

IMPLEMENTASI WORD EMBEDDING GLOVE PADA SUPPORT VECTOR MACHINE DENGAN PARTICLE SWARM OPTIMIZATION UNTUK ANALISIS SENTIMEN

Putu Ayu Indah Nazwa Usia, Isran K. Hasan, Siti Nurwardia Abdussamad*

Statistika, Universitas Negeri Gorontalo, Indonesia

sitinurwardia@ung.ac.id

Informasi Artikel	Abstrak
<p>Submitted: April 1, 2026 Revised: April 24, 2026 Accepted: May 5, 2026</p> <p>Kata Kunci Analisis Sentimen; <i>Particle Swarm Optimization</i>; <i>Support Vector Machine</i>; <i>Word Embedding GloVe</i>.</p>	<p>Tujuan: Penelitian ini menerapkan <i>word embedding</i> GloVe pada <i>Support Vector Machine</i> yang dioptimalkan memanfaatkan <i>Particle Swarm Optimization</i> untuk mengevaluasi efektivitas model klasifikasi analisis sentimen ulasan pengguna Indodax di platform media sosial X.</p> <p>Metode: Pada studi ini, data teks diklasifikasikan memanfaatkan pendekatan metodologi kuantitatif. Data berupa 877 cuitan pengguna Indodax di media sosial X yang dikumpulkan melalui teknik scraping menggunakan Python. Tahapan analisis meliputi pengumpulan data, <i>pre-processing</i> data, pelabelan sentimen, <i>word embedding</i> GloVe, pembagian data, <i>Particle Swarm Optimization</i> digunakan guna optimasi parameter, serta <i>Support Vector Machine</i> dimanfaatkan untuk klasifikasi.</p> <p>Hasil: Hasil penelitian menunjukkan bahwa analisis sentiment ulasan pengguna Indodax di media sosial X menghasilkan distribusi sentimen yang relatif seimbang, yaitu sebesar 49,7% sentimen positif dan 50,3% sentimen negatif. Selain itu, melalui skor akurasi sejumlah 82%, presisi 85%, <i>recall</i> 82%, serta <i>f1-score</i> 84%, hasil penilaian model menunjukkan bahwa penerapan <i>word embedding</i> GloVe pada SVM yang dioptimalkan dengan <i>Particle Swarm Optimization</i> dapat menghasilkan performa klasifikasi yang baik.</p> <p>Simpulan: Dengan menggabungkan GloVe <i>word embedding</i> dengan <i>Support vector Machine</i> yang dioptimasi menggunakan PSO, penelitian ini memberikan kontribusi dalam pengembangan metode analisis sentimen berbasis <i>machine learning</i>, serta memberikan implikasi praktis bagi pengelola platform dan pemerintah dalam memahami persepsi pengguna terhadap layanan investasi digital.</p>
<p>Keywords Sentiment Analysis; <i>Particle Swarm Optimization</i>; <i>Support Vector Machine</i>; <i>Word Embedding GloVe</i>.</p>	<p>Abstrack</p> <p>Purpose: This study applies GloVe word embeddings to a Support Vector Machine optimized using Particle Swarm Optimization to evaluate the effectiveness of a sentiment analysis classification model for Indodax user reviews on the social media platform X.</p> <p>Method: This study uses text data and a quantitative methodology with a classification design. The dataset consists of 877 tweets from Indodax users on the social media platform X, collected via web scraping using Python. The analysis process includes data preprocessing, sentiment labeling, GloVe word embedding, data splitting, PSO is used for parameter optimization, and Support Vector Machines are used for classification.</p> <p>Results: The results of the study indicate that sentiment analysis of Indodax user reviews on the social media platform X yielded a relatively balanced sentiment distribution, with 49.7% positive sentiment and 50.3% negative sentiment. Additionally, the model evaluation results indicate that the implementation of GloVe word embeddings on a Support Vector Machine optimized using Particle Swarm Optimization achieves good classification performance, with an accuracy of 82%, precision of 85%, recall of 82%, and an F1-score of 84%.</p> <p>Contribution: This study contributes to the development of a machine learning-based sentiment analysis method through the combination of</p>

PENDAHULUAN

Perubahan perilaku investasi masyarakat mengalami perkembangan yang cukup signifikan seiring dengan pesatnya pemanfaatan teknologi digital dalam sektor (Yusril dkk., 2025). Transformasi ini ditandai dengan hadirnya *cryptocurrency* sebagai aset digital berbasis teknologi blockchain yang menawarkan efisiensi dan transparansi dalam transaksi keuangan (Sudarwanti dkk., 2025). Kementerian Perdagangan Indonesia melaporkan bahwa hingga September 2024, terdapat lebih dari 21,27 juta investor mata uang kripto di Indonesia, meningkat dari 11,2 juta pada tahun 2021 (Faqih, 2025). Kondisi ini menunjukkan bahwa aset kripto telah menjadi bagian penting dalam strategi investasi masyarakat modern (Firgayanda dan Jumhur, 2022).

Peningkatan minat masyarakat terhadap aset kripto mendorong munculnya berbagai platform perdagangan digital, salah satunya Indonesia Digital Asset Exchange (Indodax) merupakan platform kripto tertua dan terbesar di Indonesia (Maulana dkk., 2025). Data menunjukkan bahwa pada April nilai transaksi pada platform Indodax tercatat sebesar Rp15,24 triliun, yang berkontribusi pada kisaran 42,83% terhadap keseluruhan transaksi kripto nasional sehingga menunjukkan bahwa Indodax memiliki posisi dominan dalam perdagangan kripto di Indonesia (Setiawan, 2025). Tingginya aktivitas transaksi tersebut mencerminkan intensitas interaksi pengguna, sehingga menjadikan Indodax relevan untuk dianalisis melalui ulasan pengguna. Dalam studi ini, data ulasan diperoleh melalui media sosial X.

X merupakan media sosial berbasis *microblogging* yang memungkinkan pengguna menyampaikan opini secara singkat, padat, dan cepat tersebar melalui fitur *reply*, *retweet*, dan *hashtag* (Rangga dkk., 2019). Karakteristik ini membuat ulasan pengguna di X muncul secara spontan dan real-time, sehingga mencerminkan respons alami serta menjadi sumber data yang relevan untuk memahami persepsi pengguna melalui analisis sentimen, yaitu pengklasifikasian ulasan ke dalam kategori positif atau negatif (Arsi & Waluyo, 2021). Penerapan analisis sentimen pada ulasan pengguna Indodax penting dilakukan untuk memahami persepsi publik serta memberikan masukan bagi pengelola dan pemerintah dalam meningkatkan ekosistem investasi kripto di Indonesia. Namun, proses ini menghadapi tantangan karena data bersifat tidak terstruktur, berdimensi tinggi, dan mengandung noise. Untuk mengatasi tantangan tersebut, berbagai metode *machine learning* dapat diterapkan dalam proses klasifikasi data. Salah satu metode yang dikenal efektif dan mampu menghasilkan model klasifikasi yang akurat meskipun dihadapkan pada data berdimensi tinggi adalah *Support Vector Machine* (SVM) (Sari dkk., 2025).

Metode SVM ialah bagian dari algoritma *machine learning* berbasis *supervised learning* yang banyak dimanfaatkan guna mengelompokkan data teks (Parapat dkk., 2018). SVM bekerja dengan memaksimalkan margin antar kelas dan mampu menangani data berdimensi tinggi melalui penggunaan fungsi kernel (Maulidah dkk., 2021). Beberapa penelitian menunjukkan bahwa SVM menghasilkan akurasi yang lebih tinggi dibandingkan metode lain seperti pada penelitian yang dilakukan oleh Yusupa & Tarigan (2024) yang

membandingkan metode SVM, *Random Forest*, dan *Naive Bayes* dan penelitian oleh Munawaroh (2022) yang membandingkan SVM, *Naive Bayes*, dan KNN. Berdasarkan hasil penelitian tersebut, terbukti bahwa SVM memiliki keunggulan dalam menghasilkan klasifikasi yang lebih akurat pada analisis sentimen berbasis teks. Namun, efektivitas SVM sangat bergantung pada kualitas fitur yang digunakan, karena data teks perlu direpresentasikan dalam bentuk numerik sebelum proses klasifikasi dilakukan (Tirta dkk., 2025). Oleh karena itu, penggunaan *word embedding* menjadi penting untuk menghasilkan representasi fitur yang lebih informatif.

Pada *machine learning*, *word embedding* ialah cara yang dimanfaatkan guna merepresentasikan sejumlah kata sebagai vektor numerik (Sabrila dkk., 2021). Dalam pengembangan representasi kata, *Global Vectors For Word Representation* (GloVe) merupakan metode *word embedding* yang memanfaatkan statistik global dari seluruh korpus. Berbeda dengan metode *word embedding* lain yang hanya memanfaatkan konteks lokal, GloVe menghasilkan representasi kata yang lebih konsisten dan bermakna dengan menangkap tautan semantik antar kata secara global (Nurdin dkk., 2020). Efektivitas tersebut dibuktikan oleh penelitian Anjani dan Irmanda (2024) yang menunjukkan bahwa GloVe menghasilkan kinerja terbaik dibandingkan Word2Vec dan FastText. Selain penggunaan *word embedding*, teknik optimasi juga memiliki peran penting dalam meningkatkan kinerja SVM karena optimasi membantu menemukan kombinasi hyperparameter tuning, sehingga model dapat membangun batas keputusan yang lebih tepat.

Salah satu teknik optimasi yang efisien untuk memilih *hyperparameter* guna meningkatkan performa klasifikasi SVM adalah *Particle Swarm Optimization* (PSO) (Achyani, 2018). Setiap partikel merepresentasikan kandidat solusi yang memperbarui posisi dan kecepatan berdasarkan pengalaman individu maupun kelompok, sehingga PSO dapat menentukan *hyperparameter tuning* secara efisien (Ginabila dkk., 2024). Seperti pada penelitian sebelumnya oleh Putra dkk., (2024), menunjukkan bahwa PSO mampu meningkatkan akurasi dibandingkan *Grid Search*, sehingga efektif dalam menentukan *hyperparameter* untuk klasifikasi. Maka dari itu, studi ini menekankan pada penerapan *word embedding* GloVe dengan optimasi PSO dalam meningkatkan kinerja SVM pada analisis sentimen ulasan pengguna Indodax di media sosial X. Lebih dari itu, harapannya studi ini bisa memvisualisasikan terkait penerapan teknologi pemrosesan bahasa dalam memahami kecenderungan sentimen ulasan pengguna Indodax.

METODE PENELITIAN

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain klasifikasi berbasis *machine learning* untuk mengklasifikasikan sentimen teks ke dalam kategori positif dan negatif. Untuk mendapatkan performa model sebaik mungkin, desain penelitian ini berfokus pada pengembangan model klasifikasi dengan mengimplementasikan *word embedding* GloVe pada SVM yang dioptimalkan menggunakan PSO.

Subjek/Partisipan

Subjek dalam penelitian ini berupa data teks ulasan pengguna Indodax di media sosial X yang berjumlah 877 data. Data dikumpulkan menggunakan teknik scraping melalui bahasa pemrograman Python dengan package *tweet harvest* berdasarkan kata kunci “indodax” dan

“INDODAX” pada rentang waktu 1 Januari 2024 hingga 27 Oktober 2025. Karakteristik data berupa teks tidak terstruktur yang mengandung opini pengguna terkait layanan platform Indodax. Teknik pengambilan sampel yang digunakan adalah sampling jenuh yaitu seluruh data yang memenuhi syarat penelitian digunakan sebagai sampel. Penelitian ini menggunakan cakupan unit analisis pada skala nasional karena data diperoleh dari pengguna media sosial X di Indonesia dan tidak terikat pada lokasi geografis tertentu.

Analisis Data

Penelitian ini dirancang menggunakan pendekatan pembelajaran mesin berbasis *machine learning* untuk mengklasifikasikan sentimen teks cuitan. Proses analisis dilakukan secara bertahap, dimulai dari pengumpulan data, *pre-processing*, pemberian label, implementasi *word embedding* GloVe, pembagian data, optimasi parameter menggunakan PSO, klasifikasi dengan SVM serta evaluasi kinerja model.

Pre-Processing Data

Pada tahap awal, data cuitan diproses melalui *pre-processing* untuk menyiapkan data sebelum memasuki tahap klasifikasi. Melalui tahap ini, data teks dibersihkan dan dinormalisasi sehingga dapat diolah dengan lebih baik pada proses analisis. Dalam penelitian ini, *pre-processing* dilakukan melalui lima tahapan, yaitu :

1. *Case folding*, yaitu tahap menormalisasi teks lewat pengubahan seluruh huruf kapital hingga berbentuk huruf kecil untuk mempermudah keterbacaan data teks serta keseragaman di dalam dataset (Dina dkk., 2025).
2. *Cleaning*, yaitu tahapan untuk menghapus karakter dari kata yang tidak dibutuhkan atau diperlukan, seperti tanda baca, hashtag, URL, simbol serta username, dengan tujuan mengurangi noise dalam data (Dina dkk., 2025).
3. *Tokenizing*, yaitu tahapan pembagian teks menjadi sejumlah unit kecil berupa token, sehingga teks yang semula panjang hingga berbentuk struktur data sederhana yang bisa ditangani model dengan mudah (Ama dkk., 2022).
4. *Normalization*, yaitu tahap memperbaiki kata tidak baku dengan cara mengubah ungkapan atau bentuk bahasa informal yang sering muncul dalam teks menjadi kata yang baku (Safitri dkk., 2023).
5. *Stemming*, yakni tahap pengubahan kata-kata dalam teks hingga berbentuk dasar dengan menghilangkan akhiran atau imbuhan untuk mengurangi variasi kata serupa dalam analisis (Ramadhan dkk., 2022).

Pelabelan Data

Pelabelan data dilakukan pada tahap ini dengan mengategorikan data ke sentimen positif dan negatif, yang masing-masing secara berurutan diberi label 1 dan -1. Tahap pelabelan akan dilakukan secara manual oleh penulis dan divalidasi oleh salah satu Dosen Bahasa Indonesia selaku ahli Bahasa Indonesia.

Word Embedding GloVe

Global Vectors for Word Representation (GloVe) merupakan algoritma *word embedding* yang menggunakan data statistik global dari koneksi kemunculan bersama kata pada korpus teks untuk menghasilkan representasi semantic kata. Model ini menggunakan pendekatan global matrix factorization, yaitu pemfaktoran matriks yang mengekstraksi pola makna dari hubungan antar kata dalam korpus besar dan mengubahnya menjadi representasi vektor yang lebih sederhana serta bermakna secara semantik (Nurdin dkk., 2020).

Pembagian Data

Dataset dibagi menjadi dua subset utamam yakni 20% digunakan untuk pengujian dan 80% digunakan untuk pelatihan. Data pengujian, dimanfaatkan guna menilai kinerja model yang telah dilaksanakan pelatihan sebelumnya, sedangkan data pelatihan digunakan dalam proses pembelajaran model. Proporsi pembagian 80:20 merupakan rasio paling seimbang dalam meningkatkan hasil kinerja klasifikasi SVM (Vincent dkk., 2023).

Particle Swarm Optimization

Tahap optimasi dilakukan menggunakan Particle Swarm Optimization (PSO). Pada tahap ini, PSO digunakan untuk mencari kombinasi *hyperparameter* terbaik pada model *Support Vector Machine*, yaitu parameter C dan γ pada kernel *Radial Basis Function* (RBF). Proses optimasi dilakukan secara iteratif melalui pembaruan kecepatan dan posisi partikel hingga diperoleh konfigurasi *hyperparameter* yang optimal dalam meningkatkan kinerja klasifikasi pada data latih. Pembaruan kecepatan dan posisi setiap partikel dilakukan berdasarkan persamaan berikut (Suprianto dkk., 2024) :

1. Pembaruan Kecepatan

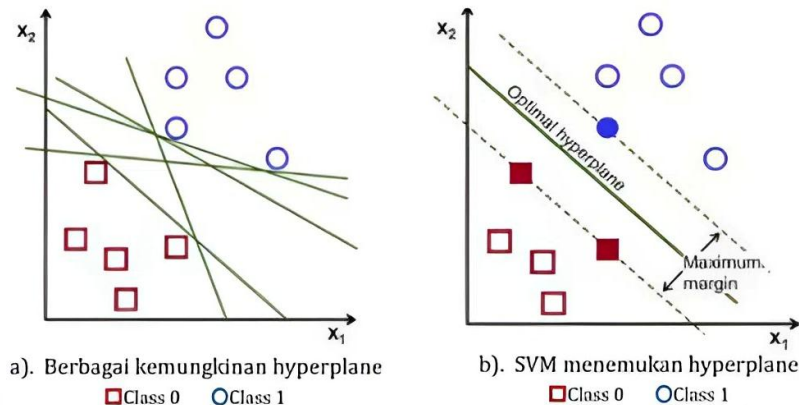
$$V_{id}^{k+1} = w \times V_{id}^k + c_1 \times rand_1 \times (P_{id} - X_{id}) + c_2 \times rand_2 \times (G_{id} - X_{id})$$

2. Pembaruan Posisi

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1}$$

Support Vector Machine

Support Vector Machine (SVM) ialah algoritma *machine learning* dan acapkali dimanfaatkan guna mengatasi permasalahan dari pengkategorian (Saputra dkk., 2025). Algoritma tersebut bekerja dengan mencari *hyperplane* optimal sebagai batas pemisah antar kelas berdasarkan fitur yang dimiliki. Margin adalah jarak di antara titik-titik data atas tiap-tiap kelas melalui garis pemisah (*hyperplane*). Jarak tersebut dibuat sebesar mungkin agar pemisah dapat memisahkan data dengan lebih jelas. Semakin besar margin, semakin baik kemampuan model dalam membedakan kedua kelas dengan benar (Drajana, 2017).



Gambar 1. SVM berupaya mengidentifikasi *hyperplane* terbaik

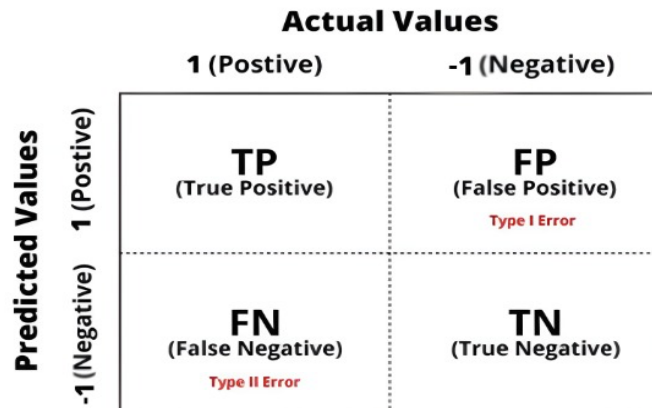
Prinsip utama SVM ialah memaksimalkan margin antar kelas menggunakan fungsi pemisah linear agar model dapat membedakan dua kelompok data secara optimal. Akan tetapi, tidak seluruh data bisa diseparasikan dengan linear di ruang vector aslinya, sehingga SVM memanfaatkan fungsi kernel guna memetakan data ke ruang dengan dimensi yang lebih tinggi sehingga *hyperplane* dapat terbentuk. Pada studi ini dimanfaatkan fungsi kernel *Radial Basis Function* (RBF) karena kemampuannya dalam meningkatkan akurasi pada proses klasifikasi sentimen menggunakan algoritma SVM. Persamaan kernel RBF ditunjukkan sebagai berikut:

$$K(x_n, x_i) = \exp(-\gamma|x_n - x_i|^2)$$

Confusion Matrix

Confusion matrix ialah bagian dari teknik pengujian yang dimanfaatkan guna melaksanakan pengukuran akan kinerja suatu metode klasifikasi. Teknik ini menyajikan perbandingan secara sistematis antara hasil prediksi model dengan hasil klasifikasi sebenarnya, sehingga dapat diketahui sejauh mana model mampu mengidentifikasi kelas secara tepat (Valerian dkk., 2025). Berikut komponen yang ada dalam confusion matrix ialah:

1. True Positive (TP), yakni nilai positif yang diprediksi positif.
2. False Positive (FP), yakni nilai negatif yang diprediksi positif.
3. False Negative (FN), yakni nilai positif diprediksi negatif.
4. True Negative (TN), yakni nilai negatif yang diprediksi negatif



Gambar 2. Confusion Matrix

Melalui komponen tersebut, rumus berikut dapat digunakan untuk menghasilkan beberapa metrik evaluasi kinerja model dengan memanfaatkan empat aspek penilaian, yaitu akurasi, presisi, *recall*, dan *F1-score* (Romadoni dkk., 2020):

1. Akurasi menunjukkan seberapa baik nilai yang diharapkan sesuai dengan nilai sebenarnya.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

2. Presisi menunjukkan tingkat ketelitian antara data yang dimintan dan hasil proyeksi model.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

3. *Recall* menunjukkan seberapa baik model tersebut mengambil kembali sebuah informasi.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

4. *F1-Score* menunjukkan metrik evaluasi yang memberikan evaluasi kinerja model yang lebih menyeluruh dengan menggabungkan nilai presisi dan *recall*. Nilai presisi dan *recall* dirata-ratakan untuk mendapatkan nilai *f1-score* (Yuranda dkk., 2023).

$$F1-Score = 2 \times \frac{precision \times recall}{precision + recall} \times 100\%$$

HASIL DAN PEMBAHASAN

Hasil Penelitian

Pre-processing Data

Tahapan *pre-processing* dilakukan untuk membersihkan dan menyiapkan data sebelum proses analisis. Hasil dari setiap tahapan *pre-processing* ditampilkan oleh Tabel 1.

Tabel 1. Hasil Tahapan *Pre-processing*

Proses	Input	Output
<i>Case Folding</i>	@indodax Kalo lg pada naik koinnya Pada mantenen terus aja minkocak!!	@indodax kalo lg pada naik koinnya pada mantenen terus aja minkocak !!
<i>Cleaning</i>	@indodax kalo lg pada naik koinnya pada mantenen terus aja minkocak!!	kalo lg pada naik koinnya pada mantenen terus aja min kocak
<i>Tokenizing</i>	kalo lg pada naik koinnya pada mantenen terus aja min kocak	['kalo', 'lg', 'pada', 'naik', 'koinnya', 'pada', 'mantenen', 'terus', 'aja', 'min', 'kocak']
<i>Normalization</i>	['kalo', 'lg', 'pada', 'naik', 'koinnya', 'pada', 'mantenen', 'terus', 'aja', 'min', 'kocak']	['kalau', 'lagi', 'pada', 'naik', 'koinnya', 'pada', 'mantenen', 'terus', 'saja', 'admin', 'kocak']
<i>Stemming</i>	['kalau', 'lagi', 'pada', 'naik', 'koinnya', 'pada', 'mantenen', 'terus', 'saja', 'admin', 'kocak']	kalau lagi pada naik koin pada mantenen terus saja admin kocak

Pelabelan Data

Tahapan pelabelan data dalam penelitian ini dilaksanakan dengan manual, kemudian hasil pelabelan tersebut divalidasi oleh validator ahli bahasa Indonesia. Contoh hasil validasi pelabelan oleh validator ahli disajikan oleh Tabel 2.

Tabel 2. Contoh Hasil Pelabelan dan Validasi

Cuitan	Label Awal	Label Validasi	Keterangan
Cuma indodax yang harga cryptonya berbeda jauh dengan harga pasaran pertukaran lain berbeda tetapi tidak terlalu jauh indodax terlalu banyak permainan liciknya dan banyak crypto scam pada indodax mungkin bappebti sudah dibayar untuk tutup mulut dan mata	Negatif	Negatif	Sesuai
indodax paling keren dan fiturnya sangat lengkap dan mudah dipahami proses transaksi lebih simpel dan cepat sukses selalu indodax sebagai exchange crypto terbaik di indonesia	Positif	Positif	Sesuai

Dalam penelitian ini, klasifikasi sentimen difokuskan pada kategori positif dan negatif. Berdasarkan hasil pelabelan, terdapat 4 cuitan yang termasuk dalam kategori netral dan 2 cuitan yang teridentifikasi sebagai data duplikat, sehingga tidak digunakan dalam proses analisis. Total data yang dimanfaatkan pada studi ini ditampilkan oleh Tabel 3.

Tabel 3. Data dan Label

Label	Jumlah Data
Positif	436
Negatif	441

Word Embedding GloVe

GloVe yang digunakan dalam prosedur *word embedding* pada penelitian ini adalah korpus *twitter.27B* yang dikembangkan dari data Twitter. Setiap kata direpresentasikan dalam bentuk vektor berdimensi 100. Contoh vektor kata hasil proses *word embedding* GloVe disajikan pada Tabel 4.

Tabel 4. Vektor Kata Model GloVe

Kata	Dimensi1	Dimensi2	Dimensi...	Dimensi100
enak	-0.0222	0.0990	...	-0.4874
bgtt	0.4989	1.5602	...	-0.1420
aplikasi	-0.4314	-0.3411	...	-0.1212
nyaa	0.3167	1.4619	...	-0.5275

Pembagian Data

Data dalam penelitian ini dibagi melalui rasio 80:20, yaitu 80% sebagai data latih dan 20% sebagai data uji. Proporsi data latih yang lebih besar digunakan untuk meningkatkan kemampuan model dalam mempelajari pola dan kosakata, sehingga proses klasifikasi dapat dilakukan dengan lebih optimal. Dari total 877 data, sebanyak 701 data digunakan sebagai data latih dan 176 data sebagai data uji.

Particle Swarm Optimization

Proses optimasi menggunakan PSO menghasilkan kombinasi *hyperparameter* terbaik pada model SVM. Proses optimasi dilakukan dengan jumlah partikel sebanyak 10 dan 30 iterasi, serta menggunakan skema *5-fold cross validation* pada data latih untuk mengevaluasi setiap partikel. Hasil optimasi disajikan pada Tabel 5.

Tabel 5. Hasil Optimasi PSO

Parameter	Nilai
C	46.852500
γ	0.878993
Akurasi	0.827396

Klasifikasi dan Evaluasi

Parameter optimal C dan γ yang diperoleh dari proses optimasi digunakan dalam pembentukan model SVM pada tahap klasifikasi. Hasil klasifikasi disajikan dalam bentuk *confusion matrix* sebagai dasar evaluasi kinerja model, yang merepresentasikan hubungan antara prediksi model dengan label aslinya. Confusion matrix dari hasil klasifikasi ditampilkan dalam tabel 6.

Tabel 6. Confusion Matrix SVM

Actual	Prediction	
	Positif	Negatif
Positif	65	14
Negatif	17	80

Berdasarkan Tabel 6, model SVM dengan kernel RBF mampu mengklasifikasikan 65 data sebagai *True Positive* dan 80 data sebagai *True Negative*, dengan 16 *False Negative* dan 17 *False Positive*. Jumlah kesalahan klasifikasi yang relatif kecil menunjukkan bahwa mayoritas data sudah dikelompokkan dengan tepat oleh model. Perolehan atas klasifikasi tersebut selanjutnya diproses pada tahapan evaluasi guna melaksanakan pengukuran akan kinerja model yang sudah dibentuk. Evaluasi akan kemampuan dilaksanakan memanfaatkan beberapa metrik penilaian, yakni akurasi, *recall*, *precision*, serta *F1-score*, yang disajikan dalam Gambar 3.

	precision	recall	f1-score	support
Negatif	0.85	0.82	0.84	97
Positif	0.79	0.82	0.81	79
accuracy			0.82	176
macro avg	0.82	0.82	0.82	176
weighted avg	0.82	0.82	0.82	176

Gambar 3. Hasil Evaluasi Klasifikasi SVM

Gambar 3. Menyajikan hasil evaluasi kinerja model SVM yang menggunakan *word embedding* GloVe dengan optimasi PSO. Model menghasilkan nilai akurasi sebesar 82%. Pada kelas sentimen negatif diperoleh nilai precision sebesar 85%, recall sebesar 82%, dan F1-score sebesar 84%, sedangkan pada kelas sentimen positif diperoleh nilai precision sebesar 79%, recall sebesar 82%, dan F1-score sebesar 81%.

Pembahasan

Perolehan atas analisis sentimen ulasan pengguna Indodax di media sosial X menunjukkan bahwa 49,7% ulasan tergolong dalam sentimen positif, sedangkan 50,3% lainnya termasuk dalam sentimen negatif. Hal ini menunjukkan bahwa persepsi pengguna terhadap Indodax terbagi secara relatif seimbang, dengan sentimen negatif sedikit lebih dominan dibandingkan sentimen positif. Dominasi sentimen negatif menunjukkan bahwa sejumlah pengguna X masih menyampaikan kritik terhadap layanan Indodax, terutama yang berkaitan dengan gangguan sistem, fluktuasi harga aset kripto, serta pengalaman transaksi yang dianggap kurang optimal. Temuan ini memberikan implikasi praktis bahwa pihak pengelola platform perlu meningkatkan stabilitas sistem dan kualitas layanan untuk menjaga kepercayaan pengguna. Namun demikian, hampir setengah dari total ulasan menunjukkan sentimen positif yang mencerminkan adanya kepercayaan pengguna terhadap aspek keamanan, kualitas layanan, serta reputasi platform dalam mendukung aktivitas perdagangan aset kripto.

Hasil evaluasi kinerja model menunjukkan bahwa penerapan word embedding *Global Vectors for Word Representation* (GloVe) pada SVM yang dioptimalkan menggunakan PSO menciptakan performa klasifikasi yang baik, dengan nilai akurasi sebesar 82%, presisi 85%, recall 82%, serta F1-score 84%. Hasil ini menunjukkan bahwa kombinasi metode *word embedding* GloVe dengan optimasi PSO mampu menghasilkan kinerja yang baik pada model SVM dalam mengklasifikasikan ulasan pengguna dengan efektif. Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Yusupa dan Tarigan (2024) yang menunjukkan bahwa metode SVM memiliki tingkat efektivitas yang lebih baik dibandingkan metode klasifikasi lainnya dalam analisis sentimen data media sosial. Persamaan penelitian terletak pada penggunaan metode SVM sebagai algoritma klasifikasi utama dalam analisis sentimen. Adapun perbedaannya karena mengintegrasikan *word embedding* GloVe serta optimasi parameter menggunakan PSO, sehingga menghasilkan performa model yang lebih optimal.

Penelitian ini juga konsisten dengan temuan oleh Anjani dan Irmanda (2024) yang membandingkan beberapa metode *word embedding* dan menemukan bahwa GloVe menghasilkan akurasi yang lebih tinggi dalam klasifikasi sentimen menggunakan SVM.

Kesamaan penelitian ini terletak pada penggunaan GloVe sebagai representasi fitur teks, sedangkan perbedaannya terletak pada penambahan optimasi PSO dalam penelitian ini yang tidak digunakan pada penelitian sebelumnya. Pada penelitian oleh Putra dkk. (2024) menunjukkan bahwa optimasi parameter menggunakan PSO pada model SVM mampu meningkatkan akurasi klasifikasi dibandingkan metode optimasi lainnya seperti *Grid Search*. Hal tersebut sejalan dengan hasil penelitian ini, di mana penggunaan PSO berkontribusi dalam meningkatkan performa model melalui pencarian parameter optimal secara lebih efisien.

Secara ilmiah, peningkatan kinerja model dalam penelitian ini dapat dijelaskan melalui kemampuan GloVe dalam merepresentasikan hubungan semantik antar kata dalam bentuk vektor yang lebih informatif, sehingga mendukung SVM dalam membedakan pola sentimen secara lebih akurat. Selain itu, PSO berperan dalam mengoptimalkan parameter SVM secara global, sehingga mampu menghindari solusi lokal dan menghasilkan model dengan performa yang lebih optimal. Meskipun demikian, penelitian ini memiliki keterbatasan pada jumlah data yang relatif terbatas, sehingga generalisasi hasil terhadap keseluruhan opini pengguna masih terbatas.

Implikasi

Studi ini dapat berkontribusi bagi pengembangan ilmu bagi peneliti yang akan melaksanakan studi di sektor *machine learning* dan analisis sentimen, khususnya dalam penerapan *word embedding* GloVe pada metode *Support Vector Machine* yang dioptimalisasi melalui *Particle Swarm Optimization*. Selain itu, studi ini memperkaya referensi empiris mengenai efektivitas kombinasi metode tersebut dalam meningkatkan performa analisis sentimen ulasan pengguna pada data media sosial yang bersifat tidak terstruktur. Studi ini bisa memberi visualisasi yang berguna untuk pihak pengelola dalam memahami ulasan pengguna serta kecenderungan opini publik terhadap peningkatan kualitas layanan platform. Hasil analisis sentimen bisa dimanfaatkan menjadi landasan pertimbangan untuk pemerintah guna membentuk kebijakan yang mendukung penguatan ekosistem investasi digital di Indonesia.

Keterbatasan dan Rekomendasi Penelitian Lanjutan

Penelitian ini mempunyai keterbatasan pada jumlah data yang digunakan, sehingga representasi persepsi publik yang dihasilkan masih dapat dikembangkan lebih luas. Maka dari itu, studi berikutnya diberi masukan untuk memanfaatkan jumlah data yang lebih besar dengan periode pengambilan data yang lebih lama supaya menghasilkan representasi sentimen yang lebih aktual dan komprehensif. Selain itu, pengembangan penelitian bisa dilaksanakan lewat mengombinasikan teknik *word embedding* lain seperti Word2Vec atau FastText untuk dibandingkan dengan GloVe dalam mengevaluasi kualitas representasi fitur terhadap kinerja klasifikasi. Penelitian lanjutan juga disarankan untuk mengeksplorasi metode optimasi parameter selain PSO, seperti *Genetic Algorithm* (GA) atau *Grid Search*, untuk memperoleh kombinasi parameter yang lebih optimal dan meningkatkan performa model klasifikasi sentimen secara keseluruhan.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, tujuan penelitian untuk menerapkan *word embedding* GloVe pada *Support Vector Machine* dengan optimasi *Particle Swarm Optimization*

dalam mengevaluasi efektivitas model klasifikasi sentimen telah tercapai. Model yang dihasilkan menunjukkan kinerja yang baik melalui akurasi 82%, presisi 85%, *recall* 82%, dan *F1-score* 84%, sehingga efektif dalam mengklasifikasikan sentimen ulasan pengguna. Selain itu, hasil analisis sentimen ulasan pengguna Indodax di media sosial X menunjukkan bahwa persepsi publik relatif seimbang, dengan 50,3% sentimen negatif dan 49,7% sentimen positif. Sentimen negatif umumnya berkaitan dengan gangguan sistem, fluktuasi harga, dan pengalaman transaksi, sedangkan sentimen positif mencerminkan kepercayaan pengguna terhadap keamanan dan kualitas layanan platform. Penelitian ini menunjukkan bahwa kombinasi *word embedding* GloVe dan optimasi PSO merupakan pendekatan yang efektif dalam analisis sentimen pada data teks tidak terstruktur. Selain itu, hasil ini juga dapat menjadi pertimbangan bagi pengelola platform dalam memahami opini pengguna. Meskipun demikian, penelitian ini masih memiliki keterbatasan pada jumlah data yang digunakan. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan data yang lebih besar dan beragam, serta mengeksplorasi kombinasi metode lain untuk memperoleh performa model yang lebih optimal.

DAFTAR PUSTAKA

- Ama, A. U. T., Mulya, D. N., Astuti, Y. P. D., & Prasadhya, I. B. G. (2022). Analisis sentimen customer feedback Tokopedia menggunakan algoritma Naïve Bayes. *Jurnal Sistem Komputer dan Informatika (JSON)*, 4(1), 50–58. <https://doi.org/10.30865/json.v4i1.4783>
- Anjani, M., & Irmanda, H. N. (2024). Comparison performance of Word2Vec, GloVe, FastText using support vector machine method for sentiment analysis. *Jurnal Teknik Informatika (JUTIF)*, 5(3), 669–674. <https://doi.org/10.52436/1.jutif.2024.5.3.1366>
- Arsi, P., & Waluyo, R. (2021). Analisis sentimen wacana pemindahan ibu kota Indonesia menggunakan algoritma support vector machine (SVM). *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(1), 147–156. <https://doi.org/10.25126/jtiik.202183944>
- Bisnis, J., Syariah, D. K., Sudarwanti, M., Kholwati, S., & Lestari, A. (2025). Muamalah zaman now: Apakah kripto termasuk mata uang sah? *Jurnal Kajian Bisnis Syariah*. <https://doi.org/10.XXXXXX/XXXXXX>
- Dhafin Firgayanda, M., & Jumhur, H. M. (2022). Analisis pengaruh financial literacy terhadap keputusan investasi di pasar cryptocurrency (studi pada pengguna platform Indodax).
- Diannita Sari, F., Widayat, W., & Artikel, R. (2025). Analisis sentimen masyarakat terhadap PON 2024 Aceh dan Sumatra Utara di Twitter menggunakan algoritma SVM dan KNN. *AITI: Jurnal Teknologi Informasi*, 22(2), 102–116.
- Eka Achyani, Y. (2018). Penerapan metode particle swarm optimization pada optimasi prediksi pemasaran langsung. *Jurnal Informatika*, 5(1), 1–11.
- Franciska Mey Dina, D., Haryanti, T., & Amirul Haq, M. (2025). Analisis sentimen terhadap komentar pada media sosial TikTok yang berpotensi menyebabkan depresi menggunakan metode Naïve Bayes. *Jurnal Ilmiah Computing Insight*, 7(1).
- Ginabila, G., Fauzi, A., Pratiwi, R. L., Fauziah, S., & Alfianti, Z. I. (2024). Analisis sentimen perkembangan motor listrik menggunakan support vector machine dan optimasi particle swarm optimization. *INTI Nusa Mandiri*, 19(1), 77–83. <https://doi.org/10.33480/inti.v19i1.5579>

- Ikhwan Saputra, E., Khairul Anam, M., Yenni, H., & Zamsuri, A. (2025). Optimalisasi algoritma support vector machine pada aspect-based sentiment analysis menggunakan GridSearchCV. *Jurnal Sistem Informasi*, 7(1).
- Ivo Drajana. (2017). Metode support vector machine dan forward selection prediksi pembayaran pembelian bahan baku kopra. *ILKOM Jurnal Ilmiah*, 9(?).
- Maulidah, N., Supriyadi, R., Utami, D. Y., Hasan, F. N., Fauzi, A., & Christian, A. (2021). Prediksi penyakit diabetes melitus menggunakan metode support vector machine dan Naïve Bayes. *Indonesian Journal on Software Engineering (IJSE)*, 7(1), 63–68.
- Monika Parapat, I., & Tanzil Furqon, M. (2018). Penerapan metode support vector machine (SVM) pada klasifikasi penyimpangan tumbuh kembang anak. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(10).
- Munawaroh, K. (2022). Performance comparison of SVM, Naïve Bayes, and KNN algorithms for analysis of public opinion sentiment against COVID-19 vaccination on Twitter. *Journal of Advances in Information Systems and Technology*, 4(2).
- Putra, R. F. D., Sudewo, A. H., & Wibowo, A. (2024). Particle swarm optimization and GridSearch optimization on support vector machine algorithm on sentiment analysis of Donald Trump's assassination attempt. *Jurnal Teknik Informatika (JUTIF)*, 5(6), 1609–1619. <https://doi.org/10.52436/1.jutif.2024.5.6.2713>
- Ramadhan, B. Z., Riza, I., & Maulana, I. (2022). Analisis sentimen ulasan pada aplikasi e-commerce dengan menggunakan algoritma Naïve Bayes. *Journal of Applied Informatics and Computing (JAIC)*, 6(2).
- Rangga, M., Nasution, A., & Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen Twitter. *Jurnal Informatika*, 6(2), 212–218.
- Romadoni, F., Umaidah, Y., & Sari, B. N. (2020). Text mining untuk analisis sentimen pelanggan terhadap layanan uang elektronik menggunakan algoritma support vector machine. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 9(2), 247–253. <https://doi.org/10.32736/sisfokom.v9i2.903>
- Sabrila, T. S., Sari, V. R., & Minarno, A. E. (2021). Analisis sentimen pada tweet tentang penanganan COVID-19 menggunakan word embedding pada algoritma support vector machine dan k-nearest neighbor. *Fountain of Informatics Journal*, 6(2), 69–76. <https://doi.org/10.21111/fij.v6i2.5536>
- Safitri, T., Umaidah, Y., & Maulana, I. (2023). Analisis sentimen pengguna Twitter terhadap BTS menggunakan algoritma support vector machine. *Journal of Applied Informatics and Computing (JAIC)*, 7(1).
- Sakina Setiawan. (2025, June 5). Transaksi kripto RI capai Rp35,61 triliun per April 2025, 42,83 persen di Indodax. *Kompas.com*.
- Suprianto, O., Kiftiah, M., & Perdana, H. (2024). Penerapan particle swarm optimization dalam penyusunan jadwal mata kuliah di Prodi Matematika Universitas Tanjungpura. *Buletin Ilmiah Math. Stat dan Terapannya (BIMASTER)*, 13(5).
- Tirta, A., Subandono, P., & Ariatmanto, D. (2025). Optimalisasi seleksi fitur dalam analisis sentimen Bank Saqu: Studi perbandingan SVM dan random forest menggunakan information gain dan chi-square. *Sistemasi: Jurnal Sistem Informasi*.

- Vincent, R., Maulana, I., & Komarudin, O. (2023). Perbandingan klasifikasi Naïve Bayes dan support vector machine dalam analisis sentimen dengan multiclass di Twitter. *Jurnal Mahasiswa Teknik Informatika*, 7(4).
- Yuranda, R., Sutabri, T., & Wahyuningsih, D. (2023). Machine learning approach in evaluating news labels based on titles: Online media case study. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 12(3), 434–439. <https://doi.org/10.32736/sisfokom.v12i3.1808>
- Yusupa, A., & Tarigan, V. (2024). Perbandingan algoritma machine learning dalam analisis sentimen mobil listrik di Indonesia pada media sosial Twitter/X. *JIP (Jurnal Informatika Polinema)*.