam sektor pertanian telah banyak dilakukan, tetapi masih terbatas dalam membahas respons petani milenial terhadap AI. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi efektivitas model SVM dalam mengklasifikasikan sentimen petani milenial terhadap AI pada platform media sosial X. Dengan memanfaatkan teknik *crawling data*, TF-IDF, serta SVM, penelitian ini diharapkan dapat memberikan wawasan yang lebih akurat mengenai bagaimana petani milenial dipersepsikan di media sosial, serta membantu dalam merancang strategi pengembangan AI yang lebih tepat guna di sektor pertanian.

2. METODOLOGI PENELITIAN

Penelitian ini bertujuan untuk menganalisis sentimen *tweet* petani milenial, dimulai dari pengumpulan data, dilanjutkan dengan *preprocessing*, *labeling*, dan ekstraksi fitur menggunakan TF-IDF. Model kemudian diklasifikasikan menggunakan SVM, dievaluasi, dan hasilnya divisualisasikan untuk mendapatkan wawasan yang lebih jelas. Tahapan penelitian dapat dilihat pada gambar 1.



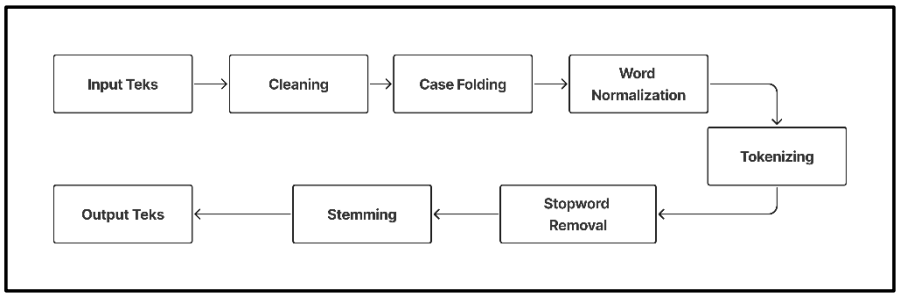
Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Pada penelitian ini, langkah pertama melibatkan pengumpulan data *tweet* melalui proses *crawling* menggunakan pustaka *Tweet Harvest* di Python. *Library* ini dibuat untuk mempermudah ekstraksi data dari media sosial, termasuk X. Sebelum digunakan, diperlukan instalasi Node.js guna mendukung fungsionalitasnya. Penulis menjalankan kode berbasis Python menggunakan Google Colab sebagai platform eksekusi.[15]. Data yang dikumpulkan berasal dari tweet berbahasa Indonesia dalam rentang waktu 1 November 2024 hingga 1 Januari 2025, dengan total 3.805 *tweet.* Data tersebut kemudian disimpan dalam format Excel untuk memudahkan proses analisis lebih lanjut.

2.2. *Preprocessing*

Preprocessing data adalah tahap awal dalam pengolahan data. Pada tahap ini, data yang akan diproses diolah untuk menghindari adanya gangguan dari data yang mengandung *noise* atau data yang tidak konsisten [16]. Tahapan preprocessing dapat dilihat pada gambar 2.



p-ISSN: 2775-4227
e-ISSN: 2775-4219

Gambar 2. Tahapan *preprocessing*

Data yang telah dikumpulkan kemudian melewati beberapa tahapan *preprocessing* untuk memastikan kualitas data yang optimal, yaitu:

a. *Cleaning*: Membersihkan data dari elemen-elemen yang tidak relevan, seperti angka, tanda baca, simbol, atau teks yang tidak bermakna.
b. *Case Folding*: Proses ini merupakan proses menyeragamkan huruf- huruf kapital menjadi huruf kecil. dan pembersihan dokumen dari kata- kata yang tidak dibutuhkan untuk mengurangi *noise*. Kata yang dihilangkan merupakan suatu karakter HTML, kata kunci, *emoticon*, *hastag (#)*, *mention username (@username)*, angka dan url[17].
c. *Word Normalization*: Melakukan normalisasi teks untuk mengubah istilah tidak baku menjadi bentuk baku, misalnya "gpp" menjadi "tidak apa-apa".
d. *Tokenizing*: Memecah teks menjadi unit-unit kecil seperti kata atau frasa untuk mempermudah pengolahan data.
e. *Stopword Removal* dan *Filtering*: Menghapus kata-kata umum yang tidak memiliki makna penting dalam analisis sentimen, seperti "dan", "di", "yang".
f. *Stemming*: Mengubah kata-kata ke bentuk dasar atau akar katanya, misalnya "bermain" menjadi "main".

2.3. *Labeling*

Labeling adalah proses lanjutan setelah *preprocessing* yang bertujuan untuk mempermudah analisis serta pengolahan data pada tahap berikutnya[18]. Proses ini dimulai dengan mengumpulkan data, misalnya melalui *crawling* atau *scraping*, yang kemudian menjalani tahap pembersihan untuk menghilangkan simbol atau elemen yang tidak relevan. Setelah itu, data akan dilabeli untuk menilai apakah sentimen yang terkandung bersifat positif, netral atau negatif. Pendekatan penelitian yang digunakan dalam studi ini merupakan kombinasi antara metode *lexicon based* dan *Support Vector Machine* (SVM). Pada tahap awal, metode *lexicon based* diterapkan untuk melakukan pelabelan data berdasarkan nilai sentimen yang terkandung di dalamnya. Data yang telah diberikan label kemudian digunakan sebagai masukan bagi model SVM, yang berfungsi untuk melakukan analisis lebih lanjut dengan tingkat akurasi yang lebih tinggi. Integrasi kedua metode ini memungkinkan pemrosesan data yang lebih sistematis dan berbasis pada prinsip pembelajaran mesin guna meningkatkan keandalan hasil analisis.[19]

2.4. *Term Frequency – Inverse Document Frequency* (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) adalah teknik yang digunakan untuk menentukan bobot pentingnya kata dalam sebuah dokumen dengan mempertimbangkan dua faktor utama: seberapa sering kata tersebut muncul dalam dokumen tertentu (*Term Frequency*) dan seberapa jarang kata tersebut ditemukan dalam kumpulan dokumen (*Inverse Document Frequency*). Kata yang sering muncul dalam satu dokumen akan memiliki bobot lebih tinggi, sedangkan kata yang muncul di banyak dokumen akan memiliki bobot lebih rendah karena dianggap kurang efektif dalam membedakan satu dokumen dari yang lain[20]. Pemberian bobot ini bertujuan untuk menetapkan nilai frekuensi kata yang nantinya akan dipakai dalam tahap proses selanjutnya.

2.5. Klasifikasi Model dengan SVM

Algoritma *Support Vector Machine* (SVM) adalah metode yang bekerja dengan mencari *hyperplane* atau bidang pemisah optimal untuk memisahkan dua kelas secara maksimal dalam ruang berdimensi tinggi [21]. Dalam *SVM*, support vector mengacu pada data yang berada paling dekat dengan *hyperplane*. Algoritma ini mengoptimalkan *hyperplane* dengan hanya memperhitungkan *support vector*, sehingga meningkatkan efisiensi karena sebagian besar data yang tidak berpengaruh langsung terhadap penentuan *hyperplane* optimal dapat diabaikan. Pendekatan ini memungkinkan *SVM* untuk bekerja lebih efektif.

Dalam proses klasifikasi menggunakan *SVM*, terkadang kernel *linear* tidak mampu memisahkan data dengan optimal, yang dapat menurunkan kinerja model. Hal ini sering terjadi karena teknik linear tidak cukup baik dalam menangani data dengan pola yang kompleks atau tidak teratur. Untuk mengatasi masalah tersebut, penggunaan kernel Non-Linear menjadi alternatif yang efektif. Dengan memanfaatkan kombinasi kernel, *SVM* dapat mentransformasi data ke ruang berdimensi lebih tinggi, sehingga kemampuan pemisahan antar kelas meningkat dan kinerja keseluruhan model menjadi lebih baik.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

2.6. Evaluasi Model

Evaluasi dilakukan dengan menggunakan metode *confusion matrix* untuk menilai performa model dalam mengklasifikasikan sentimen. *Confusion matrix* menyediakan informasi rinci tentang hasil klasifikasi yang dihasilkan model, termasuk jumlah prediksi yang benar serta kesalahan untuk setiap kelas [19]. Pada pengujian, kelas positif mencakup *True Positive* (TP) dan *False Negative* (FN), sedangkan kelas negatif mencakup *True Negative* (TN) dan *False Positive* (FP). Berdasarkan nilai-nilai ini, *confusion matrix* digunakan untuk menghitung metrik evaluasi seperti *accuracy, Precision, recall*, dan *F1-score* [22].

Akurasi dihitung sebagai proporsi prediksi yang benar terhadap seluruh prediksi. *Precision* mengukur seberapa baik model dalam mengidentifikasi data yang benar-benar positif dari semua hasil yang diprediksi positif. *Recall* menunjukkan sejauh mana model mampu mendeteksi semua kasus positif yang sebenarnya. Sementara itu, *F1-score* menghitung rata-rata harmonik antara *Precision* dan *recall*, memberikan evaluasi yang seimbang antara keduanya.

Berikut adalah rumus-rumus perhitungan berdasarkan *confusion matrix*:

a. *Precision* dihitung dengan membagi jumlah *True Positive (TP)* dengan jumlah *True Positive* ditambah *False Positive (FP)*, lalu dikalikan 100%.

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (1)$$

b. *Recall* dihitung dengan membagi jumlah *True Positive (TP)* dengan jumlah *True Positive ditambah False Negative (FN*), lalu dikalikan 100%

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (2)$$

c. *F1-score* dihitung dengan mengalikan dua kali hasil kali *Precision* dan *recall*, lalu dibagi dengan jumlah *Precision* dan *recall*, kemudian dikalikan 100%.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \times 100\% \quad (3)$$

2.7. Visualisasi

Bagian visualisasi dalam penelitian ini menyajikan berbagai grafik dan diagram guna memperjelas hasil analisis sentimen petani milenial di media sosial X dengan menggunakan algoritma *Support Vector Machine* (SVM). *Word cloud* digunakan untuk menampilkan kata-kata yang paling sering muncul dalam percakapan, memberikan gambaran mengenai topik yang dominan. *Confusion matrix* dimanfaatkan untuk menilai kinerja model dalam mengklasifikasikan sentimen, sedangkan grafik akurasi dan loss memperlihatkan performa model selama pelatihan.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Pada penelitian ini, data dikumpulkan menggunakan metode *crawling* dengan *library Tweet Harvest* dan token *API* yang di proses melalui google colab, kemudian disimpan dalam format CSV(*Comma Separated Values*) untuk keperluan analisis lebih lanjut.berikut hasil dari pengumpulan data ditunjukan pada tabel 1.

Tabel 1. Hasil pengumpulan data.

No	*Username*	*Tweet*
1	komunalll	Food estate terus dilanjutkan dengan pelbagai permasalahan. Ini bukan karena anak muda enggan menjadi petani tetapi keterbatasan akses lahan untuk bertani sementara mencetak lahan untuk lumbung pangan terus dilakukan dengan dalih ketahanan pangan .
2	Thesenuttzz	@khanifirsyad emg milenial masih muda ya? yg kelahiran 96 aja udh mau 29 tahun semua. apa program ini buat memberdayakan mereka yg gagal bersaing di lapangan

p-ISSN: 2775-4227
e-ISSN: 2775-4219

		kerja ya harusnya klo mau ngomong anak muda ya genZ dong yg didorong untuk bertani
3.805	SaifulRizal1908	@yusuf_dumdum Lahan pertanian yang seharusnya tumbuh padi malah tumbuh rumah anak muda kurang berminat bertani dan alat pertanian masih jadul walaupun sudah ada beberapa upgrade pupuk mahal dll

Berdasarkan tabel 1 yang menunjukan hasil pengumpulan data yang dilakukan dengan Proses *crawling* dalam rentang waktu 1 November 2024 hingga 1 Januari 2025 dengan menggunakan kata kunci "petani milenial". Dari hasil pengumpulan data, diperoleh 3.805 data *tweet.* Data yang telah dikumpulkan selanjutnya akan melalui tahap pembersihan dan pemrosesan sebelum dilakukan analisis sentimen guna mendapatkan wawasan yang lebih mendalam terkait opini publik mengenai petani milenial.

3.2. *Preprocessing*

Setelah melalui tahap pengumpulan data, diperoleh sebanyak 3.805 data tweet. Selanjutnya, dilakukan Tahap *preprocessing* bertujuan untuk membersihkan dan menyiapkan data sebelum dianalisis lebih lanjut. Cleaning dilakukan untuk menghapus karakter atau simbol yang tidak diperlukan, seperti tanda baca, angka, dan tautan. Case folding mengubah seluruh teks menjadi huruf kecil agar konsistensi data terjaga. Normalisasi kata bertujuan untuk mengganti kata-kata tidak baku atau slang menjadi bentuk standar. Tokenisasi memecah teks menjadi kata-kata atau unit linguistik yang lebih kecil. *Stopword removal* menghilangkan kata-kata umum yang tidak memiliki makna signifikan dalam analisis, seperti "dan," "di," atau "yang." Terakhir, *stemming* mengubah kata ke bentuk dasarnya untuk menyederhanakan analisis teks. Hasil dari tahap *preprocessing* dapat dilihat pada tabel 2.

Tabel 2. Hasil *Preprocessing*

Tahapan	Tweet
Data Tweet	@yusuf_dumdum Lahan pertanian yang seharusnya tumbuh padi malah tumbuh rumah anak muda kurang berminat bertani dan alat pertanian masih jadul walaupun sudah ada beberapa upgrade pupuk mahal dll
Data Cleaning	Lahan pertanian yang seharusnya tumbuh padi malah tumbuh rumah anak muda kurang berminat bertani dan alat pertanian masih jadul walaupun sudah ada beberapa upgrade pupuk mahal dll
Case Folding	lahan pertanian yang seharusnya tumbuh padi malah tumbuh rumah anak muda kurang berminat bertani dan alat pertanian masih jadul walaupun sudah ada beberapa upgrade pupuk mahal dll
Normalization	lahan pertanian yang seharusnya tumbuh padi malah tumbuh rumah anak muda kurang berminat bertani dan alat pertanian masih jadul walaupun sudah ada beberapa upgrade pupuk mahal dll
Tokenizing	['lahan', 'pertanian', 'yang', 'seharusnya', 'tumbuh', 'padi', 'malah', 'tumbuh', 'rumah', 'anak', 'muda', 'kurang', 'berminat', 'bertani', 'dan', 'alat', 'pertanian', 'masih', 'jadul', 'walaupun', 'sudah', 'ada', 'beberapa', 'upgrade', 'pupuk', 'mahal', 'dll']
Stopword	['lahan', 'pertanian', 'tumbuh', 'padi', 'tumbuh', 'rumah', 'anak', 'muda', 'berminat', 'bertani', 'alat', 'pertanian', 'jadul', 'upgrade', 'pupuk', 'mahal', 'dll']
Stemming	lahan tani tumbuh padi tumbuh rumah anak muda minat tan alat tani jadul upgrade pupuk mahal dll

Berdasarkan Tabel 2, tahap *preprocessing* diawali dengan *data cleaning*, yaitu menghapus elemen yang tidak relevan seperti simbol "@" pada tweet asli "@yusuf_dumdum", serta menghilangkan data duplikat dan data *null*, sehingga jumlah data berkurang dari 3.805 *tweet* menjadi 2.430 *tweet.* Selanjutnya, proses *case folding* mengonversi semua huruf menjadi kecil untuk menjaga konsistensi, misalnya kata "Lahan" menjadi "lahan". Pada tahap *tokenizing*, kalimat dipecah menjadi kata-kata terpisah, seperti "Lahan pertanian" menjadi ['lahan', 'pertanian']. Kemudian, *stopword removal* menghapus kata-kata yang kurang bermakna seperti "yang" dan "dan", sehingga hanya menyisakan kata penting seperti ['lahan', 'pertanian']. Selanjutnya, *normalization* mengganti kata tidak baku dengan bentuk standar, misalnya "nyoba" diubah menjadi "mencoba". Terakhir, *stemming* menyederhanakan kata ke bentuk dasar, seperti "berminat" menjadi "minat". Seluruh proses ini bertujuan untuk menyederhanakan dan menstandarkan teks agar siap untuk dianalisis.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

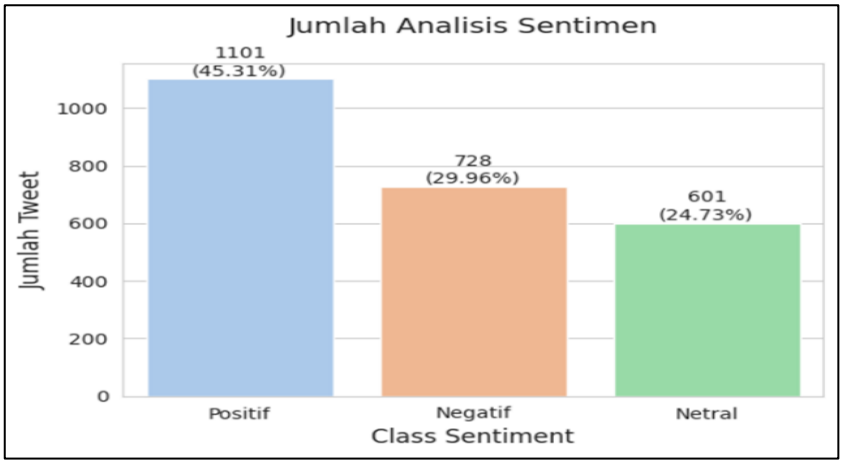
3.3. *Labeling*

Setelah tahap prapemrosesan, data *tweet* sebanyak 2.430 *tweet* diberi label menggunakan pendekatan *lexicon-based* dengan kamus *Indonesian Sentiment Lexicon* (INSET). Setiap kata dalam *tweet* dibandingkan dengan daftar kata positif dan negatif dalam kamus tersebut. Jika kata ditemukan dalam daftar positif, diberi skor +1, dan jika dalam daftar negatif, diberi skor -1. Skor sentimen dihitung dengan mengurangi jumlah kata negatif dari kata positif. *Tweet* dikategorikan sebagai positif jika skor lebih dari 0, negatif jika kurang dari 0, dan netral jika skor sama dengan 0. Hasilnya disajikan dalam Tabel 3.

Tabel 3. Hasil penegumpulan data.

No	*Username*	*Tweet*	*Score*	*Sentiment*
1	komunalll	food estate lanjut pelbagai masalah anak muda enggan tani batas akses lahan tan cetak lahan lumbung pangan dalih tahan pangan	-5	Negatif
2	Thesenuttzz	milenial muda ya lahir program daya gagal saing lapang kerja ya omong anak muda ya genz dorong tani	1	Positif
2.430	SaifulRizal1908	lahan tani tumbuh padi tumbuh rumah anak muda minat tan alat tani jadul upgrade pupuk mahal dll	0	Netral

Tabel 1 menunjukkan hasil pelabelan sentimen pada data tweet. Setiap tweet diberi skor dan kategori sentimen berdasarkan analisis lexicon-based. Contohnya, tweet dari "komunalll" diberi skor -5 dan kategori "Positif", "Thesenuttzz" dengan skor 1 dan kategori "Negatif", serta "SaifulRizal1908" dengan skor 0 dan kategori "Netral". Untuk memberikan pemahaman yang lebih jelas mengenai distribusi hasil pelabelan, perbandingan jumlah data pada setiap kategori ditampilkan dalam sebuah diagram. Diagram tersebut menunjukkan proporsi data untuk masing-masing kategori, seperti yang terlihat pada Gambar 3.



Gambar 3**.** Hasil *labelling*

Berdasarkan Gambar 2, hasil labeling menunjukkan bahwa dari total data, 1101 *tweet* diklasifikasikan sebagai positif (45.31%), 728 *tweet* sebagai negatif (29.96%), dan 601 *tweet* sebagai netral (24.73%). Hal ini mengindikasikan bahwa sebagian besar teks dalam *dataset* memiliki sentimen yang cenderung positif.

3.4. *Term Frequency – Inverse Document Frequency* (TF-IDF)

Setelah proses pelabelan, tahap selanjutnya adalah perhitungan TF-IDF (*Term Frequency-Inverse Document Frequency*). TF-IDF digunakan untuk mengukur seberapa penting sebuah kata dalam sebuah dokumen dibandingkan dengan keseluruhan koleksi dokumen. Hasil TF IDF dapat diluihat pada tabel 4.

Tabel 1. Hasil Tabel Ekstraksi TF-IDF

p-ISSN: 2775-4227
e-ISSN: 2775-4219

(0, 4062)	0.15396219329817856
(0, 6086)	0.3306166035465491
(0, 6251)	0.3374152761790472
(1, 4994)	0.23454317388113718
(1942, 5153)	0.2862728031713548
(1942, 3717)	0.2862728031713548
(1942, 6223)	0.2862728031713548
(1942, 3996)	0.2862728031713548

Tabel 4 menunjukkan hasil ekstraksi fitur menggunakan TF-IDF (*Term Frequency-Inverse Document Frequency*), yang mengukur pentingnya suatu kata dalam dokumen. Setiap baris menunjukkan pasangan indeks dokumen dan indeks kata, bersama dengan nilai TF-IDF yang menunjukkan seberapa relevan kata tersebut dalam dokumen. Nilai yang lebih tinggi menunjukkan bahwa kata tersebut lebih penting dalam dokumen tersebut. Misalnya, pada dokumen ke-0, kata dengan indeks 6251 memiliki nilai TF-IDF 0.337, yang menunjukkan relevansi tinggi kata tersebut dalam dokumen tersebut.

3.5. Klasifikasi Model dengan SVM

Setelah melalui tahap prapemrosesan dan pelabelan, langkah selanjutnya adalah melakukan klasifikasi menggunakan model *Support Vector Machine* (SVM). Pada tahap ini, data dibagi menjadi dua bagian: 80% digunakan untuk data pelatihan dan 20% untuk data pengujian. Kinerja model SVM kemudian diukur dengan menggunakan beberapa metrik, yaitu *Accuracy, Precision, Recall*, dan *F1-Score*. Metrik-metrik ini memberikan gambaran yang lebih komprehensif tentang seberapa baik model dalam mengklasifikasikan sentimen *tweet*, baik dalam hal ketepatan prediksi, kemampuan model untuk mengenali *tweet* positif atau negatif, serta keseimbangan antara *precision* dan *recall*. Hasil klasifikasi model dapat dilihat pada gambar 4.

	precision	recall	f1-score	support
Negatif	0.72	0.69	0.71	142
Netral	0.56	0.45	0.50	126
Positif	0.74	0.84	0.79	218
accuracy			0.70	486
macro avg	0.67	0.66	0.66	486
weighted avg	0.69	0.70	0.69	486

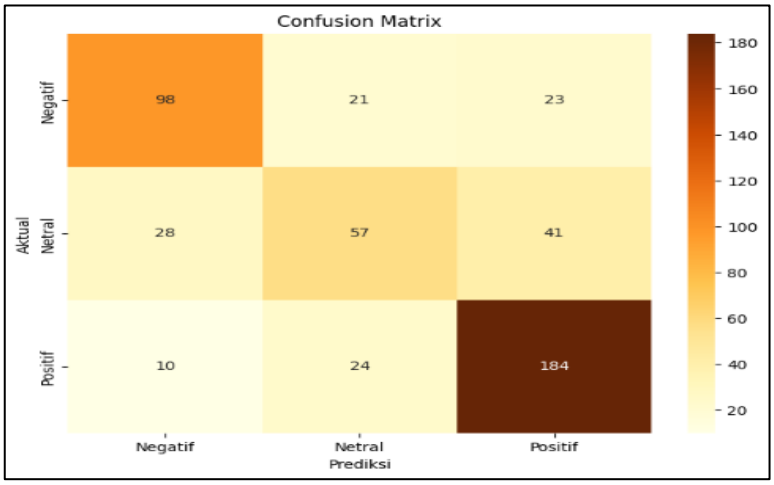
Gambar 4. Hasil *Classification Report*

Berdasarkan gambar 4 yang menunjukkan hasil *Classification Report*, model klasifikasi yang digunakan dalam penelitian ini memiliki tingkat akurasi sebesar 70%, yang menunjukkan bahwa model cukup baik dalam mengklasifikasikan data secara keseluruhan. Dari tiga kelas yang diuji, kelas Positif memiliki performa terbaik dengan precision 0.74, recall 0.84, dan f1-score 0.79, yang menunjukkan bahwa model mampu mengenali sebagian besar sampel positif dengan baik. Sementara itu, kelas Negatif juga memiliki performa yang cukup baik dengan precision 0.72, recall 0.69, dan f1-score 0.71, meskipun masih ada beberapa sampel yang salah diklasifikasikan. Namun, model memiliki kesulitan dalam mengenali kelas Netral, dengan nilai *precision* 0.56, *recall* 0.45, dan f1-score 0.50, yang menunjukkan bahwa model kurang akurat dalam mengidentifikasi data netral, kemungkinan karena karakteristiknya yang lebih ambigu dibandingkan kelas lain. Nilai rata-rata makro (macro avg) untuk precision, recall, dan f1-score masing-masing sebesar 0.67, 0.66, dan 0.66, yang mencerminkan keseimbangan antara kelas meskipun ada perbedaan performa antar kelas. Secara keseluruhan, model bekerja dengan baik tetapi masih dapat ditingkatkan, terutama dalam membedakan data netral dengan lebih akurat.

3.6. Evaluasi Model

Penelitian ini juga mengevaluasi model menggunakan *Confusion Matrix* untuk menganalisis seberapa baik algoritma mengklasifikasikan data dan mengidentifikasi kesalahan yang terjadi. *Confusion Matrix* memberikan gambaran rinci tentang jumlah prediksi yang benar dan salah untuk setiap kelas, membantu memahami pola kesalahan yang sering muncul. Dengan membandingkan hasil *Confusion Matrix*, penelitian ini dapat menilai keakuratan model dalam membedakan kelas negatif, netral, dan positif, serta menentukan algoritma yang paling optimal dengan tingkat kesalahan paling rendah. Hasil evaluasi model dapat dilihat pada gambar 3.

p-ISSN: 2775-4227
e-ISSN: 2775-4219



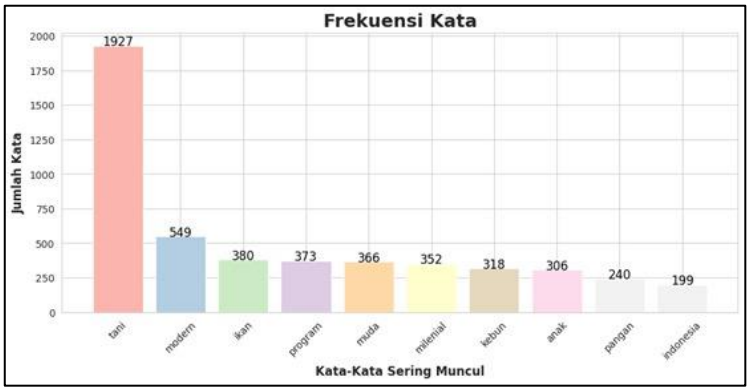
Gambar 5. Hasil *confusion matrix*

Berdasarkan gambar 5 menunjukkan *Confusion Matrix* yang menggambarkan kinerja model klasifikasi sentimen. Model mengklasifikasikan 98 data negatif, 57 netral, dan 184 positif dengan benar. Namun, terdapat beberapa kesalahan prediksi, seperti 21 data negatif diprediksi sebagai netral, 41 data netral diprediksi sebagai positif, dan 10 data positif diprediksi sebagai negatif. Model lebih akurat dalam mengklasifikasikan sentimen positif, tetapi masih sering keliru dalam membedakan netral dan negatif.

Evaluasi model menunjukkan bahwa *Support Vector Machine* (SVM) mencapai akurasi 70% dengan rata-rata *F1-score* 0,69, yang mengindikasikan kinerja cukup baik dalam mengklasifikasikan sentimen petani milenial di media sosial X. *Confusion matrix* mengungkapkan bahwa model dapat mengenali sentimen positif dan negatif dengan cukup baik, tetapi masih mengalami kesulitan dalam membedakan sentimen netral yang sering salah diklasifikasikan. *Precision* dan *recal*l tertinggi ditemukan pada kelas positif dengan nilai masing-masing 0,74 dan 0,84, sementara kelas netral hanya mencapai precision 0,56 dan recall 0,45, menunjukkan bahwa model belum mampu mengenali sentimen netral dengan akurat. Untuk meningkatkan performa, beberapa langkah dapat diterapkan, seperti menambah jumlah data pelatihan, mengeksplorasi model berbasis *deep learning*, atau menggunakan *word embedding* seperti *Word2Vec* atau *BERT* agar model dapat memahami konteks sentimen dengan lebih baik.

3.7. Visualisasi

Setelah melalui tahap klasifikasi dan evaluasi model, selanjutnya adalah tahap visualisasi. Pada tahap ini, beberapa metode digunakan untuk menggambarkan hasil analisis data teks. Antaranya adalah sebagai berikut :

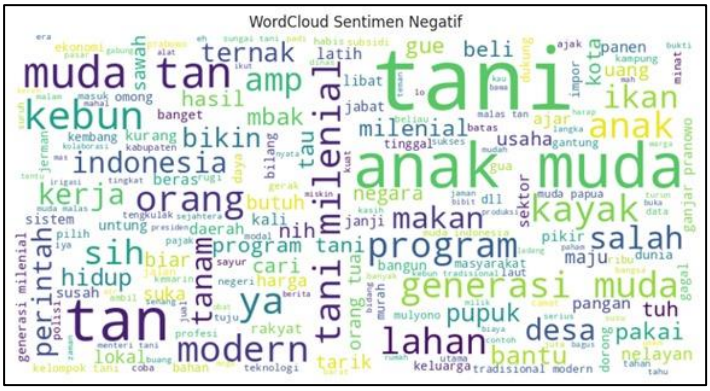


Gambar 6. Frekuensi kata setelah dilakukan *preprocessing data*

Gambar 6 menunjukkan Grafik frekuensi kata menunjukkan bahwa kata tani memiliki kemunculan tertinggi (1927 kali), menandakan bahwa pertanian menjadi topik utama dalam analisis ini. Kata modern (549 kali) dan ikan (380 kali) juga sering muncul, menunjukkan perhatian terhadap modernisasi pertanian dan sektor perikanan. Selain itu, kata program (373 kali) dan muda (366 kali) mengindikasikan adanya keterlibatan generasi muda dalam

p-ISSN: 2775-4227
e-ISSN: 2775-4219

program pertanian. Kata-kata lain seperti milenial, kebun, dan anak menunjukkan fokus pada peran anak muda dan lahan pertanian. Secara keseluruhan, analisis ini mengonfirmasi bahwa pembahasan terkait pertanian, modernisasi, dan keterlibatan generasi muda menjadi isu utama dalam dataset yang dianalisis.



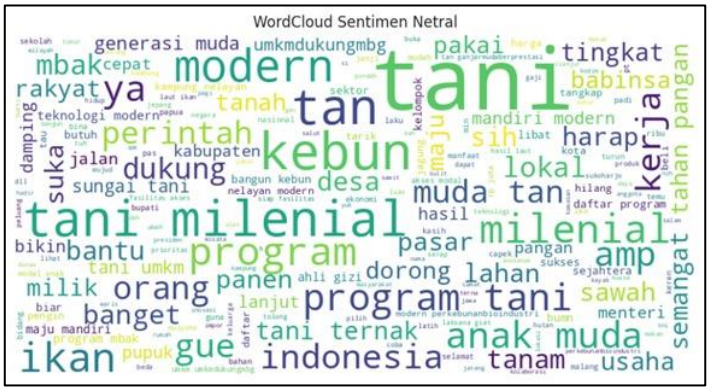
Gambar 7. *Wordcloud* Negatif

Berdasarkan gambar 7 menunjukkan *Word cloud* sentimen negatif terkait pertanian dan program tani milenial. Kata-kata utama seperti tani, muda, anak muda, modern, lahan, salah, kerja, bikin, butuh, tarik, cari, dan ternak menunjukkan adanya berbagai tantangan dan kesulitan dalam sektor ini. Kata "salah" dapat mencerminkan adanya ketidakpuasan atau permasalahan dalam pelaksanaan program, sementara "butuh", "tarik", dan "cari" mengindikasikan adanya kebutuhan yang belum terpenuhi, baik dari segi lahan, tenaga kerja, maupun dukungan lainnya. Secara keseluruhan, *word cloud* ini menggambarkan berbagai kendala yang dihadapi dalam program pertanian dan keterlibatan generasi muda di sektor ini.

Gambar 8. *Wordcloud* Positif

Berdasarkan gambar 7 menunjukkan *Word cloud* sentimen positif terhadap pertanian, dengan fokus pada inovasi, dukungan pemerintah, dan keberhasilan program tani. Kata-kata seperti inovasi, dukungan, dan hasil menunjukkan optimisme terhadap perkembangan sektor ini. Kehadiran kampung nelayan, pasar, dan harga menegaskan dampak positif pada produksi dan distribusi hasil tani, sementara modern dan milenial mencerminkan peran generasi muda dalam pertanian.

p-ISSN: 2775-4227
e-ISSN: 2775-4219



Gambar 9. *Wordcloud* Netral

Berdasarkan gambar 9 menunjukkan *Word cloud* ini mencerminkan opini netral tentang pertanian dan program tani milenial. Kata-kata utama seperti tani, program, milenial, kebun, modern, desa, tanam, ikan, usaha, dan kerja menunjukkan diskusi umum tanpa kecenderungan positif atau negatif. Kehadiran kata program dan modern menandakan pembicaraan seputar perkembangan sektor pertanian, sementara usaha dan kerja mencerminkan aspek aktivitas dan keterlibatan masyarakat dalam bidang ini.

Berdasarkan hasil visualisasi pada gambar 6, 7, 8, dan 9, perbincangan mengenai petani milenial di media sosial X mencerminkan berbagai perspektif yang terbagi dalam tiga kategori sentimen: positif, negatif, dan netral. Gambar 5 menunjukkan bahwa kata "tani" memiliki kemunculan tertinggi, mengindikasikan bahwa diskusi utama berfokus pada pertanian dan petani milenial. *Word cloud* sentimen negatif pada gambar 6 menyoroti isu keterbatasan akses lahan, minimnya dukungan, serta tantangan dalam modernisasi pertanian, sedangkan Gambar 7 yang menampilkan *word cloud* sentimen positif mencerminkan optimisme masyarakat terhadap inovasi pertanian, dukungan pemerintah, serta keberhasilan beberapa program yang telah dijalankan. Adapun gambar 8 dengan word cloud sentimen netral menunjukkan diskusi yang lebih bersifat informatif mengenai program pertanian, peran generasi muda, serta perkembangan teknologi pertanian. Secara keseluruhan, visualisasi ini menggambarkan bahwa meskipun sektor pertanian bagi petani milenial menghadapi berbagai tantangan, tetap terdapat harapan dan dukungan terhadap inovasi serta modernisasi yang terus berkembang.

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *Support Vector Machine* (SVM) untuk menganalisis sentimen pengguna media sosial X terhadap petani milenial dengan menggunakan 2.430 tweet yang dikumpulkan melalui *crawling*. Proses *preprocessing* dilakukan untuk membersihkan data, menormalkan kata, dan melakukan tokenisasi sebelum diubah menjadi representasi numerik menggunakan pembobotan TF-IDF. Hasil pelatihan model SVM menunjukkan akurasi sebesar 70% dengan rata-rata *F1-score* 0,69, yang mengindikasikan bahwa model cukup efektif dalam mengklasifikasikan sentimen positif, netral, dan negatif. Evaluasi menggunakan *confusion matrix* mengungkapkan bahwa model memiliki performa tinggi dalam mendeteksi sentimen positif, tetapi masih kesulitan membedakan sentimen netral dari kategori lain, kemungkinan karena distribusi data sentimen netral yang tidak seimbang. Studi ini berkontribusi dalam pengembangan metode analisis sentimen berbasis SVM yang dapat digunakan untuk mendukung pengambilan keputusan di sektor pertanian modern, meskipun masih memiliki keterbatasan, seperti distribusi data yang kurang merata serta belum mengeksplorasi kernel *Non-Linear* dalam SVM. Oleh karena itu, penelitian di masa depan dapat difokuskan pada peningkatan akurasi model dengan menggunakan *dataset* yang lebih besar, menerapkan metode pembelajaran mendalam, atau mengadopsi pendekatan *ensemble* untuk meningkatkan efektivitas klasifikasi sentimen. Kesimpulannya, SVM menunjukkan efektivitas dalam mengklasifikasikan sentimen petani milenial terhadap AI dengan akurasi 70%, tetapi untuk meningkatkan performa dan mengatasi kesulitan dalam membedakan sentimen netral, eksplorasi lebih lanjut terhadap teknik *deep learning* dan perluasan *dataset* direkomendasikan untuk penelitian selanjutnya.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

DAFTAR PUSTAKA

[1] L. Gkinko and A. Elbanna, "The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users," *Int J Inf Manage*, vol. 69, no. August 2022, p. 102568, 2023, doi: 10.1016/j.ijinfomgt.2022.102568.

[2] U. Darwis and L. Septia Dewi BrGinting, "IMPLEMENTASI TEKNOLOGI ARTIFICIAL INTELEGENCE DALAM BIDANG PENDIDIKAN," *Jurnal Ilmu Pendidikan*, vol. 3, pp. 72–79, Jul. 2023, doi: 10.32696/jip.v5i1.3153.

[3] H. I. Putra Ganda Dewata1, Azzar Rizky2, "Jurnal Rein," *Analisis Sentimen Terhadap Boikot Produk Israel Menggunakan Algoritma Naive Bayes Dan SMOTE*, vol. 1, no. 1, pp. 36–45, 2024.

[4] E. Y. Arvianti, H. Anggrasari, and M. Masyhuri, "Pemanfaatan Teknologi Komunikasi melalui Digital Marketing pada Petani Milenial di Kota Batu, Jawa Timur," *Agriekonomika*, vol. 11, no. 1, pp. 11–18, 2022, doi: 10.21107/agriekonomika.v11i1.10403.

[5] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Smatika Jurnal*, vol. 10, no. 02, pp. 71–76, 2020, doi: 10.32664/smatika.v10i02.455.

[6] T. Wiratama Putra, A. Triayudi, and A. Andrianingsih, "Analisis Sentimen Pembelajaran Daring Menggunakan Metode NaÃ¯ve Bayes, KNN, dan Decision Tree," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 6, no. 1, pp. 20–26, 2022, doi: 10.35870/jtik.v6i1.368.

[7] M. K. Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, pp. 478–483, 2023.

[8] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[9] Q. A. Chairunnisa, Y. Herdiyeni, M. Kusuma, D. Hardhienata, and J. Adisantoso, "Analisis Sentimen Pengguna Twitter Terhadap Program Vaksinasi Covid-19 di Indonesia Menggunakan Algoritme Support Vector Machine Sentiment Analysis of Twitter Users on COVID-19 Vaccination Program in Indonesia using Support Vector Machine Algorithm," *Jurnal Ilmu Komputer dan Agri-Komputer*, vol. 9, no. 1, p. 79, 2022.

[10] M. Azimi-Pour, H. Eskandari-Naddaf, and A. Pakzad, "Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete," *Constr Build Mater*, vol. 230, p. 117021, Jan. 2020, doi: 10.1016/J.CONBUILDMAT.2019.117021.

[11] F. Nie, W. Zhu, and X. Li, "Decision Tree SVM: An extension of linear SVM for non-linear classification," *Neurocomputing*, vol. 401, pp. 153–159, Aug. 2020, doi: 10.1016/J.NEUCOM.2019.10.051.

[12] A. N. Putri, A. Aryanti, and S. Soim, "Implementasi Algoritma SVM Non-Linear Pada Klasifikasi Analisis Sentimen Perkembangan AI di Sektor Pendidikan," vol. 6, no. 2, pp. 703–711, 2024, doi: 10.47065/bits.v6i2.5522.

[13] I. S. Tomagola, A. Id Hidiana, and P. Nurul Sabrina, "Analisis Sentimen Terhadap Pangan Nasional Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 5, pp. 3350–3356, 2024, doi: 10.36040/jati.v7i5.7473.

[14] D. Atmajaya, A. Febrianti, and H. Darwis, "Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *The Indonesian Journal of Computer Science*, vol. 12, no. 4, pp. 2173–2181, 2023, doi: 10.33022/ijcs.v12i4.3341.

[15] Y. Mayasari, R. Nasution, and N. Sumatera Utara, "Post-Election Sentiment Analysis 2024 via Twitter (X) Using the Naïve Bayes Classifier Algorithm," *Journal of Dinda Data Science, Information Technology, and Data Analytics*, vol. 4, no. 2, pp. 123–134, Aug. 2024, doi: 10.20895/dinda.v4i2.1582.

[16] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *Jurnal Ilmiah Informatika*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.

[17] A. A. Pratiwi and M. Kamayani, "Perbandingan Pelabelan Data dalam Analisis Sentimen Kurikulum Proyek di platform TikTok: Pendekatan Naïve Bayes," *Jurnal Eksplora Informatika*, vol. 14, no. 1, pp. 96–107, Sep. 2024, doi: 10.30864/eksplora.v14i1.1093.

[18] R. Hilma, M. Ula, and S. Fachrurrazi, "Analisis Sentimen Cyberbullying pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Naïve Bayes Classifier," *e-jurnal TECHSI*, vol. 14, pp. 107–23, Dec. 2023, doi: 10.29103/techsi.v14i2.12103.

p-ISSN: 2775-4227
e-ISSN: 2775-4219

[19] R. H. Muhammadi, T. G. Laksana, and A. B. Arifa, "Combination of Support Vector Machine and Lexicon-Based Algorithm in Twitter Sentiment Analysis," *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, vol. 8, pp. 59–71, Apr. 2022, doi: 10.23917/khif.v8i1.15213.

[20] E. Harieby and M. Walid, "TWITTER TEXT MINING MENGENAI ISU VAKSINASI COVID-19 MENGGUNAKAN METODE TERM FREQUENCY, INVERSE DOCUMENT FREQUENCY (TF-IDF)," *Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 532–537, Aug. 2022, doi: 10.36040/jati.v6i2.5129.

[21] T. Safitri, Y. Umaidah, and I. Maulana, "Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine," *Journal of Applied Informatics and Computing*, vol. 7, no. 1, pp. 28–35, 2023, doi: 10.30871/jaic.v7i1.5039.

[22] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm," *Jurnal Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi: 10.14710/j.gauss.v9i3.28932.