



INNOVATIVE: Journal Of Social Science Research

Volume 4 Nomor 5 Tahun 2024 Page 6159-6172

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

## Analisis Sentimen *Non-Fungible Token Solana* Di Media Sosial X Dengan *Random Forest* Dan *Support Vector Machine*

Muhammad Verel Prisco Alfito Devani<sup>1✉</sup>, Mukh. Taufik Chulkamdi<sup>2</sup>, Mohammad Faried  
Rahmat<sup>3</sup>

Universitas Islam Balitar

Email: [mhdverel27@gmail.com](mailto:mhdverel27@gmail.com)<sup>1✉</sup>

### Abstrak

Penelitian ini menganalisis sentimen terhadap NFT pada blockchain Solana di platform media sosial X menggunakan Support Vector Machine (SVM) dan Random Forest. Analisis ini penting bagi pengembang, investor, dan pemangku kepentingan lainnya untuk memahami sentimen publik, merancang strategi pemasaran, mengidentifikasi kebutuhan pasar, serta mengantisipasi tren masa depan. Data yang digunakan berupa tweet terkait NFT Solana. Metode SVM dengan kernel linear dan Random Forest masing-masing mencapai tingkat akurasi 69,56%, presisi 56,46%, recall 69,56%, dan F1-score 57,41%. Visualisasi kata-kata dominan dalam tweet menggunakan wordcloud menambah wawasan tentang sentimen publik. Meski efektif, kedua metode ini masih perlu peningkatan, terutama dalam mengatasi kesalahan klasifikasi *False Negatives*. Penelitian selanjutnya disarankan untuk meningkatkan volume data, menggunakan ensemble learning seperti SMOTE untuk menangani ketidakseimbangan kelas, dan mengeksplorasi algoritma lain. Studi ini memberikan wawasan berharga untuk pemangku kepentingan dalam ekosistem NFT Solana.

Kata Kunci: Analisis Sentimen, *NFT*, *Blockchain Solana*, Media Sosial X, *Support Vector Machine (SVM)*, *Random Forest*, *Wordcloud*.

## Abstract

This research analyzes sentiment towards NFTs on the Solana blockchain on the social media platform X using Support Vector Machine (SVM) and Random Forest. This analysis is crucial for developers, investors, and other stakeholders to understand public sentiment, design marketing strategies, identify market needs, and anticipate future trends. The data used consists of tweets related to Solana NFTs. The SVM method with a linear kernel and Random Forest each achieved an accuracy rate of 69.56%, precision of 56.46%, recall of 69.56%, and F1-score of 57.41%. Visualization of dominant words in the tweets using word clouds adds insights into public sentiment. Although effective, both methods still need improvement, especially in addressing False Negative misclassifications. Future research is recommended to increase the data volume, use ensemble learning like SMOTE to handle class imbalance, and explore other algorithms. This study provides valuable insights for stakeholders in the Solana NFT ecosystem.

Keyword: *Sentiment Analysis, NFT, Solana Blockchain, Social Media X, Support Vector Machine (SVM), Random Forest, Wordcloud.*

## PENDAHULUAN

Dalam era digital yang sedang berkembang, teknologi *blockchain* telah mencapai kemajuan signifikan dan memberikan dampak yang luas pada berbagai aspek kehidupan manusia. Sebagai sebuah teknologi yang mendasari berbagai mata uang kripto, *blockchain* adalah sebuah ledger terdistribusi yang menyimpan catatan transaksi dalam blok-blok yang dihubungkan satu sama lain secara kronologis. Dengan struktur ini, setiap blok dalam *blockchain* mengandung serangkaian transaksi yang terenkripsi dan tidak dapat diubah, menciptakan sistem yang aman dan transparan di mana informasi dapat disimpan tanpa kebutuhan akan otoritas pusat (Nakamoto, 2009).

Salah satu sektor yang terpengaruh secara signifikan adalah sektor keuangan dan investasi (Irawan dkk., 2020). Sebagai salah satu platform blockchain paling populer, Solana telah menjadi sorotan bagi banyak investor dan pengembang, terutama dalam konteks *Non-Fungible Tokens (NFT)*. Karakteristik unik dari NFT yang tidak dapat dipertukarkan secara langsung dengan aset lain telah menjadikannya topik menarik yang berpotensi memunculkan beragam sentimen di kalangan masyarakat (Yuda Lesmana & Andarsyah, 2022).

X, yang didirikan oleh Jack Dorsey, berfungsi sebagai platform media sosial terkemuka yang memungkinkan pengguna mengirim pesan melalui format yang dikenal sebagai "*tweet*" dan membagikan ulang ("*retweet*"). Saat ini, platform ini telah menjadi salah satu yang paling populer dan banyak digunakan untuk mengekspresikan opini dan pandangan

dari para pengguna (Krisdiyanto, 2021). Mayoritas isi di X berkisar pada penggunaannya sendiri, yang menggunakan platform tersebut untuk membagikan cerita, pendapat, dan kegiatan mereka kepada masyarakat. Pengguna X dibatasi hanya bisa mengirim pesan dengan panjang maksimal 280 karakter per kiriman. (Primadhani Tirtopangarsa & Maharani, 2021). Dengan adanya platform ini, masyarakat lebih cenderung untuk mengungkapkan pandangan mereka melalui media sosial daripada melakukannya secara langsung (Anggreni dkk., 2019).

Dalam konteks ini, pendekatan analisis sentimen menggunakan metode *Support Vector Machine (SVM)* telah terbukti efektif. SVM, sebuah algoritma dalam pembelajaran mesin, mampu mengelompokkan data dengan tingkat ketepatan yang tinggi (Ramadhanty, 2021). Pemanfaatan metode *Support Vector Machine (SVM)* dalam analisis sentimen untuk mengidentifikasi dan mengelompokkan opini atau pandangan pengguna X terhadap *NFT* di platform *blockchain* Solana dapat memberikan wawasan yang berharga bagi para pengembang, investor, dan stakeholder lainnya.

Selain SVM, metode Random Forest juga telah mendapatkan perhatian signifikan dalam analisis sentimen. Random Forest adalah algoritma ensemble learning yang terdiri dari beberapa pohon keputusan untuk meningkatkan akurasi prediksi dan mengurangi overfitting (Rahim dkk., 2023). Dengan menggabungkan hasil dari banyak pohon keputusan, Random Forest mampu memberikan prediksi yang lebih stabil dan andal. Dalam konteks analisis sentimen, metode Random Forest dapat digunakan untuk mengklasifikasikan opini pengguna dengan tingkat akurasi yang tinggi, serta memberikan wawasan tambahan melalui analisis fitur yang berkontribusi pada klasifikasi sentiment (Jannah dkk., 2024).

Semua individu yang terlibat dalam platform *blockchain Solana*, termasuk investor dan penggemar NFT, perlu memiliki pemahaman yang mendalam tentang dinamika sentimen ini. Analisis sentimen yang tepat dapat menjadi kunci dalam merancang strategi pemasaran yang efektif, mengidentifikasi kebutuhan pasar, dan mengantisipasi tren masa depan. Penelitian ini juga berperan penting dalam pengembangan metodologi analisis sentimen di platform media sosial.

Dengan demikian, keterkaitan antara objek (teks yang berisi opini pengguna) dan metode klasifikasi Random Forest serta SVM sangat penting dalam konteks analisis sentimen NFT pada *blockchain Solana* melalui media sosial X. SVM berfungsi sebagai alat yang kuat untuk mengklasifikasikan sentimen yang terkandung dalam objek teks, yang pada gilirannya memberikan wawasan berharga tentang pandangan pengguna terhadap NFT di platform tersebut.

## METODE PENELITIAN

Penelitian ini dilakukan di Pusat Komputer Universitas Islam Balitar (PuskomUnisba), yang menyediakan fasilitas dan infrastruktur yang diperlukan untuk analisis data dan komputasi. Objek penelitian dalam kajian ini adalah data teks yang berasal dari *platform* media sosial X, yang mencakup *tweet*, *retweet*, dan komentar yang berkaitan dengan topik *Non-FungibleTokens* (NFT) di *Blockchain Solana*. Subjek penelitian meliputi pengguna aktif platform media sosial X yang secara aktif berpartisipasi dalam diskusi, berbagi informasi, atau menyampaikan pendapat terkait dengan NFT di *Blockchain Solana* selama periode penelitian. Penelitian ini memiliki deskriptif pendekatan kuantitatif, yang mengutamakan pengumpulan data numerik dan analisis statistik untuk memahami fenomena yang diamati. Pengumpulan data dalam penelitian ini dilakukan dengan cara tidak langsung, misalnya melalui Studi Literatur, Observasi, untuk memperoleh informasi yang spesifik dan relevan. Dalam penelitian ini, digunakan perangkat lunak *Orange 3* untuk mengumpulkan data dari X, sementara *Jupyter Lab* digunakan untuk tahap preprocessing data dan evaluasi teknik klasifikasi. Penelitian ini memanfaatkan tweet yang diposting oleh pengguna media sosial X dengan menggunakan kata kunci terkait NFT *Solana* sebagai populasi. Sampel yang dianalisis mencakup *tweet-tweet* dari periode Januari hingga 1 Juli 2023 hingga 25 Januari 2024.

## HASIL DAN PEMBAHASAN

### a. Klasifikasi

#### 1. *Random Forest*

Selanjutnya pengujian algoritma Random Forrest dengan menginisialisasi algoritma menggunakan "*RandomForestClassifier(random\_state=42)*". Yang dimana "*random\_state*" digunakan untuk menentukan seed atau bilangan acak yang digunakan untuk menginisialisasi generator angka acak, sehingga hasil yang dihasilkan akan konsisten jika kode dijalankan kembali, dapat dilihat pada gambar.

```
# Membuat objek Random Forest Classifier
rf_classifier = RandomForestClassifier(random_state=42)

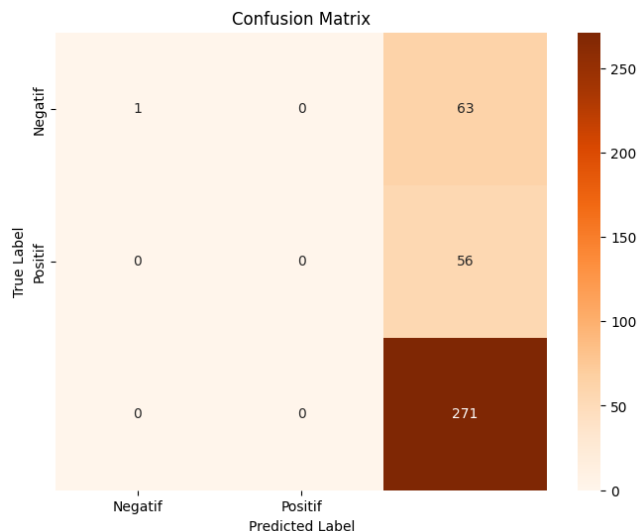
# Melatih model dengan data pelatihan
rf_classifier.fit(X_train, y_train)

# Melakukan prediksi pada data pengujian
y_pred = rf_classifier.predict(X_test)
```

Gambar 1. Implementasi Algoritma Random Forest

Setelah kode pada gambar dijalankan, algoritma Random Forest telah dilatih dan melakukan prediksi pada data pengujian. Hasil dari prediksi ini kemudian dapat digunakan

untuk mengevaluasi kinerja model dengan menggunakan confusion matrix dan evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*.



Gambar 2. Hasil Algoritma Random Forest

Berdasarkan hasil evaluasi model pada gambar confusion matrix menyatakan bahwa secara benar sebanyak 1 data sebagai negatif dan 271 sebagai data positif. Selain itu model salah dalam memprediksi 56 data ke dalam kelas positif yang seharusnya negatif (False Negatif).

Berdasarkan penjelasan confusion matrix dapat dideskripsikan bahwa klasifikasi algoritma Random Forrest memiliki tingkat akurasi: 69.56%, presisi: 56.46%, *recall*: 69.56%, dan *f1-scores*: 57.41% dalam mengklasifikasikan *tweet* yang mengandung nft solana atau tidak mengandung nft solana pada media sosial X.

## 2. Support Vector Machine

Pada tahap ini peneliti membagi data latih (80%) dan data uji (20%). selanjutnya pengujian algoritma *Support Vector Machine (SVM)* dengan menggunakan *SVC (Support Vector Classification)* dari scikit-learn, dapat dilihat pada gambar.

```
# Membuat objek SVM
svm_classifier = SVC(kernel='linear', random_state=42)

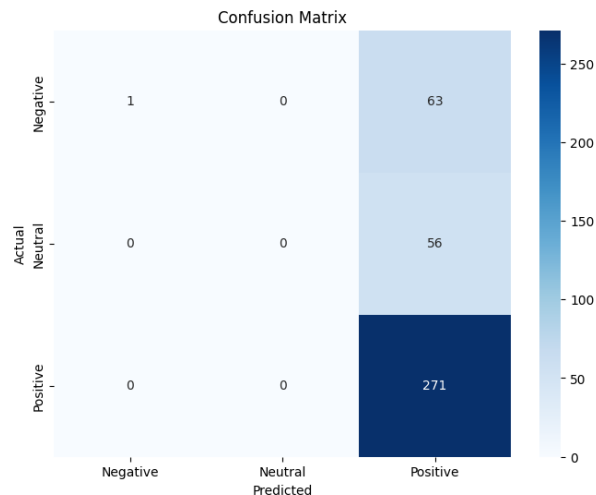
# Melatih model dengan data pelatihan
svm_classifier.fit(X_train, y_train)

# Melakukan prediksi pada data pengujian
y_pred = svm_classifier.predict(X_test)
```

Gambar 3. Confusion Matrix Random Forest

Dari gambar 3 bagian proses pemodelan menggunakan algoritma *Support Vector Machine* dengan kode "svm\_classifier" adalah objek untuk *SVM* dengan menentukan kernel yang digunakan. Dalam penelitian ini, peneliti menggunakan "kernel='linear'", yang artinya peneliti menggunakan kernel linear untuk melakukan pemetaan fitur. Dari pengujian tersebut dapat digunakan untuk melakukan evaluasi algoritma *SVM*, untuk menampilkan

*confusion matrix* dan menghitung metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*.



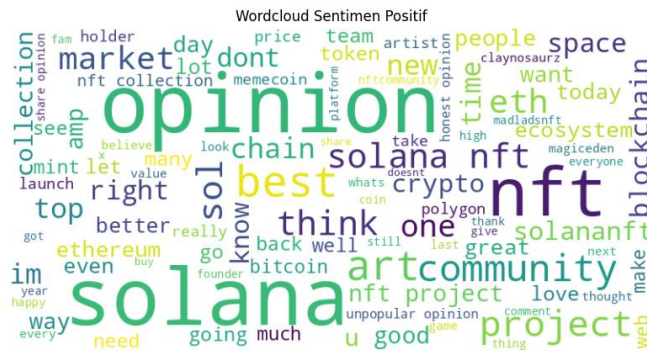
Gambar 4. Confusion Matrix SVM

Berdasarkan hasil evaluasi model pada Gambar *confusion matrix* menyatakan bahwa secara benar sebanyak 1 data sebagai negatif dan 271 sebagai data positif. Selain itu model salah dalam memprediksi 63 data ke dalam kelas positif yang seharusnya negatif (False Negatif).

Berdasarkan penjelasan confusion matrix dapat dideskripsikan bahwa klasifikasi algoritma Support Vector Machine memiliki tingkat akurasi: 69.56%, presisi: 56.46%, *recall*: 69.56%, dan *f1-scores*: 57.41% dalam mengklasifikasikan *tweet* yang mengandung nft solana atau tidak mengandung nft solana pada media sosial X.

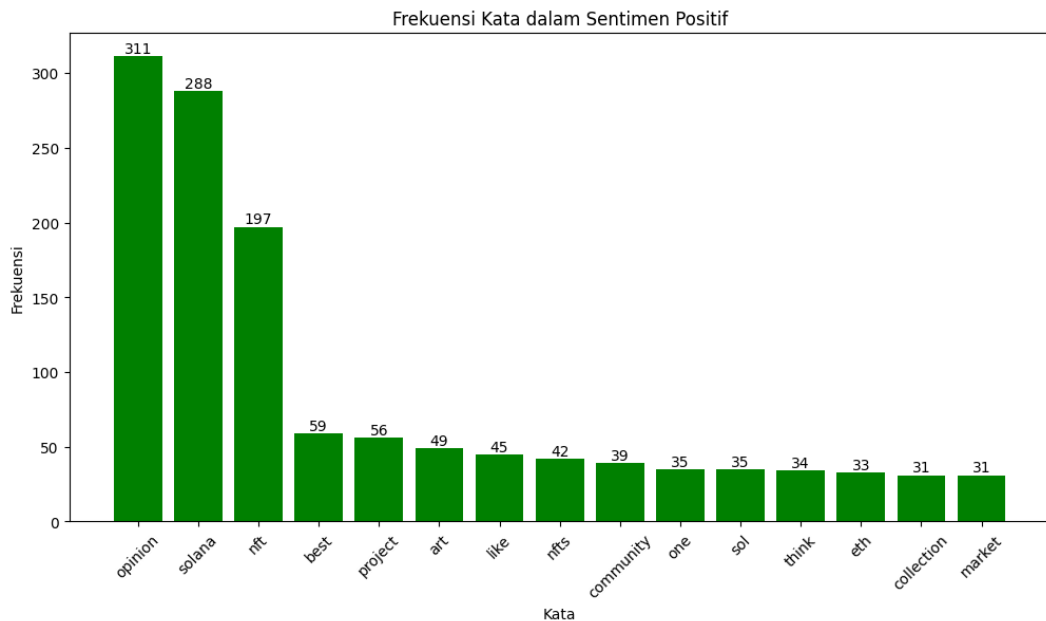
#### b. Visualisasi

Setelah mengetahui label atau sentiment pada data dan menguji algoritma klasifikasi, selanjutnya pada penelitian ini memvisualisasikan kata kata yang sering muncul dengan membuat *wordcloud*. Memvisualisasikan dengan *wordcloud* dapat melihat kata-kata yang dominan dalam teks. *Wordcloud* memberikan representasi visual yang menarik dari frekuensi kemunculan kata-kata, di mana kata-kata yang paling sering muncul ditampilkan dengan ukuran yang lebih besar. Berikut ini tampilan *wordcloud* yang diperoleh menggunakan *library* matplotlib dan *wordcloud* pada python.



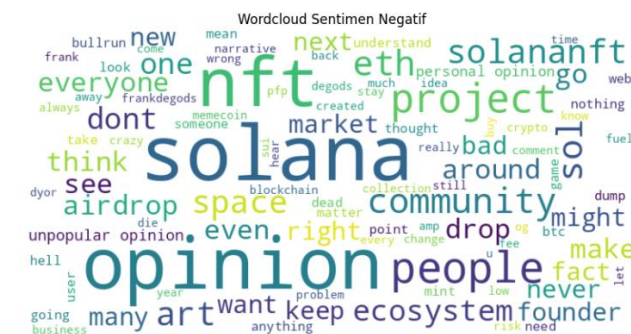
Gambar 5. Hasil Wordcloud Tweet Positif

Pada gambar 5 dapat dideskripsikan kata "opinion" dan "best" lebih dominan yang menunjukkan bahwa banyak pengguna media sosial X yang berbicara atau mengekspresikan perhatian mereka terhadap opini NFT Solana di X.



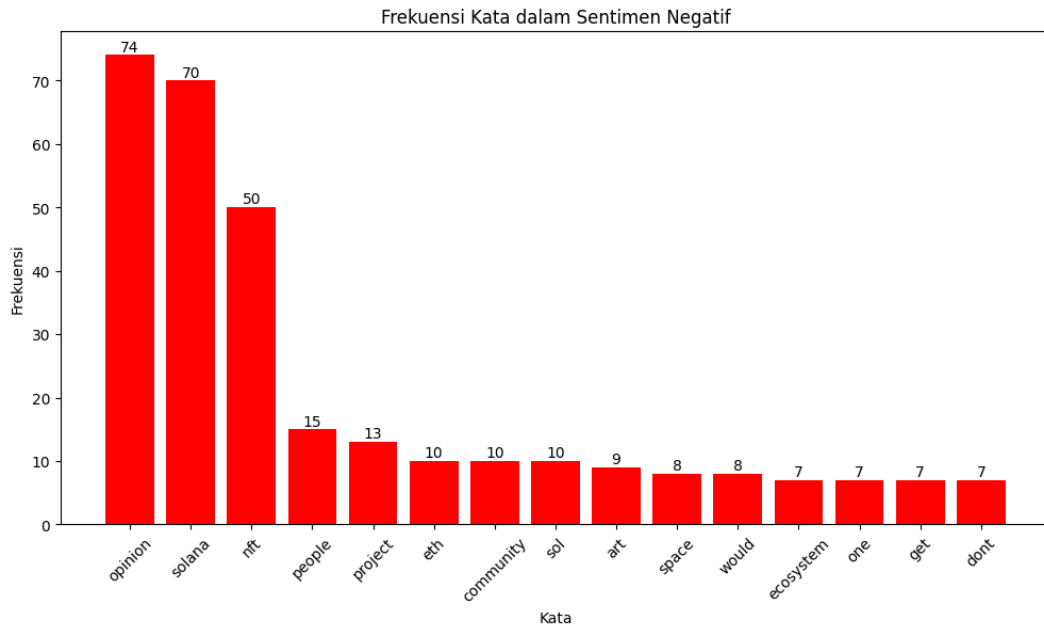
Gambar 6. Frekuensi kemunculan terbanyak kata sentimen positif

Hasil gambar 6. menunjukkan bahwa frekuensi kemunculan kata dalam sentimen positif dapat divisualisasikan dalam diagram batang. Dari diagram tersebut, terlihat bahwa kata "opinion" muncul sebanyak 311 kali, sedangkan kata yang mengandung "Solana" muncul sebanyak 288 kali.



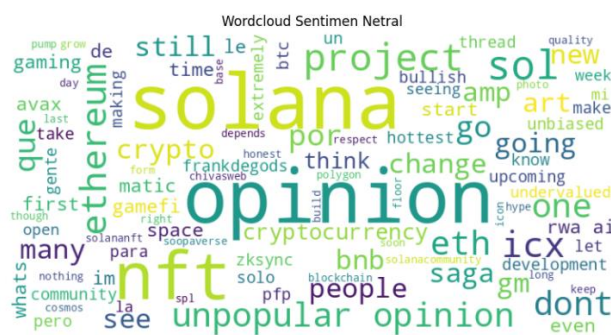
Gambar 7. Hasil Wordcloud Tweet Negatif

Pada gambar 7 dapat dideskripsikan kata "drop" pada *wordcloud* menunjukkan bahwa kata ini merupakan kata kunci yang sangat sering muncul dalam *tweet-tweet* dengan label negatif. Hal ini menunjukkan bahwa *drop* adalah topik yang sangat dominan dalam kategori *tweet-tweet* tersebut. Kemunculan kata "drop" pada word cloud yang berlabel negatif menunjukkan adanya sentimen negatif terkait dengan topik NFT Solana.



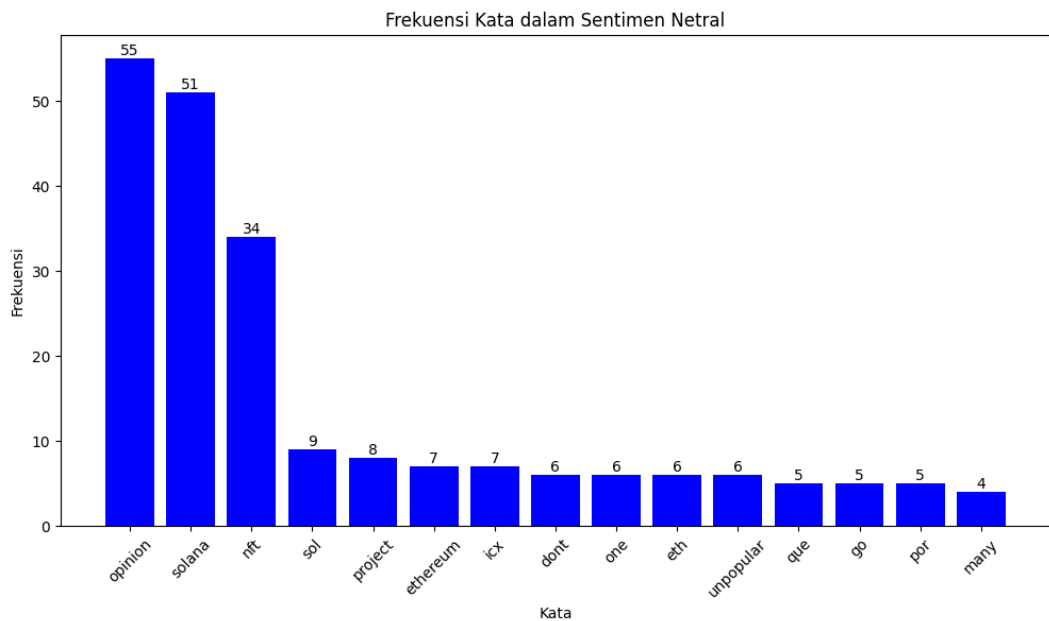
Gambar 8. Frekuensi kemunculan terbanyak kata sentimen negatif

Hasil gambar 8 menunjukkan bahwa frekuensi kemunculan kata dalam sentimen negatif dapat divisualisasikan dalam diagram batang. Dari diagram tersebut, terlihat bahwa kata "opinion" muncul sebanyak 74 kali, sedangkan kata yang mengandung "Solana" muncul sebanyak 70 kali



Gambar 9. Hasil Wordcloud Tweet Netral

Pada gambar 9 dapat dideskripsikan kata "opinion" pada *wordcloud* netral, artinya bahwa ada sentimen netral yang terkait dengan kata "nft solana" dalam *tweet-tweet* tersebut. Seperti membahas tentang NFT Solana atau *project-project* NFT solana kedepannya.



Gambar 10. Frekuensi kemunculan terbanyak kata sentimen netral

Hasil gambar 10 menunjukkan bahwa frekuensi kemunculan kata dalam sentimen netral dapat divisualisasikan dalam diagram batang. Dari diagram tersebut, terlihat bahwa kata "opinion" muncul sebanyak 55 kali, sedangkan kata yang mengandung "51" muncul sebanyak 288 kali

#### Pembahasan

Penelitian ini menerapkan pendekatan komprehensif dalam menganalisis sentimen publik terhadap NFT Solana di media sosial X (Twitter). Metodologi yang digunakan mencakup beberapa tahapan kunci, dimulai dari pengumpulan data melalui teknik crawling hingga visualisasi hasil menggunakan word cloud. Setiap tahapan dalam penelitian ini didesain dengan cermat untuk memastikan analisis yang akurat dan mendalam terhadap sentimen publik mengenai NFT Solana.

Tahap awal penelitian melibatkan pengumpulan data dari platform X menggunakan metode crawling. Pendekatan ini sudah umum digunakan dalam analisis sentimen media sosial dan telah terbukti efektif dalam menangkap opini publik secara real-time. Seperti yang ditunjukkan oleh (Azhari, 2014) dalam penelitian mereka, penggunaan data media sosial untuk analisis sentimen memungkinkan peneliti untuk mendapatkan wawasan yang kaya dan up-to-date tentang persepsi publik terhadap suatu topik. Dalam konteks NFT Solana, pendekatan ini sangat relevan mengingat sifat dinamis dan cepat berubahnya sentimen dalam industri cryptocurrency dan NFT.

Setelah data terkumpul, dilakukan serangkaian tahapan preprocessing yang meliputi case folding, tokenizing, stopwords removal, dan stemming. Tahapan ini sangat krusial dalam mempersiapkan data untuk analisis lebih lanjut Kao & Poteet (2007) dalam buku

mereka "*Natural Language Processing and Text Mining*" menekankan bahwa preprocessing yang tepat adalah fondasi untuk analisis teks yang akurat. Mereka menjelaskan bahwa tahapan ini tidak hanya membersihkan data, tetapi juga membantu dalam mengurangi noise dan meningkatkan kualitas input untuk tahapan analisis selanjutnya. Dalam konteks NFT Solana, preprocessing yang teliti membantu dalam mengisolasi kata-kata dan frasa yang paling relevan dengan sentimen terhadap NFT, memungkinkan analisis yang lebih fokus dan akurat.

Inovasi menarik dalam penelitian ini adalah penggunaan *VADER (Valence Aware Dictionary and Sentiment Reasoner)* untuk pelabelan data secara otomatis. VADER, yang diperkenalkan oleh Hutto & Gilbert, (2014) merupakan alat analisis sentimen yang dikalibrasi khusus untuk sentimen yang diekspresikan dalam media sosial. Penggunaan VADER dalam penelitian ini menunjukkan kesadaran akan kompleksitas sentimen dalam konteks media sosial, di mana penggunaan bahasa informal, sarkasme, dan konteks yang cepat berubah sering kali mempersulit analisis sentimen tradisional.

Aspek penting dari penggunaan VADER dalam penelitian ini adalah penentuan ambang batas untuk klasifikasi sentimen. Penelitian ini mengadopsi pendekatan yang lebih nuansir dibandingkan dengan penggunaan VADER yang umum. Umumnya, VADER mengklasifikasikan sentimen sebagai positif jika skor compound lebih besar dari 0, dan negatif jika kurang dari 0. Namun, penelitian ini menggunakan ambang batas yang lebih ketat: positif untuk skor di atas 0.05, negatif untuk skor di bawah -0.05, dan netral untuk skor di antara kedua nilai tersebut.

Penentuan ambang batas ini sejalan dengan rekomendasi dari beberapa penelitian terkini. Misalnya, Elbagir & Yang, (2019) dalam penelitian mereka "*Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment*" menyarankan penggunaan ambang batas yang lebih ketat untuk meningkatkan akurasi klasifikasi, terutama dalam konteks yang kompleks seperti diskusi tentang *cryptocurrency* dan NFT. Mereka menemukan bahwa penggunaan ambang batas 0.05 dan -0.05 membantu dalam mengurangi kesalahan klasifikasi untuk tweet dengan sentimen yang lemah atau ambigu.

Lebih lanjut Liu (2020) dalam bukunya "*Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*" mendiskusikan pentingnya kalibrasi ambang batas dalam analisis sentimen, terutama ketika berhadapan dengan domain spesifik seperti NFT. Liu menekankan bahwa penggunaan ambang batas yang lebih ketat dapat membantu dalam mengidentifikasi sentimen yang benar-benar signifikan, mengurangi noise dari tweet dengan sentimen yang tidak jelas atau netral.

Selanjutnya, penelitian menggunakan metode *TF-IDF (Term Frequency-Inverse Document Frequency)* untuk pembobotan kata. Pilihan ini didukung oleh penelitian Trstenjak dkk., (2014) yang mendemonstrasikan efektivitas TF-IDF dalam klasifikasi teks. Penggunaan TF-IDF membantu dalam mengidentifikasi kata-kata yang paling signifikan dalam dataset, meningkatkan kualitas input untuk proses klasifikasi. Dalam konteks NFT Solana, TF-IDF memungkinkan peneliti untuk mengidentifikasi istilah-istilah kunci yang sering dikaitkan dengan sentimen positif atau negatif terhadap NFT Solana.

Dalam tahap klasifikasi, penelitian membandingkan dua algoritma populer: *Random Forest* dan *Support Vector Machine (SVM)*. Menariknya, kedua algoritma menunjukkan performa yang identik dengan akurasi 69.56%. Hasil ini agak berbeda dengan temuan Willy dkk. (2021) yang umumnya menemukan SVM lebih unggul dalam tugas klasifikasi sentimen. Perbedaan ini membuka peluang untuk diskusi lebih lanjut tentang karakteristik spesifik dataset NFT Solana dan bagaimana hal ini mungkin mempengaruhi kinerja algoritma. Kemungkinan besar, kompleksitas dan keunikan diskusi seputar NFT Solana mempengaruhi pola sentimen yang mungkin tidak sepenuhnya tertangkap oleh model klasifikasi tradisional.

Aspek visualisasi penelitian menggunakan *word cloud*, memberikan representasi visual yang intuitif dari kata-kata yang sering muncul dalam dataset. Pendekatan ini, seperti yang dibahas oleh Heimerl dkk. (2014) sangat berguna dalam membantu interpretasi data teks yang kompleks, memungkinkan peneliti dan pembaca untuk dengan cepat mengidentifikasi tema-tema dominan dalam dataset. Dalam konteks NFT Solana, visualisasi ini dapat membantu mengidentifikasi tren, isu-isu yang sering dibicarakan, dan kata kunci yang mungkin mempengaruhi sentimen publik.

Secara keseluruhan, penelitian ini memberikan wawasan berharga tentang sentimen publik terhadap NFT Solana di media sosial. Pendekatan yang digunakan sejalan dengan tren penelitian yang lebih luas dalam analisis sentimen cryptocurrency dan blockchain, seperti yang dibahas oleh Lubis & Nasution (2023) dalam studi mereka tentang sentimen *Bitcoin* di Twitter. Mereka menekankan pentingnya analisis sentimen real-time dalam memahami dinamika pasar *cryptocurrency*, yang juga relevan untuk NFT Solana.

Meskipun akurasi yang dicapai (69.56%) menunjukkan ada ruang untuk peningkatan, penelitian ini telah meletakkan dasar yang kuat untuk studi lebih lanjut dalam domain ini. Untuk pengembangan ke depan, penelitian dapat mempertimbangkan penggunaan metode deep learning seperti LSTM atau transformers, yang telah menunjukkan performa yang menjanjikan dalam tugas-tugas NLP kompleks. Devlin dkk. (2018) dalam penelitian mereka tentang BERT (*Bidirectional Encoder Representations from Transformers*)

menunjukkan bagaimana model berbasis *transformer* dapat secara signifikan meningkatkan akurasi dalam berbagai tugas NLP, termasuk analisis sentimen.

Selain itu, penambahan fitur kontekstual dan analisis temporal mungkin dapat meningkatkan pemahaman tentang dinamika sentimen NFT Solana dari waktu ke waktu. Seperti yang diusulkan oleh Saif dkk. (2016) dalam penelitian mereka tentang analisis sentimen kontekstual, mempertimbangkan konteks dan evolusi temporal dari sentimen dapat memberikan wawasan yang lebih mendalam dan akurat tentang opini publik.

Kesimpulannya, penelitian ini tidak hanya memberikan wawasan berharga tentang sentimen publik terhadap NFT Solana, tetapi juga mendemonstrasikan pendekatan metodologis yang kuat dalam analisis sentimen media sosial. Dengan mengadopsi teknik-teknik canggih seperti VADER dengan ambang batas yang disesuaikan, penelitian ini membuka jalan untuk pemahaman yang lebih nuansir tentang sentimen dalam domain yang kompleks seperti NFT dan *cryptocurrency*. Meskipun masih ada ruang untuk peningkatan, terutama dalam hal akurasi klasifikasi, penelitian ini memberikan kontribusi signifikan dalam memahami dinamika sentimen publik terhadap teknologi blockchain dan aset digital.

## SIMPULAN

Berdasarkan hasil dan pembahasan mengenai analisis sentimen menggunakan *Support Vector Machine* (SVM) dan *Random Forest*, dapat disimpulkan bahwa:

1. Algoritma *Random Forest* dan *Support Vector Machine* (SVM) telah berhasil diterapkan untuk klasifikasi sentimen tweet terkait NFT pada *Blockchain* Solana. Proses *preprocessing* yang digunakan meliputi *case folding*, *tokenizing*, *stopwords removal*, *stemming*, serta pembobotan kata dengan metode *TF-IDF*. Kedua algoritma ini diuji pada *dataset* yang sama, setelah melalui tahapan *preprocessing* dan ekstraksi fitur yang identik.
2. Hasil pengujian menunjukkan bahwa kedua algoritma memiliki performa yang identik, dengan akurasi 69.56%, presisi 56.46%, *recall* 69.56%, dan *F1-score* 57.41%. Hasil ini menunjukkan bahwa baik *Random Forest* maupun SVM mampu menganalisis sentimen dengan kinerja yang sebanding, meskipun penelitian sebelumnya seringkali menunjukkan keunggulan SVM dalam tugas klasifikasi sentimen.

Secara keseluruhan, kedua model menunjukkan performa yang hampir serupa dengan akurasi yang dapat diterima, namun masih ada ruang untuk meningkatkan akurasi dan ketepatan prediksi, terutama dalam mengatasi kesalahan *False Negative* yang signifikan.

## DAFTAR PUSTAKA

- Anggreni, L. S., Nugroho, R. A., Luthfi, H. S., Kresna, I. M., & Santoso, T. B. (2019). Penggunaan kata umpatan di Twitter berdasarkan gender di pilkada Sumatera Utara 2018. *Jurnal Kajian Komunikasi*, 7(1), 121. <https://doi.org/10.24198/jkk.v7i1.18447>
- Azhari, F. (2014). Skripsi. Universitas Islam Negeri Syarif hidayatullah. *Hubungan Kadar Timbal dalam Urin dan Karakteristik Individu dengan Kejadian Anemia pada Pedagang Wanita di Terminal Kampung Rambutan Jakarta Timur*.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Elbagir, S., & Yang, J. (2019). Twitter sentiment analysis using natural language toolkit and VADER sentiment. *Proceedings of the international multiconference of engineers and computer scientists*, 122(16).
- Heimerl, F., Lohmann, S., Lange, S., & Ertl, T. (2014). Word Cloud Explorer: Text Analytics Based on Word Clouds. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2014.231>
- Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the international AAAI conference on web and social media*, 8(1), 216–225.
- Jannah, S. F., Astuti, R., & Basysyar, F. M. (2024). Implementasi Algoritma Random Forest Pada Aplikasi Picsart Berdasarkan Respon Pengguna. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 274–283.
- Kao, A., & Poteet, S. (2007). *Natural Language Processing and Text Mining*. <https://doi.org/10.1007/978-1-84628-754-1>
- Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 7(1), 32. <https://doi.org/10.24014/coreit.v7i1.12945>
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge university press.
- Lubis, A. R., & Nasution, M. K. M. (2023). Twitter Data Analysis and Text Normalization in Collecting Standard Word. *Journal of Applied Engineering and Technological Science (JAETS)*, 4(2), 855–863.
- Nakamoto, S. (2009). *Bitcoin: A Peer-to-Peer Electronic Cash System*.
- Rahim, A. M. A., Pratiwi, I. Y. R., & Fikri, M. A. (2023). Klasifikasi Penyakit Jantung

Menggunakan Metode Synthetic Minority Over-Sampling Technique Dan Random Forest Classifier. *Indonesian Journal of Computer Science*, 12(5).

Ramadhanty, D. R. (2021). *Implementasi Algoritma Support Vector Machine Pada Analisis Sentimen Data Twitter*. 1–83.

Saif, H., He, Y., Fernandez, M., & Alani, H. (2016). Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management*, 52(1), 5–19.

Willy, W., Rini, dian P., & Samsuryadi, S. (2021). Perbandingan Algoritma Random Forest Classifier, Support Vector Machine dan Logistic Regression Clasifier Pada Masalah High Dimension (Studi Kasus: Klasifikasi Fake News). *Jurnal Media Informatika Budidarma*, 5, 1720. <https://doi.org/10.30865/mib.v5i4.3177>

Yuda Lesmana, R., & Andarsyah, R. (2022). Model Klasifikasi Multinomial Naïve Bayes Untuk Analisis Sentiment Terkait Non-Fungible Token. *Jurnal Teknik Informatika*, 14(3), 135–139.