



Analisis Sentimen Pengguna Twitter terhadap Tren *Cryptocurrency* di Indonesia dengan Metode *Support Vector Machine*

Muhammad Farhan Asshiddiq^{1*}, Betha Nurina Sari², Garno³

¹⁻³Program Studi Informatika, Universitas Singaperbangsa Karawang, Indonesia

Email: 1910631170211@student.unsika.ac.id¹, Betha.nurina@staff.unsika.ac.id², garno@staff.unsika.ac.id³

*Penulis Korespondensi: 1910631170211@student.unsika.ac.id

Abstract. The rapid growth of cryptocurrency as a digital asset has attracted increasing public attention in Indonesia, influenced by the extensive adoption of the internet and social media. Among various social media platforms, Twitter (X) has become one of the primary platforms where users express opinions and engage in discussions about cryptocurrency. The information generated from these interactions can be utilized to identify public sentiment trends. This investigation intends to assess public opinion regarding cryptocurrency trends in Indonesia by applying the Support Vector Machine (SVM) algorithm. The research adopts the Knowledge Discovery in Database (KDD) methodology, which consists of data selection, preprocessing, transformation, data mining, and evaluation stages. Data were collected through a web crawling process using the Tweet Harvest application, resulting in 7,000 Indonesian-language tweets. After duplicate removal and preprocessing, 4,502 tweets were retained as the research dataset. Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method to convert textual data into numerical representations, while sentiment classification was conducted using a linear-kernel Support Vector Machine. Model performance was evaluated using a Confusion Matrix. The experimental results demonstrated that the proposed model achieved an accuracy of 82.35%, precision of 84%, recall of 83%, and an F1-score of 84%. These findings indicate that the combination of TF-IDF and Support Vector Machine provides effective performance for classifying public sentiment regarding cryptocurrency trends in Indonesia.

Keywords: Cryptocurrency; Sentiment Analysis; Support Vector Machine; TF-IDF; Twitter.

Abstrak. Perkembangan *cryptocurrency* sebagai aset digital telah menarik perhatian masyarakat Indonesia seiring meningkatnya pemanfaatan internet dan media sosial. Salah satu platform yang banyak dipakai untuk menyampaikan opini mengenai perkembangan *cryptocurrency* adalah Twitter (X), sehingga data yang dihasilkan dapat dimanfaatkan untuk mengidentifikasi kecenderungan sentimen publik. Penelitian ini bertujuan untuk mengevaluasi sentimen publik pada tren *cryptocurrency* di Indonesia melalui penerapan algoritma *Support Vector Machine* (SVM). Metode *Knowledge Discovery in Database* (KDD) diterapkan dalam penelitian ini, yang meliputi tahapan pemilihan data, preprocessing, transformasi, data mining, dan evaluasi. Proses pengumpulan data dilakukan melalui crawling menggunakan Tweet Harvest, yang menghasilkan 7.000 tweet berbahasa Indonesia. Setelah dilakukan penghapusan data duplikat dan tahapan preprocessing, diperoleh 4.502 tweet yang digunakan sebagai dataset penelitian. Representasi dengan fitur yang dihasilkan dari *Term Frequency–Inverse Document Frequency* (TF-IDF), sedangkan untuk klasifikasi menggunakan algoritma SVM dengan kernel linear. Evaluasi kinerja model dilakukan menggunakan *Confusion Matrix*. Pengujian yang dilakukan menunjukkan bahwa model memperoleh nilai *accuracy* sebesar 82,35%, *precision* sebesar 84%, *recall* sebesar 83%, dan *F1-score* sebesar 84%. Temuan tersebut menunjukkan bahwa kombinasi TF-IDF dan SVM efektif untuk mengklasifikasikan sentimen masyarakat terkait tren *cryptocurrency* di Indonesia.

Kata kunci: Analisis Sentimen; *Cryptocurrency*; *Support Vector Machine*; TF-IDF; Twitter.

1. LATAR BELAKANG

Perkembangan teknologi informasi telah membawa banyak perubahan di berbagai bidang, terutama pada sektor keuangan. Salah satu inovasi yang berkembang pesat adalah *cryptocurrency*, yaitu aset digital berbasis teknologi *blockchain* yang memungkinkan terjadinya transaksi secara terdesentralisasi tanpa memerlukan pihak ketiga sebagai perantara. Selain dimanfaatkan sebagai instrumen investasi, *cryptocurrency* juga mulai digunakan

sebagai alternatif aset keuangan digital di berbagai negara, termasuk Indonesia (Liestyowati et al., 2023; Darmawansyah & Kadyanan, 2024).

Meningkatnya minat masyarakat terhadap cryptocurrency tidak terlepas dari perkembangan teknologi finansial, kemudahan memperoleh informasi, serta bertambahnya jumlah investor aset digital. Di sisi lain, pesatnya perkembangan internet dan media sosial mempercepat penyebaran berbagai informasi mengenai *cryptocurrency*. Kondisi tersebut memungkinkan masyarakat untuk mengikuti berita, berdiskusi, serta bertukar pengalaman terkait investasi aset digital dengan lebih mudah. Berdasarkan laporan Digital 2022: Di Indonesia, meningkatnya pengguna internet dan media sosial menjadi salah satu sumber informasi utama dalam memperoleh informasi mengenai perkembangan cryptocurrency (Kepios, 2022). Akibatnya, informasi yang beredar melalui media sosial turut memengaruhi persepsi serta keputusan masyarakat dalam berinvestasi pada aset digital (Kına & Biçek, 2024).

Fenomena *Fear of Missing Out* (FOMO) juga menjadi salah satu faktor yang mendorong meningkatnya minat masyarakat terhadap investasi *cryptocurrency*. FOMO menggambarkan kondisi ketika seseorang terdorong mengikuti suatu tren karena khawatir kehilangan kesempatan memperoleh keuntungan. Oleh sebab itu, berbagai informasi dan opini yang tersebar di media sosial sering kali berpengaruh terhadap cara pandang serta keputusan masyarakat dalam melakukan investasi cryptocurrency (Kholilullah et al., 2024).

Salah satu media sosial yang banyak dimanfaatkan untuk membahas cryptocurrency adalah Twitter (X). Platform ini menjadi sarana bagi pengguna untuk menyampaikan opini, pengalaman, maupun tanggapan terkait perkembangan cryptocurrency. Besarnya jumlah data yang dihasilkan pengguna bersifat tidak terstruktur sehingga analisis secara manual menjadi kurang optimal (Ningsih et al., 2024). Maka dari itu, diperlukan teknik analisis sentimen untuk mengelompokkan opini masyarakat dengan kategori sentimen positif maupun negatif (Kholilullah et al., 2024). Dalam *Natural Language Processing* (NLP), analisis sentimen bertujuan untuk mengekstraksi informasi dan memberikan gambaran tentang kecenderungan persepsi masyarakat terhadap sebuah topik (Puspitasari & Sutabri, 2024; Aryanti & Mahendra, 2023).

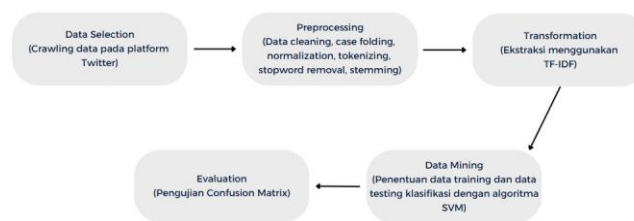
Berbagai penelitian sebelumnya menunjukkan bahwa algoritma *Support Vector Machine* (SVM) memiliki kinerja klasifikasi yang baik. Penelitian oleh Sumartini dan Wisudawati (2024) menunjukkan bahwa penerapan TF-IDF yang dikombinasikan dengan SVM mampu meningkatkan performa klasifikasi pada analisis sentimen aplikasi investasi. Penelitian lain oleh Wahid et al. (2024) juga menyatakan bahwa SVM dapat memberikan hasil yang lebih baik dibandingkan algoritma *Naïve Bayes* dalam menganalisis sentimen Twitter.

Selain itu, penelitian Noor dan Normawati (2025) membuktikan bahwa SVM dapat menghasilkan klasifikasi yang lebih baik dalam menganalisis sentimen pada isu *Artificial Intelligence* di Twitter.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis hasil sentimen pengguna Twitter terhadap tren *cryptocurrency* di Indonesia dengan algoritma *Support Vector Machine* (SVM). Proses penelitian dilakukan dengan menerapkan metode *Knowledge Discovery in Database* (KDD) yang meliputi tahapan pengumpulan data, preprocessing, ekstraksi fitur menggunakan TF-IDF, proses klasifikasi menggunakan SVM, dan melakukan evaluasi model melalui *Confusion Matrix*. Hasil penelitian ini diharapkan dapat memberikan gambaran kepada publik mengenai kecenderungan sentimen terhadap *cryptocurrency* sekaligus dapat menjadi referensi pada penelitian selanjutnya dalam menganalisis sentimen pada media sosial.

2. METODE PENELITIAN

Metode *Knowledge Discovery in Database* (KDD) diterapkan dalam penelitian ini sebagai basis untuk analisis sentimen. Metode ini dipilih karena kemampuannya dalam mengolah data dalam skala besar secara sistematis melalui tahapan *data selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation* sehingga sesuai untuk penelitian berbasis data teks dari media sosial (Joseph, 2022). Tahapan-tahapan yang akan dilakukan sesuai pada Gambar 1.



Gambar 1. Rancangan Penelitian KDD

Data Selection

Tahap *data selection* bertujuan memperoleh data yang selanjutnya akan dipakai dalam penelitian. Dataset diperoleh pada platform Twitter (X) dengan bantuan aplikasi *Tweet Harvest* dengan kata kunci "kripto" dan filter bahasa Indonesia. Proses *crawling* menghasilkan sebanyak 7.000 *tweet*. Selanjutnya dilakukan proses seleksi dengan menghapus data duplikat, *tweet* kosong, serta data yang tidak relevan sehingga diperoleh 4.502 *tweet* yang akan dipakai sebagai dataset.

Data Preprocessing

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum memasuki proses klasifikasi. Tahapan ini dilakukan untuk mengurangi *noise* menjadi lebih terstruktur dan mudah digunakan oleh algoritma klasifikasi (Sebastian et al.)

Tahapan preprocessing meliputi: *Cleaning*, yaitu menghapus URL, *mention*, *hashtag*, emoji, angka, tanda baca, dan karakter khusus yang tidak memiliki makna dalam proses analisis sentimen. *Case Folding*, yaitu proses mengganti semua huruf menjadi huruf kecil (*lowercase*) agar penulisan kata menjadi seragam. *Tokenizing*, yaitu proses memisahkan kalimat menjadi kumpulan kata (*token*). *Stopword Removal*, yaitu proses menghapus kata-kata umum seperti *dan*, *yang*, *di*, serta *ke* karena kata-kata tersebut tidak memberikan informasi penting terhadap proses klasifikasi. *Stemming*, yaitu mengganti kata yang memiliki imbuhan menjadi kata dasar (Maulana et al.).

Pelabelan Sentimen

Pada tahap ini menggunakan pendekatan *lexicon-based* dengan memanfaatkan *InSet Lexicon* Bahasa Indonesia. Setiap kata pada *tweet* dicocokkan bersama kamus sentimen yang memiliki bobot positif maupun negatif. Skor sentimen setiap *tweet* diperoleh dari akumulasi bobot setiap kata yang terdapat pada *tweet* tersebut. *Tweet* dengan nilai sentimen lebih besar dari nol diberi label positif, sedangkan *tweet* dengan nilai kurang dari nol diberi label negatif (Fernandes et al.).

Pembobotan TF-IDF

Pada tahap ini, data teks yang sudah diproses diubah menjadi representasi numerik menggunakan metode TF-IDF. Metode ini digunakan untuk menetapkan bobot pada setiap kata sesuai dengan frekuensi kemunculannya dalam dokumen dan secara keseluruhan, sehingga kata yang lebih representatif memiliki bobot yang lebih besar (Wahyuningsih & Chen, 2024). Perhitungan *Term Frequency* (TF) ditunjukkan pada Persamaan .

Perhitungan *term frequency* (TF)

$$tf = tf_{i,j} \dots\dots\dots(i)$$

Perhitungan *inverse document frequency* (IDF) ditunjukkan pada Persamaan .

$$idf_i = \log (N/df_i) \dots\dots\dots(ii)$$

Perhitungan persamaan TF-IDF menggunakan Persamaan (3).

$$W_{ij} = tf_{i,j} \cdot \log (N/df_i) \dots\dots\dots(iii)$$

dengan:

TF = frekuensi kemunculan kata pada dokumen

DF = jumlah dokumen yang memuat kata

N = jumlah seluruh dokumen

Klasifikasi Menggunakan Support Vector Machine

Tahap klasifikasi ini menggunakan algoritma *Support Vector Machine* (SVM). Algoritma ini bekerja dengan membentuk *hyperplane* optimal yang dapat memisahkan dua kelas berdasarkan margin maksimum sehingga proses klasifikasi dapat dilakukan secara efektif (Kristiyanto, 2025; Amelia et al., 2024). Kernel yang digunakan adalah kernel *linear* karena sesuai untuk data teks yang telah diekstraksi menggunakan TF-IDF dan memiliki jumlah fitur yang tinggi. Dataset kemudian dipecah menjadi 80% data latih dan 20% data uji. Data latih yang akan digunakan dalam pelatihan model klasifikasi, sedangkan data uji dimanfaatkan untuk mengukur kemampuan model dalam mengklasifikasikan data yang belum pernah diproses sebelumnya (Joseph, 2022).

Evaluasi Model

Tahap terakhir adalah mengevaluasi performa model menggunakan *Confusion Matrix*. Evaluasi dilakukan berdasarkan nilai *accuracy*, *precision*, *recall*, dan F1-Score untuk mengetahui tingkat keberhasilan model dalam mengklasifikasikan sentimen pengguna Twitter (Prapasta & Zainuddin, 2025). Pengklasifikasian *tuple Confusion Matrix* dapat dilihat pada Tabel 1 dibawah ini.

Tabel 1. Confusion Matrix

	<i>Predicted Class (Positive)</i>	<i>Predicted Class (Negative)</i>
<i>Actual Class (Positive)</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Actual Class (Negative)</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Dengan keterangan :

TP (*True Positive*) = data yang berhasil diklasifikasi menjadi kelas Positif

FP (*False Positive*) = data yang tidak berhasil diklasifikasi menjadi kelas Positif

TN (*True Negative*) = data yang berhasil diklasifikasi menjadi kelas Negatif

FN (*False Negative*) = data yang tidak berhasil diklasifikasi menjadi kelas Negatif

Persamaan *accuracy* ditunjukkan pada Persamaan (4).

$$Akurasi.= \frac{TP. + TN}{TP. +. TN. +. FP. +. FN} \times 100\% \dots \dots \dots (iv)$$

Persamaan *recall* ditunjukkan pada Persamaan (5).

$$Recall = \frac{TP.}{FN. +. TP} \times 100\% \dots \dots \dots (v)$$

Persamaan *precision* ditunjukkan pada Persamaan (6).

$$Precision.= \frac{TP}{FP + TP} \times 100\% \dots \dots \dots (vi)$$

Persamaan *F1-Score* ditunjukkan pada Persamaan (7).

$$F1-Score. = 2 \times \frac{TP \times TN}{TP + TN + FP + FN} \times 100\% \dots\dots\dots(vii)$$

3. HASIL DAN PEMBAHASAN

Berikut ini hasil dan pembahasan analisis sentimen menggunakan algoritma *Support Vector Machine* (SVM).

Data Selection

Proses pengumpulan data ini dilakukan pada platform X (Twitter) dengan memanfaatkan aplikasi *Tweet Harvest*. Pengambilan data diambil dari kata kunci yang berhubungan dengan *cryptocurrency* sehingga *tweet* yang diperoleh sesuai dengan topik penelitian. Hasil proses *crawling* berhasil mengumpulkan 7.000 *tweet* berbahasa Indonesia. Selanjutnya dilakukan tahap *remove duplicate* untuk menghilangkan data yang berulang sehingga jumlah dataset yang layak digunakan berkurang menjadi 4.502 *tweet*. Contoh hasil pengumpulan data ditampilkan pada Gambar 2 dibawah ini.

	username	created_at	full_text
0	direkturcrypto	2026-07-08T13:53:49.000Z	@aryakhadi iyaa jadi yg punya toko mah ga per...
1	tbaryagunawan	2026-07-08T13:25:34.000Z	@Novaex_AI Mekanisme seperti ini tidak lazim d...
2	ibnu1981	2026-07-08T13:20:32.000Z	@BitgetWalletID Min, kalo kita bayar2 pake QR...
3	Airdropfinds	2026-07-08T13:16:11.000Z	RBI kembali menunjukkan sikap keras terhadap k...
4	Lordittothegod	2026-07-08T13:06:23.000Z	@indodax USDT Flash Ini adalah token kripto p...

Gambar 2. Hasil Pengumpulan Dataset

Penghapusan data duplikat bertujuan memastikan setiap *tweet* yang digunakan memiliki informasi yang unik. Dengan demikian, dataset menjadi lebih representatif dan dapat mengurangi potensi bias pada saat proses pelatihan maupun pengujian model klasifikasi.

Preprocessing

Tahap preprocessing dilakukan sebelum proses ekstraksi fitur dengan tujuan meningkatkan kualitas dataset agar lebih siap untuk dianalisis. Pada tahap ini, data teks diproses melalui beberapa tahapan, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Setiap tahap memiliki fungsi untuk mereduksi elemen yang tidak diperlukan sekaligus menyeragamkan bentuk teks sehingga lebih mudah diproses pada tahap berikutnya.

Tabel 2. Contoh Hasil Preprocessing

Tahapan	Contoh
Dataset	@cryptorover @aixbt agent Karena obligasi melonjak, apa dampaknya terhadap dollar as dan pasar kripto?
Cleaning	Karena obligasi melonjak apa dampaknya terhadap dollar as dan pasar kripto

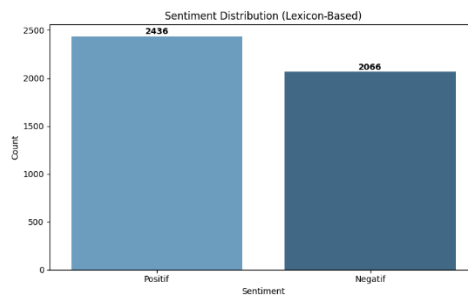
Pelabelan Sentimen

Proses pelabelan sentimen pada penelitian ini menerapkan pendekatan InSet Lexicon, sehingga setiap tweet dapat dikategorikan sebagai sentimen positif atau negatif berdasarkan skor sentimen yang diperoleh. Penentuan label dilakukan dengan menghitung akumulasi bobot setiap kata yang terdapat pada tweet, kemudian hasilnya digunakan sebagai dasar dalam menetapkan kelas sentimen. Distribusi hasil pelabelan disajikan pada Tabel 3.

Tabel 3. Hasil labeling

Sentimen	Jumlah
Positif	2436
Negatif	2066
Total	4502

Berdasarkan hasil pelabelan, diperoleh 2.436 *tweet* sentimen positif dan 2.066 *tweet* dengan sentimen negatif dari total 4.502 *tweet*. Untuk memudahkan analisis, distribusi tersebut juga ditampilkan dalam bentuk *bar chart* pada Gambar 5.



Gambar 5. Perbandingan Data

Hasil distribusi menunjukkan bahwa jumlah tweet dengan sentimen positif lebih banyak dibandingkan sentimen negatif. Kondisi ini menandakan jika sebagian besar pengguna Twitter yang berpendapat cenderung positif terhadap perkembangan *cryptocurrency* di Indonesia. Meskipun demikian, masih ditemukan sejumlah opini negatif yang umumnya berkaitan dengan risiko investasi, fluktuasi harga aset, serta potensi penipuan (*scam*) dalam ekosistem *cryptocurrency*.

Pembobotan TF-IDF

Tahap berikutnya adalah melakukan ekstraksi fitur dengan TF-IDF. Metode ini berfungsi untuk mengkonversi data teks menjadi bentuk representasi numerik sehingga setiap kata memperoleh bobot sesuai tingkat kepentingannya dalam keseluruhan dokumen. Ringkasan hasil pembobotan disajikan pada Tabel 4.

Tabel 4. Hasil pembobotan TF-IDF

Keterangan	Nilai
Jumlah Tweet	4502
Jumlah Fitur	5199

Hasil pembobotan menunjukkan bahwa dataset yang terdiri atas 4.502 tweet menghasilkan 5.199 fitur. Banyaknya jumlah fitur tersebut mencerminkan keberagaman kosakata yang terdapat pada dataset penelitian. Oleh karena itu, penggunaan metode TF-IDF menjadi penting karena mampu memberikan bobot yang lebih besar pada setiap kata yang memiliki kontribusi tinggi dalam proses klasifikasi, sekaligus mengurangi pengaruh kata-kata yang kurang representatif.

Klasifikasi Menggunakan *Support Vector Machine* (SVM)

Tahap klasifikasi dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel linear. Sebelum model dibangun, dataset dibagi menjadi 80% data latih dan 20% data uji. Pembagian ini bertujuan agar model dapat dipelajari menggunakan sebagian besar data, kemudian dievaluasi menggunakan data yang belum pernah diproses sebelumnya.

```
Jumlah Data Training : (3601, 5199)
Jumlah Data Testing  : (901, 5199)

Distribusi Data Training
sentimen
Positif    1948
Negatif    1653
Name: count, dtype: int64

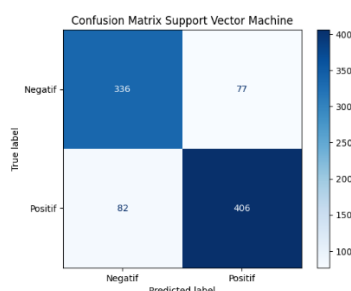
Distribusi Data Testing
sentimen
Positif     488
Negatif     413
Name: count, dtype: int64
```

Gambar 6. Pembagian data latih dan data uji

Berdasarkan Gambar 6 diatas, model dilatih menggunakan 3.601 data latih, sedangkan 901 data uji digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan sentimen. Selanjutnya, model melakukan prediksi terhadap data uji, kemudian hasil prediksi tersebut dievaluasi dengan *Confusion Matrix* untuk memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebagai indikator performa model klasifikasi.

Evaluation

Evaluasi terhadap model dilakukan menggunakan *Confusion Matrix* untuk mengetahui tingkat keberhasilan algoritma SVM dalam mengklasifikasikan sentimen pengguna Twitter. Berdasarkan hasil evaluasi yang ditampilkan pada Gambar 8, model mendapatkan akurasi sebesar 82,35%, *precision* 84%, *recall* 83%, dan *F1-score* 84%.



Gambar 7. Confusion Matrix

Berdasarkan hasilnya, diperoleh nilai evaluasi pada Gambar 8 dibawah ini..

Accuracy : 0.8235294117647058				
Classification Report				
	precision	recall	f1-score	support
Negatif	0.80	0.81	0.81	413
Positif	0.84	0.83	0.84	488
accuracy			0.82	901
macro avg	0.82	0.82	0.82	901
weighted avg	0.82	0.82	0.82	901

Gambar 8. Nilai Evaluasi

Nilai akurasi ini menunjukkan bahwa model mampu dengan baik dalam mendeteksi sentimen terkait tren *cryptocurrency* di Indonesia. Selain itu, nilai *precision*, *recall*, dan *F1-score* yang seimbang menunjukkan bahwa model tidak hanya dapat memberikan prediksi yang tepat, tetapi juga cukup konsisten dalam mengidentifikasi data di setiap kategori sentimen. Kombinasi *preprocessing*, TF-IDF, dan algoritma SVM terbukti mampu menghasilkan performa klasifikasi yang baik pada berbagai penelitian analisis sentimen berbasis Twitter [20].

Hasil penelitian ini memperlihatkan bahwa kombinasi TF-IDF sebagai metode ekstraksi fitur dan algoritma SVM mendapatkan performa yang efektif dalam menganalisis sentimen pada data teks berbahasa Indonesia. Temuan tersebut juga selaras dengan penelitian terdahulu yang menyatakan bahwa TF-IDF mampu meningkatkan kualitas representasi fitur, sedangkan SVM efektif dalam menangani proses klasifikasi pada data teks dengan dimensi fitur yang tinggi.

4. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan algoritma SVM dalam melakukan analisis sentimen terhadap opini pengguna Twitter mengenai tren *cryptocurrency* di Indonesia. Proses penelitian mengacu pada metodologi KDD yang memiliki beberapa tahapan, yaitu *data selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation*. Data penelitian dikumpulkan melalui aplikasi *Tweet Harvest*, menghasilkan 7.000 tweet berbahasa Indonesia. Setelah dilakukan proses penghapusan data duplikat dan tahapan *preprocessing*, diperoleh 4.502 *tweet* yang selanjutnya digunakan sebagai dataset penelitian.

Pada tahap ekstraksi fitur, metode TF-IDF berhasil mengonversi data teks menjadi bentuk numerik dengan menghasilkan 5.199 fitur yang digunakan sebagai masukan bagi algoritma SVM. Berdasarkan hasil evaluasi dengan menggunakan *Confusion Matrix*, model memperoleh akurasi sebesar 82,35%, *precision* 84%, *recall* 83%, dan *F1-score* 84%. Hasil tersebut menunjukkan bahwa kombinasi TF-IDF dan SVM berhasil memberikan hasil yang

baik dalam mengklasifikasikan sentimen masyarakat terhadap perkembangan *cryptocurrency* di Indonesia.

Untuk penelitian berikutnya, disarankan menambah jumlah dataset yang lebih besar agar variasi data yang dianalisis semakin beragam. Selain itu, penelitian selanjutnya dapat membandingkan performa *Support Vector Machine* dengan algoritma klasifikasi lain, seperti *Random Forest*, *Naïve Bayes*, serta metode ekstraksi fitur lainnya juga dapat dipertimbangkan untuk meningkatkan kualitas representasi data serta performa klasifikasi.

DAFTAR REFERENSI

- Amelia, I., Sugiyono, Sarimole, F. M., & Tundo. (2024). Analisis sentimen tanggapan pengguna media sosial X terhadap program Beasiswa KIP-Kuliah dengan menggunakan algoritma *Support Vector Machine* (SVM). *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 5(3), 2994–3003. <https://doi.org/10.35870/jimik.v5i3.990>
- Anggraini, J., & Alita, D. (2024). Implementasi metode SVM pada sentimen analisis terhadap pemilihan presiden (Pilpres) 2024 di Twitter. *Jurnal Informatika: Jurnal Pengembangan IT*. <https://doi.org/10.30591/jpit.v9i2.6560>
- Aryanti, P. A. N., & Mahendra, I. B. M. (2023). Analisis sentimen opini berbahasa Indonesia pada sosial media menggunakan TF-IDF dan *Support Vector Machine*. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, 12(1), 45–52.
- Darmawansyah, D., & Kadyanan, I. G. A. G. A. (2024). Analisis prediktif Bitcoin dengan metode *Support Vector Machine* serta pembobotan TF-IDF berbasis data narrative. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(1), 81–90.
- Fernandes, Y. B., Haerani, E., Syafria, F., Fikry, M., & Oktavia, L. (2025). Penerapan metode *Support Vector Machine* untuk analisis sentimen pada komentar Bitcoin di aplikasi X. *Bulletin of Computer Science Research*, 6(1).
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, 15(4), 531–538.
- Kepios. (2022). *Digital 2022: Indonesia*.
- Kholilullah, M., Martanto, M., & Hayati, U. (2024). Analisis sentimen pengguna Twitter (X) tentang Piala Dunia U-17 menggunakan metode *Naïve Bayes*. *JATI*, 8(1). <https://doi.org/10.36040/jati.v8i1.8378>
- Kına, E., & Biçek, E. (2024). Machine learning approach for emotion identification and classification in Bitcoin sentiment analysis. *Yüzüncü Yıl University Journal of the Institute of Natural and Applied Sciences*, 29(3), 913–926. <https://doi.org/10.53433/yyufbed.1532649>
- Kristiyanto, R. (2025). *Implementasi algoritma Support Vector Machine untuk klasifikasi opini publik terhadap insiden peretasan Indodax* (Skripsi, Universitas Multimedia Nusantara).
- Liestyowati, E., Sudarmanto, E., Ramadhani, H., Rijal, S., & Nurdiani, T. W. (2023). Tren investasi aset digital: Studi tentang perilaku investor muda terhadap *cryptocurrency* di

- tengah perubahan pasar keuangan di Kota Bandung. *Jurnal Akuntansi dan Keuangan West Science*, 2(3). <https://doi.org/10.58812/jakws.v2i03.639>
- Maulana, B. A., Fahmi, M. J., Imran, A. M., & Hidayati, N. (2024). Analisis sentimen terhadap aplikasi Pluang menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine*. *MALCOM*, 4(2), 375–384. <https://doi.org/10.57152/malcom.v4i2.1206>
- Ningsih, W., Alfianda, B., Rahmadden, & Wulandari, D. (2024). Perbandingan algoritma SVM dan *Naïve Bayes* dalam analisis sentimen Twitter pada penggunaan mobil listrik di Indonesia. *MALCOM*, 4(2). <https://doi.org/10.57152/malcom.v4i2.1253>
- Noor, A. R., & Normawati, D. (2025). Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* pada analisis sentimen pelaksanaan pemilihan presiden tahun 2024. *JUPI*. <https://doi.org/10.29100/jupi.v1i1i2.8120>
- Prapasta, D., & Zainuddin, M. (2025). Klasifikasi sentimen terhadap ulasan pengguna aplikasi DANA menggunakan metode *Support Vector Machine*. *JUSIFOR*, 5(1). <https://doi.org/10.70609/jusifor.v5i1.8626>
- Puspitasari, D., & Sutabri, T. (2024). Analisis sentimen menggunakan *Natural Language Processing* pada media sosial Twitter. *Jurnal Media Informatika Budidarma*, 8(1).
- Sebastian, D. F., Sulistiani, H., & Isnain, A. R. (2024). Sentiment analysis of public opinion on the right of inquiry in Indonesia using *Support Vector Machine*. *JUTIF*, 5(4). <https://doi.org/10.52436/1.jutif.2024.5.4.1968>
- Sumartini, D., & Wisudawati, L. M. (2024). Analisis sentimen pada ulasan aplikasi Tokocrypto menggunakan *Support Vector Machine*. *Jurnal Ilmiah Informatika Komputer*, 29(3). <https://doi.org/10.35760/ik.2024.v29i3.12915>
- Wahid, Y. A., Sanatang, & Andayani, D. D. (2024). Performance comparison of SVM and *Naïve Bayes* for Indonesian-language sentiment analysis using TF-IDF and SMOTE. *Journal of Embedded Systems, Security and Intelligent Systems*, 6(4). <https://doi.org/10.59562/jessi.v6i4.10818>
- Wahyuningsih, T., & Chen, S. C. (2024). Analyzing sentiment trends and patterns in Bitcoin-related tweets using TF-IDF vectorization. *Journal of Current Research in Blockchain*, 1(1). <https://doi.org/10.47738/jcrb.v1i1.11>