Journal Software, Hardware and Information Technology
p-ISSN: 2808-3385 | e-ISSN: 2776-8961
Volume 5 Issue 2, June 2025

Penerapan *Machine Learning* untuk Klasifikasi Teks Depresi pada Kesehatan Mental dengan SVM, TF-IDF, dan *Chi-Square*

Application of Machine Learning for Depression Text Classification in Mental Health Using SVM, TF-IDF, and Chi-Square

Muhammad Ridha, Muhammad Kholil Abdur Rohman, Dian Agustin, Deni Handika Shaputra, Yasmine Manayla, Amiroh Hanan Malikah
Program Studi Manajemen Informatika, Politeknik Astra, Jakarta Utara, Indonesia
m.ridha@polytechnic.astra.ac.id, 0320240048@polytechnic.astra.ac.id, 0320240021@polytechnic.astra.ac.id, 0320240018@polytechnic.astra.ac.id, 0320240084@polytechnic.astra.ac.id, 0320240009@polytechnic.astra.ac.id

Notes

Submitted: 20-06-2025
Revised: 26-06-2025
Accepted: 29-06-2025
Published: 30-06-2025

How to cite: M. Ridha, M. K. A. Rohman, D. Agustin, D. H. Shaputra, Y. Manayla, and A. H. Malikah, "Penerapan Machine Learning untuk Klasifikasi Teks Depresi pada Kesehatan Mental dengan SVM, TF-IDF, dan Chi-Square," *Journal Software, Hardware and Information Technology (SHIFT)*, vol. 5, no. 2, pp. 180–193, Jun. 2025, doi: 10.24252/shift.v5i2.210.

DOI: 10.24252/shift.v5i2.210

Copyright 2025© the Author(s)

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

ABSTRACT

Mental health has become a crucial global issue, with increasing numbers of individuals expressing their psychological conditions openly on social media platforms. This study aims to classify tweets related to mental health, specifically depression, using a combination of Support Vector Machine (SVM), Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction, and Chi-Square feature selection techniques. Although this approach has been widely applied in domains such as product and movie reviews, its application in the mental health context remains limited. The main challenge lies in capturing implicit psychological nuances and indirect expressions frequently present in platforms like Twitter, unlike the explicit text in other domains. Moreover, most prior studies have not integrated comprehensive preprocessing stages including lemmatization, stopword removal, and duplicate elimination for mental health data on social media. This research employs a dataset of 26,448 tweets derived from Kaggle and self-crawled data. The best result was achieved using an SVM with an RBF kernel without Chi-Square feature selection, yielding an accuracy of 74.93%. The study demonstrates that a comprehensive preprocessing pipeline can enhance classification performance. However, the model still struggles with sarcastic or ironic contexts. Future research is recommended to adopt deep learning approaches such as BERT or LSTM to capture more complex textual contexts.

Keywords: chi-square, machine learning, SVM, TF-IDF.

ABSTRAK

Kesehatan mental menjadi isu global yang penting, seiring dengan meningkatnya ekspresi individu mengenai kondisi psikologis mereka melalui media social pada aplikasi X. Penelitian ini bertujuan untuk mengklasifikasikan tweet terkait isu kesehatan mental, khususnya depresi, menggunakan kombinasi algoritma *Support Vector Machine* (SVM), metode ekstraksi fitur TF-IDF, dan teknik seleksi fitur *Chi- Square*. Meskipun pendekatan ini telah banyak digunakan di domain seperti ulasan produk dan film, penerapannya dalam konteks kesehatan mental masih terbatas. Tantangan utama terletak pada nuansa psikologis implisit dan ekspresi tidak langsung yang sering muncul di platform seperti Twitter, yang berbeda dari teks eksplisit pada domain lain. Selain itu, sebagian besar penelitian sebelumnya belum mengintegrasikan tahapan prapemrosesan data secara menyeluruh termasuk lemmatization, stopword removal, dan penghapusan duplikasi untuk data *mental health* di media social aplikasi X. Penelitian ini menggunakan 26.448 tweet hasil gabungan dari dataset Kaggle dan data crawling. Hasil terbaik diperoleh pada konfigurasi SVM kernel RBF tanpa seleksi fitur *Chi-Square* dengan akurasi 74,93%. Penelitian ini menunjukkan bahwa rangkaian proses prapemrosesan data yang komprehensif dapat meningkatkan performa klasifikasi. Namun, model masih menghadapi keterbatasan dalam memahami konteks sarkastik atau ironi. Untuk penelitian selanjutnya disarankan menggunakan pendekatan deep learning seperti BERT atau LSTM untuk menangkap konteks yang lebih kompleks.

Kata kunci: chi-square, machine learning, SVM, TF-IDF.

1. PENDAHULUAN

Kesehatan mental telah menjadi perhatian utama dalam konteks global karena berkaitan langsung dengan kualitas hidup dan produktivitas individu. Seiring dengan perkembangan teknologi digital dan meningkatnya penggunaan media sosial, banyak individu mengekspresikan kondisi psikologis mereka secara terbuka melalui platform daring [1]. Ekspresi ini, apabila dianalisis secara sistematis, dapat memberikan wawasan penting untuk memahami persebaran isu-isu kesehatan mental, serta menjadi dasar bagi pengembangan intervensi preventif dan kuratif. Namun, volume data teks yang sangat besar dan tidak terstruktur menjadikan analisis manual tidak efisien dan rentan bias.

Dalam ranah pengolahan data teks, pembelajaran mesin (*machine learning*) telah digunakan secara luas sebagai pendekatan untuk melakukan klasifikasi otomatis terhadap data besar. Klasifikasi teks merupakan proses penting dalam mengekstrak informasi bermakna dari dokumen tidak terstruktur [2]. *Support Vector Machine* (SVM) menjadi salah satu algoritma yang populer digunakan untuk tugas ini karena kemampuannya menangani data berdimensi tinggi

dan memberikan margin keputusan yang optimal. Akan tetapi, keberhasilan model SVM juga sangat bergantung pada representasi fitur dan seleksi fitur yang digunakan [3].

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan salah satu metode representasi teks yang efektif karena mampu menangkap bobot kata secara kontekstual dalam korpus. Pada penelitian [4] terbukti bahwa TF-IDF lebih baik dibandingkan dengan Word2Vec. Namun, metode ini menghasilkan banyak fitur yang bersifat sparse dan redundan. Oleh karena itu, diperlukan teknik seleksi fitur seperti *Chi-Square* [5] untuk memilih fitur yang paling relevan dengan kelas target. Kombinasi antara SVM, TF-IDF, dan *Chi-Square* telah digunakan dalam berbagai penelitian sebelumnya dan menunjukkan hasil yang kompetitif.

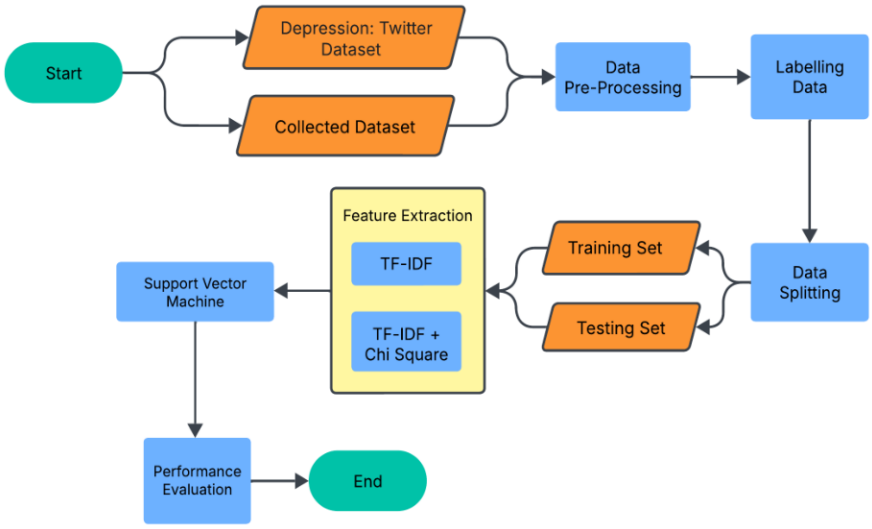
Penelitian [8] menunjukkan bahwa model SVM menghasilkan akurasi 68% dalam analisis klasifikasi tingkat kesehatan mental dengan SVM. Studi lainnya [6], memperlihatkan bahwa SVM dengan feature extraction TF-IDF lebih baik dibandingkan menggunakan Bag of Words. Sementara itu, penelitian lain telah menunjukkan bahwa mengintegrasikan TF-IDF dengan pemilihan fitur *Chi-Square* meningkatkan kinerja model secara signifikan. Misalnya, Irhab dkk. melaporkan peningkatan akurasi SVM sebesar 11,5%—dari 68,7% menjadi 80,2% ketika menganalisis ulasan film [5], sementara penelitian terpisah tentang ulasan Genshin Impact mencapai akurasi 69% menggunakan metode yang sama [7].

Namun demikian, sebagian besar penelitian sebelumnya lebih berfokus pada domain lain seperti ulasan produk [8], ulasan film [9], maupun analisis sentimen umum [10][11], yang konteksnya berbeda dengan isu kesehatan mental. . Hal ini menimbulkan tantangan tersendiri dalam ekstraksi makna dan proses klasifikasi karena model harus mampu memahami nuansa psikologis dan latar belakang pengguna yang tidak eksplisit di dalam teks. Selain itu, hingga saat ini masih terbatas penelitian yang mengintegrasikan teknik prapemrosesan data secara menyeluruh (seperti lemmatization, stopword removal, hingga penghapusan duplikasi) dengan kombinasi SVM, TF-IDF, dan seleksi fitur *Chi-Square* secara sistematis pada data terkait isu kesehatan mental di media sosial pada aplikasi X.

Berangkat dari kondisi tersebut, penelitian ini berfokus pada penerapan dan evaluasi kombinasi metode SVM, TF-IDF, dan *Chi-Square* dalam klasifikasi teks terkait isu kesehatan mental. Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan pendekatan klasifikasi yang seimbang dalam hal akurasi, efisiensi, dan kemudahan implementasi. Dengan demikian, hasilnya dapat dimanfaatkan untuk mendukung sistem analitik yang dapat digunakan oleh praktisi kesehatan, akademisi, maupun pengembang sistem dalam memahami dan merespons isu-isu kesehatan mental secara lebih efektif. Penelitian ini menjadi salah satu dari sedikit studi yang secara sistematis mengevaluasi ketiga pendekatan tersebut dalam konteks mental health, khususnya pada platform media sosial aplikasi X.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode pembelajaran mesin untuk melakukan klasifikasi teks terkait isu kesehatan mental. Tahapan dalam metodologi penelitian ini digambarkan secara sistematis dalam Gambar 1 berikut:



Gambar 1. Metodologi Penelitian

2.1. Dataset

Secara keseluruhan, total data yang digunakan dalam penelitian ini berjumlah 28.492 tweet. Data diperoleh dari dua sumber utama:

- Depression: Twitter Dataset ini diunduh dari situs Kaggle melalui tautan https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media. Dataset ini berisi kumpulan tweet yang telah diklasifikasikan ke dalam kategori "depressed" dan "non-depressed". Dataset berjumlah 20.000 yang sangat relevan karena mencakup beragam ekspresi pengguna Twitter terkait isu kesehatan mental khususnya depresi.
- Collected Dataset via Tweet Harvest dengan mengumpulkan data tambahan secara langsung dari platform Twitter menggunakan tools Tweet Harvester. Proses pengambilan data ini dilakukan dengan memasukkan kata kunci yang berhubungan dengan kesehatan mental, seperti "*depression*", "*anxiety*", "*mental health*", "*stress*", dan "*suicidal thoughts*". Data yang diambil mencakup tweet berbahasa Inggris agar dapat menyesuaikan dengan dataset *Depression*. Setelah data berhasil dikumpulkan, dilakukan proses verifikasi untuk memastikan bahwa tweet yang digunakan benar-benar relevan dengan konteks kesehatan mental. Verifikasi ini diawali dengan penyaringan otomatis berdasarkan konteks penggunaan kata kunci untuk mengeliminasi tweet yang bersifat umum, sarkastik, atau tidak terkait secara langsung dengan kondisi psikologis. Selain itu, duplikasi data juga dihilangkan untuk menjaga keunikan konten. Sebagian sampel dari tweet yang telah difilter kemudian diperiksa secara manual oleh peneliti untuk memastikan bahwa kontennya benar-benar mencerminkan ekspresi, opini, atau pengalaman pribadi mengenai kesehatan mental. Proses verifikasi ini penting untuk menjamin kualitas dan relevansi data yang akan digunakan dalam tahap analisis dan klasifikasi. Jumlah dataset yang berhasil dikumpulkan melalui proses ini adalah sebanyak 8.492 tweet.

2.2. Data Pre-processing

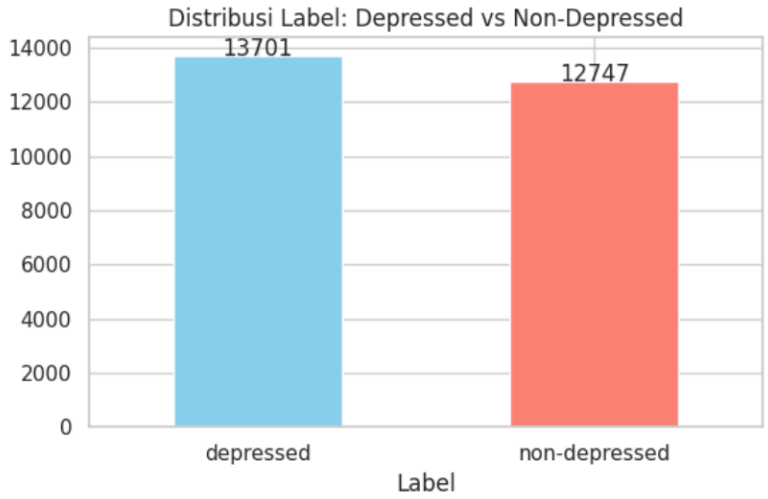
Data mentah yang dikumpulkan selanjutnya melalui tahap prapemrosesan untuk meningkatkan kualitas data. Proses ini meliputi beberapa tahapan utama, yaitu *case folding* untuk mengubah seluruh teks menjadi huruf kecil agar tidak terjadi perbedaan antar huruf kapital dan huruf kecil, *punctuation removal* untuk menghapus semua tanda baca, simbol, dan karakter khusus yang tidak berkontribusi pada pemaknaan teks, serta *number removal* untuk menghapus angka atau digit numerik yang tidak memiliki relevansi dalam proses analisis teks. Selanjutnya dilakukan *lemmatization* untuk mengubah kata ke bentuk dasar (lemma) sehingga berbagai bentuk kata yang memiliki makna serupa dapat disamakan. Selain itu, dilakukan pula *stopword removal* untuk menghapus kata-kata umum seperti "*and*", "*with*", "*to*" yang tidak memiliki makna penting dalam proses klasifikasi. Proses prapemrosesan ini juga mencakup *duplicate removal* untuk menghapus data duplikat atau tweet yang sama agar tidak terjadi redundansi yang dapat memengaruhi hasil analisis. Setelah seluruh tahapan prapemrosesan dilakukan, jumlah data yang semula sebanyak 28.492 tweet berkurang menjadi 26.448 tweet yang siap digunakan untuk tahap ekstraksi fitur selanjutnya. Tabel 1 di bawah ini merupakan contoh data dalam tahap preprocessing.

Tabel 1. Contoh Data Pre-Processing

No	Sebelum	Sesudah
1	I feel so boring now. Like all i am is a mom. I have no hobbies never see friends i feel weird with wearing makeup/dressing up i have nothing to ever fucking talk about. I ve completely lost myself	i feel so boring now like all i am is a mom i have no hobbies never see friends i feel weird with wearing makeupdressing up i have nothing to ever fucking talk about i ve completely lost myself
2	i feel so empty man like if you ask me am I happy or sad I don't even know what to answer you . I'm just existing merely . Like nothing can break or hurt me now or probably nothing really can make me happy now idk what happened but I guess that's a part of making mistakes	i feel so empty man like if you ask me am i happy or sad i dont even know what to answer you im just existing merely like nothing can break or hurt me now or probably nothing really can make me happy now idk what happened but i guess thats a part of making mistakes
3	Organizing the 2022 recap and I'm overwhelmed How did I make so much art!? I thought this was a slow year	organizing the recap and im overwhelmed how did i make so much art i thought this was a slow year

2.3. Labelling Data

Setelah tahap prapemrosesan selesai, data kemudian diberi label sesuai dengan kategori yang telah ditentukan, yaitu "depressed" atau "non-depressed". Proses pelabelan dilakukan menggunakan pendekatan berbasis leksikon (lexicon-based approach). Dalam pendekatan ini, telah disusun daftar kata kunci (lexicon) yang secara spesifik berhubungan dengan kondisi depresi maupun non-depresi. Kata-kata seperti "sad", "desperate", "lost hope", "useless" dimasukkan dalam kategori depresi, sementara kata-kata bernada positif atau netral seperti "happy", "excited", "optimistic" masuk ke kategori non-depresi. Dengan bantuan lexicon ini, tweet yang mengandung kata-kata dari daftar tersebut secara otomatis diberi label yang sesuai. Gambar 2 di bawah ini distribusi label.



Gambar 2. Distribusi Label

2.4. Data Splitting

Setelah proses pelabelan selesai, dataset yang telah diproses dibagi menjadi dua subset, yaitu data latih (training set) dan data uji (testing set) dengan perbandingan 90:10. Sebanyak 90% yaitu 23.803 digunakan untuk melatih model klasifikasi, sementara 10% yaitu 2.645 sisanya digunakan untuk menguji performa model. Pembagian ini dilakukan secara acak (randomized splitting) untuk memastikan bahwa distribusi kelas ("depressed" dan "non-depressed") tetap seimbang pada kedua subset.

2.5. Support Vector Machine

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang dirancang untuk mengklasifikasikan data ke dalam kategori-kategori yang berbeda dengan cara mengidentifikasi hyperplane optimal yang memaksimalkan margin antara kelas-kelas tersebut. Untuk data yang dapat dipisahkan secara linear, SVM mendefinisikan hyperplane menggunakan persamaan 1:

$$w \cdot x + b = 0 \qquad (1)$$

di mana:

- w adalah vektor bobot,
- x adalah vektor fitur input,
- b adalah bias.

Margin ini merepresentasikan jarak terpendek antara titik-titik data dari masing-masing kelas dengan hyperplane, yang dimaksimalkan untuk meningkatkan kemampuan generalisasi model. Ketika data tidak dapat dipisahkan secara linear, SVM menggunakan fungsi kernel seperti *Radial Basis Function* (RBF) untuk memetakan data ke dalam ruang berdimensi lebih tinggi sehingga data menjadi lebih mudah dipisahkan. Algoritma ini juga mengintegrasikan konsep *support vectors*, yaitu titik-titik data yang paling dekat dengan hyperplane dan sangat berperan dalam menentukan posisinya.

Salah satu keunggulan utama dari SVM adalah kemampuannya dalam menangani data berdimensi tinggi secara efisien serta ketahanannya terhadap overfitting, asalkan parameter kernel dan regularisasi yang tepat dipilih. Oleh karena itu, SVM secara luas digunakan dalam

tugas-tugas klasifikasi, termasuk analisis sentimen, di mana dataset sering kali bersifat kompleks dan berdimensi tinggi [3].

Dalam konteks penelitian ini, SVM digunakan untuk mengklasifikasikan teks yang mengandung indikasi depresi atau masalah kesehatan mental. Proses klasifikasi dimulai dengan mengubah teks mentah menjadi vektor fitur menggunakan metode seperti TF-IDF, sehingga setiap dokumen direpresentasikan dalam ruang berdimensi tinggi. SVM kemudian mempelajari pola distribusi dari teks yang termasuk dalam kategori "depresi" dan "tidak depresi", serta menentukan hyperplane yang paling optimal untuk memisahkan kedua kelas tersebut. Titik-titik data yang berada paling dekat dengan hyperplane, yaitu *support vectors*, menjadi kunci dalam membentuk garis pemisah yang akurat. Ketika model dilatih dengan data yang telah direpresentasikan secara numerik, SVM dapat mengklasifikasikan teks baru berdasarkan sisi hyperplane tempat vektor tersebut berada. Dengan pendekatan ini, SVM mampu mendeteksi tanda-tanda linguistik yang mengindikasikan kondisi mental tertentu, sehingga efektif digunakan dalam sistem deteksi dini terhadap depresi melalui analisis teks.

Untuk meningkatkan kinerja model dan mengurangi kompleksitas data berdimensi tinggi, penelitian ini juga menerapkan metode seleksi fitur Chi-Square. Metode ini digunakan untuk mengidentifikasi fitur-fitur (kata-kata) yang secara statistik memiliki hubungan signifikan dengan label kelas. Fitur yang tidak relevan atau tidak memiliki kekuatan diskriminatif terhadap kelas "depresi" dan "tidak depresi" dieliminasi, sehingga hanya fitur-fitur penting yang dipertahankan. Dengan demikian, noise dalam data dapat diminimalkan dan model SVM dapat bekerja lebih efisien dan akurat. Kombinasi antara TF-IDF sebagai metode representasi fitur, Chi-Square sebagai teknik seleksi fitur, dan SVM sebagai algoritma klasifikasi menghasilkan pendekatan yang kuat untuk mendeteksi indikasi depresi dari teks secara otomatis.

2.6. TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) adalah algoritma yang digunakan untuk mengevaluasi pentingnya sebuah kata dalam suatu dokumen relatif terhadap seluruh korpus teks. TF-IDF menggabungkan dua komponen utama, yaitu *Term Frequency* (TF) yang mengukur seberapa sering suatu kata muncul dalam sebuah dokumen, dan Inverse *Document Frequency* (IDF) yang menilai seberapa jarang kata tersebut muncul di seluruh dokumen dalam korpus. Kombinasi kedua komponen ini memberikan bobot lebih tinggi pada kata-kata yang sering muncul dalam suatu dokumen tertentu tetapi jarang muncul di dokumen lainnya, sehingga menyoroti relevansi kontekstual kata tersebut. Rumus TF-IDF dapat dilihat pada persamaan 2.

Dalam konteks penelitian ini, TF-IDF digunakan untuk merepresentasikan teks-teks yang berkaitan dengan depresi dan isu kesehatan mental dalam bentuk vektor numerik, agar dapat diproses oleh algoritma klasifikasi seperti *Support Vector Machine* (SVM). Melalui pendekatan ini, kata-kata yang memiliki keterkaitan kuat dengan gejala atau ekspresi depresi seperti *empty*, *tired*, *worthless*, atau *can't sleep* akan diberikan bobot yang lebih tinggi jika kata tersebut sering muncul dalam teks tertentu namun jarang muncul di dokumen lain. Dengan demikian, TF-IDF membantu dalam menyoroti kata-kata yang relevan secara emosional atau psikologis, yang sangat penting untuk mendeteksi indikasi depresi dari data teks. Teknik ini memungkinkan model untuk fokus pada fitur-fitur yang informatif dan mendukung peningkatan akurasi dalam mengidentifikasi status mental seseorang berdasarkan tulisan mereka.

$$TF - IDF(t,d) = TF(t,d).IDF(t) \qquad (2)$$

Dengan:

Term Frequency (TF): Mengukur seberapa sering term t muncul dalam dokumen d dengan persamaan 3:

$$TF(t,d) = \frac{f_{t,d}}{\sum_k f_{k,d}} \quad (3)$$

di mana:

- $f_{t,d}$ = jumlah kemunculan *term* t dalam dokumen d,
- penyebut adalah total semua frekuensi *term* dalam dokumen d.

Inverse Document Frequency (IDF): Mengukur seberapa jarang term t muncul di seluruh dokumen dalam korpus dengan persamaan 4:

$$IDF(t) = \log\left(\frac{N}{1+ n_t}\right) \quad (4)$$

di mana:

- N = jumlah total dokumen dalam korpus,
- n_t = jumlah dokumen yang mengandung *term* t,
- logaritma biasanya menggunakan basis 10 atau natural (ln).

TF-IDF banyak digunakan dalam analisis teks untuk proses ekstraksi fitur, klasterisasi dokumen, dan pengambilan informasi karena kemampuannya dalam mengidentifikasi kata kunci yang signifikan. Dalam praktiknya, TF-IDF mengubah data teks yang tidak terstruktur menjadi representasi numerik yang dapat diproses oleh algoritma pembelajaran mesin [12].

2.7. Chi-Square

Chi-Square adalah metode statistik yang digunakan untuk mengukur hubungan antara fitur (kata) dan kategori (kelas) dalam sebuah dataset teks. Metode ini menghitung nilai *Chi-Square* [13] untuk setiap fitur, yang merepresentasikan signifikansi statistik dari fitur tersebut terhadap kelas target. Semakin tinggi nilai *Chi-Square* suatu fitur, maka fitur tersebut dianggap semakin signifikan untuk proses klasifikasi. Persamaan 5 merupakan formula untuk *Chi-Square*.

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

Dengan:

- O_i: frekuensi observasi pada kategori tertentu,
- E_i: frekuensi harapan (ekspektasi) jika tidak ada hubungan antara kata dengan kategori.

Metode *Chi-Square* memprioritaskan fitur-fitur dengan nilai yang lebih tinggi karena fitur-fitur tersebut dianggap memiliki kontribusi paling signifikan dalam membedakan antara kelas-kelas yang ada. Dalam praktiknya, *Chi-Square* sangat efektif dalam mengurangi dimensi data dengan cara mengeliminasi fitur-fitur yang tidak relevan atau kurang signifikan, sehingga model dapat berfokus pada fitur-fitur yang paling informatif.

Dalam konteks penggunaan *Support Vector Machine* (SVM), seleksi fitur menggunakan *Chi-Square* dapat meningkatkan akurasi model dengan mengurangi noise dan menghapus fitur yang kurang relevan. Dengan demikian, model menjadi lebih efisien terutama ketika menghadapi data berdimensi tinggi. Berdasarkan studi oleh Afriyani et al., integrasi antara *Chi-Square* dengan algoritma seperti SVM terbukti meningkatkan akurasi klasifikasi secara signifikan. Sebagai contoh, kombinasi ini berhasil mencapai akurasi hingga 92% dalam tugas klasifikasi data teks opini politik di Twitter [14]. Proses pengurangan noise dan penghapusan fitur yang kurang

relevan dilakukan dengan menghitung nilai Chi-Square untuk setiap fitur (seperti kata atau token dalam data teks) terhadap label kelas. Fitur-fitur yang menunjukkan korelasi statistik rendah dengan kelas target akan dieliminasi karena dianggap tidak memberikan kontribusi berarti terhadap proses klasifikasi. Dengan demikian, fitur-fitur yang bersifat umum atau muncul secara acak di seluruh kelas—yang berpotensi menjadi noise—dapat disaring. Hasilnya, model SVM hanya akan dilatih pada fitur-fitur yang paling informatif dan relevan, sehingga dapat menghasilkan margin pemisah yang lebih akurat dan meningkatkan kemampuan generalisasi model terhadap data baru.

2.8. Evaluasi Model

Pada penelitian ini, evaluasi dilakukan dengan menggunakan *Confusion Matrix*, yang memberikan gambaran performa model berdasarkan jumlah prediksi yang benar dan salah terhadap kelas target. Metrik kinerja yang akan diukur seperti persamaan 6 - 9 yaitu *Accuracy* pada persamaan 6, *Precision*, *Recall*, dan *F1-Score* [15].

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

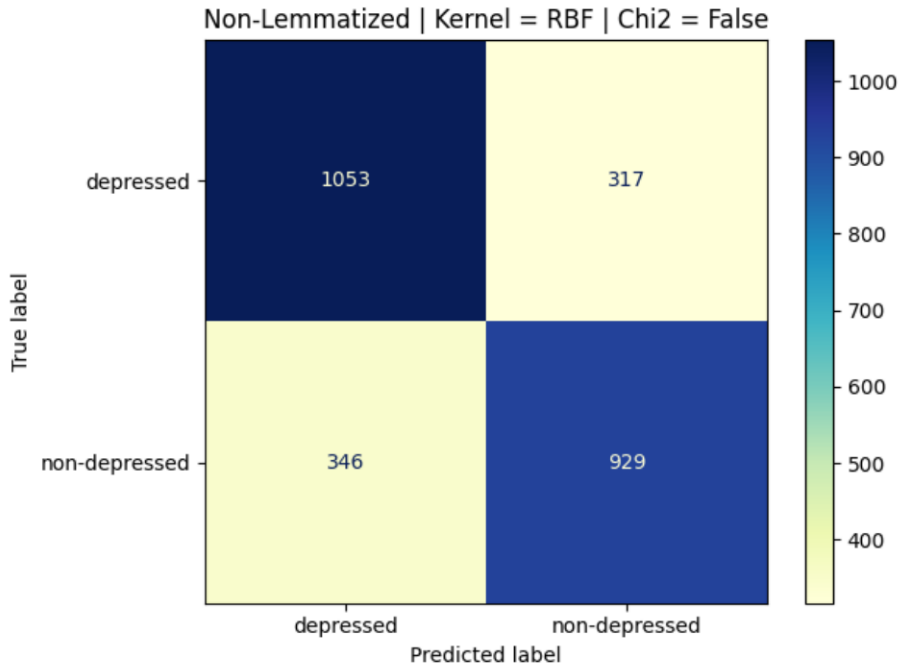
$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1 - Score = 2\,x\,\frac{Precision\ x\ Recall}{Precision+Recall} \quad (9)$$

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1. Hasil Evaluasi Model

Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) dengan tiga jenis kernel, yaitu Linear, *Radial Basis Function* (RBF), dan *Polynomial* (Poly). Evaluasi dilakukan pada dua kondisi utama: data tanpa lemmatization dan data dengan lemmatization, serta menggunakan atau tanpa menggunakan metode feature selection *Chi-Square* (Chi2). Pada Gambar 3 adalah hasil dari Confusion Matrix SVM dengan Non-Lemmatized | Kernel = RBF | Chi2 = False. Berikut adalah persamaan *accuracy*-nya:



Gambar 3. Confusion Matrix SVM dengan Non-Lemmatized | Kernel = RBF | Chi2 = False

$$Accuracy = \frac{1053 + 929}{1053 + 929 + 346 + 317}$$

$$Accuracy = \frac{1982}{2645}$$

$$Accuracy = 0.7493$$

Berikut adalah ringkasan hasil akurasi seluruh model yang dapat dilihat pada Tabel 2 di bawah ini:

Tabel 2. Ringkasan Hasil Akurasi Model

Data	Kernel	Chi2	Akurasi
Non-Lemmatized	Linear	False	74.10%
Non-Lemmatized	Linear	True	71.61%
Non-Lemmatized	RBF	False	**74.93%**
Non-Lemmatized	RBF	True	72.36%
Non-Lemmatized	Poly	False	73.91%
Non-Lemmatized	Poly	True	70.36%
Lemmatized	Linear	False	73.91%
Lemmatized	Linear	True	71.42%
Lemmatized	RBF	False	74.71%
Lemmatized	RBF	True	72.17%
Lemmatized	Poly	False	73.88%
Lemmatized	Poly	True	71.12%

3.2. Analisis Hasil

Pengaruh Lemmatization

Dari hasil di atas, terlihat bahwa proses lemmatization tidak memberikan peningkatan signifikan terhadap akurasi model. Bahkan, akurasi model non-lemmatized sedikit lebih tinggi dibanding lemmatized, terutama pada kernel RBF (74.93% vs 74.71%). Hal ini menunjukkan bahwa bentuk kata asli dalam tweet sudah cukup representatif untuk proses klasifikasi, mengingat bahasa Twitter cenderung ringkas dan informal.

Pengaruh Kernel SVM

Jenis kernel yang digunakan memberikan pengaruh besar terhadap performa model:

- Kernel RBF menghasilkan akurasi tertinggi, yaitu 74.93% pada data non-lemmatized tanpa Chi2.
- Kernel Poly dan Linear memiliki akurasi yang lebih rendah, meskipun masih berada pada kisaran yang cukup baik (sekitar 70%-74%).

Penggunaan kernel RBF memberikan hasil lebih baik karena mampu memetakan data ke dimensi lebih tinggi sehingga dapat memisahkan data yang tidak linear.

Pengaruh Chi-Square (Chi2) Feature Selection

Penggunaan metode feature selection *Chi-Square* justru menurunkan akurasi model pada semua kernel. Sebagai contoh:

- Kernel RBF Non-Lemmatized: dari 74.93% turun menjadi 72.36%.
- Kernel Poly Lemmatized: dari 73.88% turun menjadi 71.12%.

Hal ini menunjukkan bahwa proses feature selection dengan Chi2 kemungkinan besar mengeliminasi fitur-fitur penting yang seharusnya dipertahankan untuk meningkatkan performa klasifikasi. Dengan demikian, penggunaan Chi2 dalam kasus ini tidak disarankan.

Evaluasi Berdasarkan Confusion Matrix

Evaluasi model juga didasarkan pada matriks kebingungan (confusion matrix) untuk melihat detail prediksi benar (*True Positive/True Negative*) dan salah (*False Positive/False Negative*) dari masing-masing kelas. Persamaan *Precision, Recall*, *F1-Score* dari Non-Lemmatized | Kernel = RBF | Chi2 = False yang berdasarkan confusion matrix dapat dilihat di bawah ini.

$$Precision_{depressed} = \frac{1053}{1053+346} = 0.7527$$

$$Precision_{non-depressed} = \frac{929}{929+317} = 0.7456$$

$$Recall_{depressed} = \frac{1053}{1053+317} = 0.7686$$

$$Recall_{non-depressed} = \frac{929}{929+346} = 0.7286$$

$$F1-Score_{depressed} = 2\ x \frac{0.7527\ x\ 0.7686}{0.7527+0.7686} = 0.7606$$

$$F1-Score_{non-depressed} = 2\ x \frac{0.7456\ x\ 0.7286}{0.7456+0.7286} = 0.7372$$

Tabel 3 merupakan hasil *precision*, *recall* dan *f1-score* dari Non-Lemmatized | Kernel = RBF | Chi2 = False.

Tabel 3. *Precision, Recall*, *F1-Score* dari SVM yang Non-Lemmatized | Kernel = RBF | Chi2 = False

Kelas	*Precision*	*Recall*	*F1-Score*
Depressed	0.75	0.77	0.76
Non-Depressed	0.75	0.73	0.74

Nilai precision dan recall kelas "Depressed" sedikit lebih tinggi, artinya model lebih sensitif dalam mendeteksi tweet bertema depresi. Sebaliknya, kelas "Non-Depressed" cenderung sedikit lebih sering salah klasifikasi.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma Support Vector Machine (SVM) dalam klasifikasi tweet terkait kesehatan mental, khususnya untuk mendeteksi indikasi depresi, memberikan performa yang cukup baik. Penelitian ini menggunakan dua sumber data utama, yaitu *Depression: Twitter Dataset* dari Kaggle sebanyak 20.000 tweet serta data tambahan sebanyak 8.492 tweet yang dikumpulkan secara langsung dari Twitter melalui tools Tweet Harvester dengan kata kunci seperti *"depression"*, *"anxiety"*, *"mental health"*, *"stress"*, dan *"suicidal thoughts"*. Setelah melalui proses prapemrosesan yang meliputi *case folding*, *punctuation removal*, *number removal*, *stopword removal*, *lemmatization*, dan penghapusan duplikasi data, total data yang digunakan dalam proses pelatihan dan pengujian model berjumlah 26.448 tweet.

Dalam tahap pemodelan, dilakukan berbagai konfigurasi model SVM dengan menggunakan kernel linier, RBF, dan polinomial, baik dengan maupun tanpa seleksi fitur menggunakan Chi-Square. Berdasarkan hasil evaluasi menggunakan confusion matrix, model terbaik diperoleh dari konfigurasi model SVM tanpa lemmatization, kernel RBF, dan tanpa seleksi fitur Chi-Square, dengan akurasi sebesar 74,93%. Pada konfigurasi model terbaik ini, precision untuk kelas *depressed* adalah 0,75, recall sebesar 0,77, dan f1-score sebesar 0,76, sedangkan untuk kelas *non-depressed*, precision sebesar 0,75, recall sebesar 0,73, dan f1-score sebesar 0,74. Hal ini menunjukkan bahwa model mampu membedakan tweet terkait indikasi depresi dengan tingkat ketelitian dan keseimbangan yang baik antar kelas.

Namun demikian, penelitian ini masih memiliki beberapa kelemahan. Pertama, meskipun data telah dibersihkan dan diproses, adanya noise dalam data seperti sarkasme, ironi, atau ambiguasi makna dalam teks Twitter masih berpotensi memengaruhi hasil klasifikasi. Selain itu, pendekatan pemodelan ini hanya menggunakan fitur tekstual sederhana dan belum melibatkan fitur tambahan seperti *user profile*, fitur berbasis waktu, atau fitur jaringan sosial (*network features*) yang mungkin dapat meningkatkan akurasi model. Penggunaan SVM linier juga mungkin kurang efektif dalam menangkap hubungan non-linier yang kompleks dalam data teks jika dibandingkan dengan model berbasis deep learning seperti LSTM atau BERT.

Sebagai saran untuk penelitian selanjutnya, disarankan untuk mencoba pendekatan berbasis *deep learning* yang mampu menangkap konteks kalimat secara lebih dalam dan memperhitungkan urutan kata seperti model LSTM atau *transformer-based models* (misalnya BERT) [16]. Selain itu, eksplorasi terhadap teknik *data augmentation* dapat digunakan untuk menyeimbangkan jumlah data antar kelas agar model tidak bias. Penambahan fitur non-teks seperti waktu posting, interaksi pengguna, atau metadata tweet lainnya juga diharapkan dapat

meningkatkan performa klasifikasi. Terakhir, perlu dilakukan validasi data secara manual agar dapat mengurangi kesalahan label yang mungkin masih terdapat pada dataset hasil crawling. Secara keseluruhan, penelitian ini berhasil menunjukkan bahwa SVM dapat digunakan sebagai metode klasifikasi awal yang cukup efektif untuk mendeteksi tweet terkait depresi, meskipun masih terdapat ruang untuk perbaikan di masa yang akan datang.

PERNYATAAN PENGGUNAAN ARTIFICIAL INTELLIGENCE

Penulis menggunakan ChatGPT dan DeepSeek untuk mendukung penyusunan naskah ilmiah serta membantu pengembangan dan pemeriksaan kode pemrograman dalam penelitian. Keluaran dari kedua alat tersebut telah dievaluasi dan diverifikasi sebelum diterapkan. Penulis bertanggung jawab atas validitas kode, interpretasi hasil, dan keseluruhan isi artikel.

KONTRIBUSI PENULIS

Muhammad Ridha: Konsep, Metode, Analisis Data, Review.

Muhammad Kholil Abdur Rohman: Pemrograman, Metode, Penulisan Artikel, Review.

Dian Agustin: Pengumpulan Data, Penulisan Artikel, Editing.

Deni Handika Shaputra: Penulisan Artikel, Visualisasi, Review.

Yasmine Manayla: Penulisan Artikel, Analisa Data.

Amiroh Hanan Malikah: Konsep, Metode, Pengumpulan Data, Editing.

PENDANAAN

Penelitian ini tidak menerima hibah khusus dari lembaga pendanaan mana pun, baik sektor publik, komersial, maupun nirlaba.

KONFLIK KEPENTINGAN

Penulis menyatakan tidak terdapat konflik kepentingan terkait publikasi artikel ini.

DAFTAR PUSTAKA

[1] S. K. Lipson et al., “Trends in college student mental health and help-seeking by race/ethnicity: Findings from the national healthy minds study, 2013–2021,” J Affect Disord, vol. 306, pp. 138–147, Jun. 2022, doi: 10.1016/j.jad.2022.03.038.

[2] M. A. Mansoor and K. H. Ansari, “Early Detection of Mental Health Crises through Artifical-Intelligence-Powered Social Media Analysis: A Prospective Observational Study,” J Pers Med, vol. 14, no. 9, p. 958, Sep. 2024, doi: 10.3390/jpm14090958.

[3] E. Hokijuliandy, H. Napitupulu, and Firdaniza, “Application of SVM and Chi-Square Feature Selection for Sentiment Analysis of Indonesia’s National Health Insurance Mobile Application,” Mathematics, vol. 11, no. 17, Sep. 2023, doi: 10.3390/math11173765.

[4] Ahmad Hilman Dani, Eva Yulia Puspaningrum, and Retno Mumpuni, “Studi Performa TF-IDF dan Word2Vec Pada Analisis Sentimen Cyberbullying”, *Router*, vol. 1, no. 4, pp. 20–32, Dec. 2023.

[5] U. Ikhsani Larasati, M. Aziz Muslim, and R. Arifudin, “Improve the Accuracy of Support Vector Machine Using Chi Square Statistic and Term Frequency Inverse Document Frequency on Movie Review Sentiment Analysis,” Scientific Journal of Informatics, vol. 6, no. 1, pp. 2407–7658, 2019, [Online]. Available: http://journal.unnes.ac.id/nju/index.php/sji

[6] K. Tri Putra, M. Amin Hariyadi, and C. Crysdian, “Perbandingan Feature Extraction TF-IDF Dan BoW Untuk Analisis Sentimen Berbasis SVM,” J. Cahaya MAndalika. 2023 Nov;1449.

[7] P. Subarkah, P. Arsi, D. I. S. Saputra, A. Aminuddin, Berlilana, and N. Hermanto, “Indonesian Police in the Twitterverse: A Sentiment Analysis Perspectives,” in 2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), IEEE, Nov. 2023, pp. 76–81. doi: 10.1109/ICITISEE58992.2023.10405357.

[8] W. Andriyani, Y. Astuti, B. A. Wisesa, and D. Hengki, "Analisis Sentimen pada Ulasan Produk dengan SVM dan Word2Vec." *JIKO (Jurnal Informatika dan Komputer)* 9, no. 1, pp 173-185, 2025

[9] I. Sari and K. Wardhana, "Analisa Internet Movie Database (IMDb) Menggunakan Algoritma Machine Learning Super Vector Machine." *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)* 9, no. 3, pp 382-390. 2025.

[10] F. Putrawansyah, “Penerapan Metode Support Vector Machine Terhadap Klasifikasi Jenis Jambu Biji,” JIKO (Jurnal Informatika dan Komputer), vol. 8, no. 1, p. 193, Feb. 2024, doi: 10.26798/jiko.v8i1.988.

[11] Arsyah, Ulya Ilhami, Mutiana Pratiwi, and Abulwafa Muhammad, “Twitter Sentiment Analysis of Public Space Opinions using SVM and TF-IDF Methods,” Indonesian Journal of Computer Science Attribution, vol. 13, no. 1, pp. 2024–387, 2024.

[12] O. I. Gifari, M. Adha, I. Rifky Hendrawan, F. Freddy, and S. Durrand, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” JIFOTECH (Journal of Information Technology), vol. 2, no. 1, 2022.

[13] P. Indriyanti and W. Gunawan, “Pemanfaatan Chi Square dan Ensemble Tree Classifier pada Model SVM, KNN dan C4.5 dalam Penjualan Online,” Faktor Exacta, vol. 17, no. 3, p. 314, Oct. 2024, doi: 10.30998/faktorexacta.v17i3.24149.

[14] S. Afriyani, S. Surono, and I. M. Solihin, “Chi-Square Feature Selection with Pseudo-Labelling in Natural Language Processing,” JTAM (Jurnal Teori dan Aplikasi Matematika), vol. 8, no. 3, p. 896, Jul. 2024, doi: 10.31764/jtam.v8i3.22751.

[15] M. Rakha, M. Dwi Sulistiyo, D. Nasien, and M. Ridha, “A Combined MobileNetV2 and CBAM Model to Improve Classifying the Breast Cancer Ultrasound Images,” Journal of Applied Engineering and Technological Science, vol. 6, no. 1, pp. 561–578, 2024.

[16] M. Ridha, D. Nurjanah, and M. Rakha, “Multilabel Classification Abusive Language and Hate Speech on Indonesian Twitter using Transformer Model: IndoBERTweet & IndoRoBERTa,” in The 4th International Conference on Intelligent Cybernetics Technology & Applications 2024 (ICICyTA 2024), 2024.