OPTIMASI ALGORITMA SVM DENGAN PSO UNTUK ANALISIS SENTIMEN PADA MEDIA SOSIAL X TERHADAP KINERJA TIMNAS DI ERA SHIN TAE-YONG

Talitha Atha Anastasya, Agatha Diani Putri Saka, Milen Juventus Dappa Deke, Agung Mustika Rizki
Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur
Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur, Indonesia
22081010054@student.upnjatim.ac.id

ABSTRAK

Penelitian ini menganalisis perbandingan performa model *Support Vector Machine* (SVM) dalam klasifikasi sentimen komentar terkait Timnas Sepakbola Indonesia di era Shin Tae Yong yang diperoleh dari Twitter. Meskipun model SVM memberikan hasil yang memadai, akurasi dan efisiensi klasifikasi sentimen masih memiliki potensi untuk ditingkatkan. Oleh karena itu, tujuan penelitian ini adalah untuk meningkatkan performa model SVM dalam klasifikasi sentimen komentar dengan menggunakan optimasi parameter. Dataset yang digunakan terdiri dari 398 tweet dengan dua atribut utama yaitu komentar dan label. Atribut komentar berisi teks, sementara label menunjukkan klasifikasi sentimen, yaitu positif dan negatif. Model SVM diuji pertama kali tanpa optimasi untuk mendapatkan baseline, kemudian dilakukan optimasi menggunakan *Particle Swarm Optimization* (PSO). Pengukuran performa dilakukan dengan nilai akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model SVM tanpa optimasi menghasilkan nilai akurasi, presisi, recall, dan F1-score yang memadai namun masih dapat ditingkatkan. Setelah dilakukan optimasi menggunakan PSO, performa model meningkat secara signifikan, dengan peningkatan sebesar 3,8% pada akurasi, presisi, dan recall, serta 4% pada F1-score. Hasil ini menunjukkan bahwa optimasi parameter menggunakan PSO berhasil meningkatkan keakuratan dan efisiensi klasifikasi sentimen, memberikan wawasan lebih dalam tentang persepsi publik terhadap Timnas Sepakbola Indonesia melalui media sosial.

Kata kunci : *Analisis sentimen, Support Vector Machine (SVM), Particle Swarm Optimization (PSO), Twitter.*

1. PENDAHULUAN

Dalam perkembangan teknologi pada saat ini, yang menjadi wadah utama untuk masyarakat dalam menyampaikan opini dan sentiment terhadap berbagai isu yang ada adalah melalui media sosial, termasuk kinerja tim olahraga nasional. Media sosial adalah platform yang dimanfaatkan oleh pengguna untuk menampilkan identitas diri, berkomunikasi, berkolaborasi, berbagi informasi, dan berinteraksi dengan berbagai pengguna yang ada dimedia sosial [1].

Twitter atau X merupakan salah satu media sosial gratis yang sering digunakan oleh masyarakat untuk menyampaikan informasi secara langsung di linimasa, sekaligus menambahkan komentar terkait pengalaman dan pandangan pribadi mereka [2].

Kinerja Timnas Indonesia di era Shin Tae Yong telah memicu berbagai sentimen di media sosial X. Performa tim yang fluktuatif, harapan tinggi publik, dan strategi pelatih yang kontroversial menghasilkan perdebatan dan opini yang kompleks. Hal ini memicu berbagai reaksi dan sentimen dari para pengguna media sosial X. Opini dan komentar dari para pendukung sering kali bervariasi. Ketika Timnas memperoleh hasil positif, banyak muncul komentar yang menggambarkan kebahagiaan dan pujian. Sebaliknya, jika hasil yang diperoleh buruk, akan muncul komentar yang berisi kritik tajam bahkan cemoohan [3].

Untuk memahami sentimen publik dan pengaruhnya terhadap kinerja timnas, dibutuhkan metode analisis yang efektif. Salah satu algoritma yang dapat diterapkan adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) adalah teknik klasifikasi yang efektif untuk mengatasi masalah yang bersifat nonlinier [4].

SVM berfungsi dengan mencari proses atau fungsi pemisah (hyperplane) yang memaksimalkan jarak antar kelas dalam setiap kategori sentimen. Algoritma SVM terbukti memiliki tingkat akurasi yang lebih tinggi dalam klasifikasi teks, jika dibandingkan dengan algoritma lain seperti *K-Nearest Neighbors* (KNN), C4.5, dan *Naïve Bayes* [5].

Penelitian sebelumnya telah menerapkan algoritma SVM untuk analisis sentimen pengguna Twitter terhadap BTS dengan melakukan tiga skenario split data, yaitu, data latih sebesar 90% dan data uji 10% pada skenario 1, 80% dan 20% pada skenario 2, serta 70% dan 30% pada skenario 3, menghasilkan akurasi berturut-turut 90%, 90%, dan 89% [6].

Namun, algoritma SVM tidak selalu memberikan hasil yang optimal, sehingga diperlukan metode optimasi untuk meningkatkan kinerjanya.

Pada penelitian ini, digunakan metode optimasi *Particle Swarm Optimization* (PSO) untuk meningkatkan hasil dari kinerja algoritma *Support Vector Machine* (SVM). Berdasarkan penelitian sebelumnya, metode PSO terbukti dapat meningkatkan akurasi algoritma SVM dari 83,33% menjadi 88,89% [7], serta meningkatkan akurasi dari 79,06% menjadi 81,15% pada penelitian lainnya [8].

Hal ini menunjukkan bahwa penerapan PSO pada algoritma SVM dapat memberikan peningkatan akurasi yang signifikan, serta mengoptimalkan

parameter yang ada, sehingga menghasilkan model yang lebih efisien dan akurat dalam menyelesaikan masalah klasifikasi. Optimasi menggunakan PSO diharapkan dapat memberikan peningkatan kinerja yang lebih baik atau sebanding dari penelitian sebelumnya.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian sebelumnya yang dilakukan oleh Safitri membahas analisis sentimen terhadap pengguna Twitter yang berinteraksi dengan BTS menggunakan algoritma *Support Vector Machine* (SVM) dengan sampel data sebanyak 622 tweet [6].

Hasil akurasi yang diperoleh cukup baik, namun masih terdapat potensi untuk ditingkatkan. Mengingat hal tersebut, penelitian lain yang memiliki fokus serupa dilakukan oleh Handayani yang menganalisis sentimen pada ulasan produk di Tokopedia. Dalam penelitian ini, metode SVM dioptimasi lebih lanjut dengan menggunakan teknik *Particle Swarm Optimization* (PSO), yang berhasil meningkatkan akurasi model dari 83,33% menjadi 88,89% [7].

Peningkatan ini menunjukkan bahwa optimasi menggunakan PSO dapat memperbaiki kinerja model SVM dalam klasifikasi sentimen.

Penelitian lain yang relevan dilakukan oleh Arsi yang menganalisis sentimen terkait wacana pemindahan ibu kota Indonesia dengan pendekatan serupa. Penelitian ini juga mengimplementasikan PSO untuk mengoptimasi metode SVM dan menghasilkan peningkatan akurasi dari 79,06% menjadi 81,15% [8]. Hasil-hasil tersebut membuktikan bahwa optimasi model SVM menggunakan PSO tidak hanya efektif dalam meningkatkan akurasi, tetapi juga memberikan kontribusi signifikan dalam meningkatkan kinerja analisis sentimen di berbagai konteks.

2.2. Twitter (X)

X, yang sebelumnya dikenal sebagai Twitter, adalah platform media sosial dan layanan microblogging yang memungkinkan penggunanya mengirim pesan singkat yang disebut tweet [9]. Dengan berbagai fitur interaksi tanpa batas wilayah dan waktu, X menarik sekitar 160 juta pengguna pada tahun 2010 [10].

Seringkali, pengguna Twitter membagikan pengalaman pribadi mereka melalui cuitan yang beragam, mulai dari yang lucu hingga yang sedih [11].

X juga kerap digunakan untuk menyampaikan opini, baik positif maupun negatif, seperti pandangan masyarakat terhadap kinerja Timnas Indonesia di era Shin Tae-yong.

2.3. Analisis Sentimen

Analisis sentimen adalah metode komputasi yang digunakan untuk mengkaji pendapat, perasaan, dan emosi yang disampaikan melalui teks [12].

Dalam konteks penelitian ini, analisis sentimen diterapkan untuk menganalisis opini masyarakat terkait kinerja Timnas Indonesia di era Shin Tae-yong. Analisis ini memungkinkan identifikasi sentimen yang berkembang di media sosial X, yang menjadi platform utama bagi masyarakat untuk menyampaikan pandangan mereka mengenai hasil dan perkembangan Timnas. Analisis sentimen dimanfaatkan untuk mendapatkan gambaran keseluruhan dari media sosial, dengan tujuan mengidentifikasi apakah sentimen yang dominan bersifat positif atau negatif [13].

2.4. Algoritma *Support Vector Machine* (SVM)

Teknik klasifikasi *Support Vector Machine* (SVM) memanfaatkan ruang hipotesis dengan fungsi linear dua arah dalam ruang fitur berdimensi tinggi, yang biasanya digunakan untuk mengklasifikasikan data yang hanya memiliki dua kelas [14].

SVM sering digunakan karena memiliki akurasi yang tinggi dan kemampuan untuk menangani data dengan dimensi yang besar [15].

Hal ini sangat berguna dalam menganalisis data teks yang seringkali memiliki dimensi fitur yang besar, seperti kata-kata, frasa, atau pola tertentu dalam teks yang mencerminkan sentimen pengguna di media sosial.

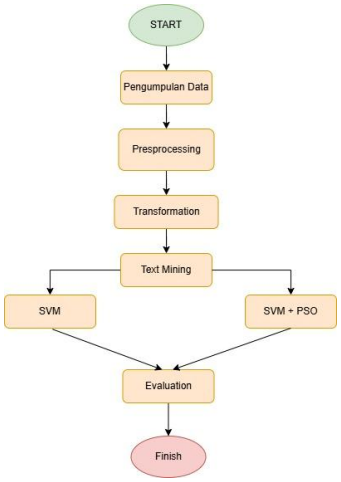
2.5. *Particle Swarm Optimization* (PSO)

Dalam penelitian ini, algoritma optimasi yang utama adalah *Particle Swarm Optimization* (PSO). Algoritma ini dipilih karena memiliki keunggulan seperti kemudahan dalam implementasi, efisiensi komputasi, fleksibilitas, serta kemampuan untuk beradaptasi dengan berbagai jenis masalah optimasi. [16].

Particle Swarm Optimization (PSO) adalah algoritma yang dikembangkan oleh Kennedy dan Eberhart pada tahun 1995. Algoritma ini terinspirasi oleh interaksi sosial dalam kelompok hewan. seperti kawanan burung atau ikan, dan dikembangkan untuk menyelesaikan masalah optimasi [17].

PSO sering dimanfaatkan untuk menyelesaikan masalah optimasi dan berfungsi sebagai metode pemilihan fitur, sebagaimana dijelaskan oleh Liu [18].

3. METODE PENELITIAN



Gambar 1. Alur penelitian

3.1. Pengumpulan Data

Kumpulan data dari penelitian ini diambil dari sumber kaggel.com. Dataset ini terdiri dari 398 data, dengan dua atribut utama, yaitu Label dan Komentar. Atribut komentar berisi teks komentar yang ditulis oleh pengguna mengenai Timnas Indonesia di bawah kepelatihan Shin Tae Yong, sementara atribut label menunjukkan klasifikasi sentimen komentar tersebut, yaitu Positif atau Negatif. Dari total 398 komentar, 51% memiliki label Positif dan 49% memiliki label Negatif.

3.2. *Preprocessing*

Data yang siap digunakan untuk langkah berikutnya akan dihasilkan dari pemrosesan data mentah yang telah dikumpulkan pada tahap preprocessing [19]. Tahapan-tahapan dalam *preprocessing* data adalah sebagai berikut:

3.2.1. *Cleansing*

Atribut yang tidak relevan dengan proses klasifikasi, seperti simbol karakter, angka, dan emotikon, akan dihapus [19].

Proses cleansing ini penting untuk memastikan bahwa hanya data yang relevan dan memiliki makna yang tersisa, sehingga tidak mengganggu analisis lebih lanjut. Pembersihan data dilakukan untuk menghilangkan gangguan yang dapat mempengaruhi performa model dalam mengklasifikasikan sentimen.

3.2.2. *Case folding*

Untuk memastikan konsistensi dalam pengolahan data teks, seluruh teks diubah menjadi huruf kecil melalui proses yang disebut *Case Folding* [20].

Langkah ini bertujuan untuk menghindari perbedaan representasi yang tidak perlu, seperti perbedaan antara huruf kapital dan huruf kecil, yang tidak mempengaruhi makna suatu kata. Dengan melakukan *case folding*, model dapat memproses data secara lebih efisien dan konsisten.

3.2.3. *Tokenization*

Tahap ini bertujuan untuk memisahkan setiap kalimat menjadi kata-kata individual, sehingga setiap kata dapat dianalisis secara mandiri dalam proses pengolahan data teks [20].

Tokenization sangat penting dalam mempersiapkan data teks untuk analisis lebih lanjut, karena model hanya dapat memproses data dalam bentuk unit terkecil yang dapat dipahami, yaitu kata-kata. Proses ini memungkinkan pengolahan teks yang lebih mendalam dan analitis.

3.2.4. *Stopword Removal*

Kata-kata yang termasuk dalam daftar stopword dari NLTK dihapus pada tahap *stopword removal*, seperti kata penghubung "dan," "atau," "yang," dan lain sebagainya. Selain itu, daftar *stopword* tambahan disusun secara manual untuk mencakup kata-kata, termasuk nama-nama yang tidak relevan dengan penelitian ini dan tidak tercakup dalam daftar stopword bawaan NLTK [21].

Penghapusan stopword bertujuan untuk mengurangi beban pemrosesan data, karena kata-kata ini tidak memberikan kontribusi signifikan terhadap makna atau sentimen yang ingin dianalisis.

3.2.5. *Normalization*

Kata-kata yang tidak sesuai dengan ejaan atau aturan dalam Kamus Besar Bahasa Indonesia (KBBI) diperbaiki pada tahap ini [21].

Normalization diperlukan untuk memastikan bahwa setiap kata yang digunakan dalam teks sesuai dengan standar bahasa yang benar. Langkah ini menghindari kesalahan yang disebabkan oleh variasi ejaan atau penggunaan bahasa yang tidak baku.

3.2.6. *Stemming*

Pada tahap *stemming*, proses pengubahan kata-kata yang memiliki imbuhan ke bentuk kata dasar dilakukan dengan memanfaatkan library Sastrawi [21].

Stemming digunakan untuk mengurangi variasi kata yang berasal dari kata dasar yang sama, misalnya "kualitasnya" akan diubah menjadi kata dasar "kualitas". Dengan demikian, model dapat lebih fokus pada makna inti dari kata-kata tersebut tanpa terpengaruh oleh variasi morfologis yang tidak penting.

3.3. *Transformation*

Pada tahap transformasi Proses ini mencari fitur-fitur yang mampu merepresentasikan data, dengan bergantung pada pola serta jenis informasi yang ada [22].

Tahap awal menghasilkan pola yang bertujuan untuk mengubah ide abstrak menjadi bentuk konkret yang dapat diwujudkan. Setelah itu, label dan dataset yang awalnya memiliki bentuk data teks diubah menjadi format numerik dengan memanfaatkan algoritma TF–IDF (Term Frequency–Inverse Document Frequency) [23].

3.4. *Text Mining*

Teknik *text mining* digunakan untuk mengelola data berupa teks, seperti tulisan atau dokumen, dalam proses pengolahan informasi, dengan tujuan mengklasifikasi, mengidentifikasi, serta mengelompokkan data tersebut [24].

Text mining, sebagai salah satu cabang data mining, menyediakan cara untuk mengekstrak informasi utama dari teks yang kompleks dan berukuran besar. Metode ini berguna untuk menganalisis umpan balik dan komentar, dengan tujuan memahami pesan serta perasaan yang disampaikan oleh pendukung [25].

Dengan demikian, *text mining* menjadi alat yang sangat bermanfaat dalam mendukung pengambilan

keputusan berbasis data teks, baik dalam konteks bisnis, penelitian, maupun analisis sosial.

3.5. *Evaluation*

Tahap akhir dari proses ini adalah menganalisis pola atau informasi yang telah dihasilkan melalui data mining digunakan untuk mengevaluasi apakah hasil yang diperoleh sesuai dengan hipotesis atau fakta yang telah ditetapkan sebelumnya. Proses ini tidak hanya memastikan validitas data yang ditemukan tetapi juga memberikan peluang untuk memperbaiki atau memperbarui hipotesis jika hasilnya menunjukkan temuan baru [22].

Pada tahap ini, kinerja algoritma SVM akan diukur menggunakan confusion matrix, dengan hasil evaluasi yang mencakup nilai akurasi, presisi, recall, dan f1-score [19].

Dalam penelitian ini, tahap ini memiliki peranan yang sangat krusial karena menjamin bahwa data yang diperoleh tidak hanya memiliki relevansi, tetapi juga dapat dipercaya untuk mendukung dalam hal mengambil keputusan yang lebih strategis dan tepat.

4. HASIL DAN PEMBAHASAN

Hasil yang di didapat pada penelitian ini mengulas penerapan algoritma *Support Vector Machine* (SVM) yang dioptimalkan menggunakan *Particle Swarm Optimization* (PSO). Penelitian dilakukan dalam dua tahap, yaitu, pertama, penerapan SVM untuk membangun model klasifikasi, dan kedua, penerapan PSO untuk mengoptimalkan parameter-parameter pada SVM. Kombinasi kedua metode ini diharapkan dapat memberikan peningkatan pada kinerja model, sehingga prediksi yang dihasilkan lebih akurat dan efisien.

4.1. *Preprocessing*

4.1.1. *Cleansing*

Tabel 1. Perbandingan proses *cleansing*

Sebelum *cleansing*	**Setelah *cleansing***
<username> <username> posisinya arhan double juga tuh sama edo, apa spesialnya bgst arhan, STY emang nih	posisinya arhan double juga tuh sama edo apa spesialnya bgst arhan STY emang nih
<username> Kenapa bukan Dendy S dn Dimas D....sty takut x sama bosnya mrk yaÃ°Å¸ËœÂ□Ã°Å¸ËœÂ□Ã°Å¸ËœÂ□	Kenapa bukan Dendy S dn Dimas Dsty takut x sama bosnya mrk ya

Proses data *cleansing* dilakukan untuk memastikan data yang digunakan bebas dari masalah seperti *noise*, nilai kosong, dan duplikasi. Sebelum *cleansing*, dataset mengandung banyak *noise* yang dapat mengganggu hasil analisis. Data kosong dihapus sehingga jumlah data berkurang dari 398 menjadi 395, dan terdapat 3 data duplikat juga dihilangkan untuk menjaga keunikan setiap entri dan data bersihnya menjadi 392. Setelah proses ini, dataset menjadi lebih terstruktur dan bersih, sehingga mempermudah pengolahan pada tahap selanjutnya.

4.1.2. *Case Folding*

Tabel 2. Perbandingan proses *case folding*

Sebelum *case folding*	**Sebelum *case folding***
Fix STY punya hutang budi sama dendy	fix sty punya hutang budi sama dendy
Satu line up naturalisasi smua	satu line up naturalisasi smua

Pada tahap *case folding*, semua huruf dalam dataset yang sebelumnya menggunakan huruf kapital (*uppercase*) diubah menjadi huruf kecil (*lowercase*). Proses ini dilakukan untuk menyamakan format penulisan sehingga data menjadi lebih konsisten. Dengan demikian, perbedaan huruf besar dan kecil tidak lagi mempengaruhi analisis pada tahap selanjutnya.

4.1.3. *Tokenization*

Table 3. Perbandingan proses *tokenization*

Sebelum *tokenization*	**Sebelum *tokenization***
tenang maseeee percaya sty	['tenang', 'maseeee', 'percaya', 'sty']
udah naturalisasi kali kayak pemain sepak bola	['udah', 'naturalisasi', 'kali', 'kayak', 'pemain', 'sepak', 'bola']

Pada tahap tokenisasi, setiap kalimat dalam dataset dipecah menjadi bagian-bagian kecil berupa kata-kata terpisah. Proses ini memisahkan setiap kata dalam kalimat menggunakan spasi sebagai pemisah. Setelah dilakukan tokenisasi, setiap kata pada dataset dapat dianalisis secara individu, memudahkan proses pengolahan data pada tahap berikutnya.

4.1.4. *Stopword Removal*

Table 4. Perbandingan proses *stopword removal*

Sebelum *stopword removal*	**Sebelum *stopword removal***
['gw', 'yakin', 'seyakin', 'yakinnya', 'sty', 'habis', 'piala', 'asia', 'di', 'pecat', 'pelatih', 'tolol']	['gw', 'yakin', 'seyakin', 'yakinnya', 'sty', 'habis', 'piala', 'asia', 'pecat', 'pelatih', 'tolol']
['siap2', 'sty', 'kebantai', 'yg', 'ke', 'dua', 'kali']	['siap2', 'sty', 'kebantai', 'yg', 'dua', 'kali']

Pada tahap *stopword removal*, kata-kata yang dianggap tidak relevan atau tidak memberikan kontribusi berarti dalam analisis, seperti kata sambung

atau kata yang sering muncul tanpa makna signifikan, dihapus dari dataset. Sebagai contoh, kata-kata seperti "di", "ke", dan sebagainya dihapus karena tidak memberikan informasi penting. Setelah proses ini, hanya kata-kata yang lebih bermakna dan relevan, seperti "gw", "yakin", "pelatih", dan lainnya, yang tetap ada. Hal ini bertujuan untuk menyaring kata-kata yang tidak perlu dan memastikan analisis lebih fokus pada informasi yang penting.

4.1.5. *Normalization*

Table 5. Perbandingan proses *normalization*

Sebelum *normalization*	**Sebelum *normalization***
['kalo', 'gak', 'bisa', 'menang', 'lawan', 'vietnam', 'mending', 'sty', 'out']	['kalau', 'tidak', 'bisa', 'menang', 'lawan', 'vietnam', 'mending', 'sty', 'keluar']
['bilang', 'aja', 'takut', 'tim', 'besutan', 'sty']	['bilang', 'saja', 'takut', 'tim', 'besutan', 'sty']

Pada tahap *normalization*, proses dilakukan dengan mengubah kata-kata yang tidak baku dalam dataset menjadi bentuk bakunya. Dengan menggunakan kamus normalization, kata-kata yang memiliki bentuk tidak standar atau variasi penulisan, seperti "kalo" yang diubah menjadi "kalau", akan disesuaikan agar lebih konsisten dan sesuai dengan kaidah bahasa yang benar. Proses ini bertujuan untuk meningkatkan kualitas data dan memastikan bahwa kata-kata yang digunakan dalam analisis memiliki bentuk yang seragam dan baku.

4.1.6. *Stemming*

Table 6. Perbandingan proses *stemming*

Sebelum *stemming*	**Sebelum *stemming***
['bagi', 'sty', 'dendi', 'kualitasnya', 'setara', 'asia', 'lilipaly', 'aff', 'saja']	['bagi', 'sty', 'dendi', 'kualitas', 'tara', 'asia', 'lilipaly', 'aff', 'saja']
['biasa', 'sty', 'kalau', 'tantrum', 'gitu', 'semua', 'hal', 'dibawa', 'mana', 'tidak', 'yang', 'ingat', 'pula']	['biasa', 'sty', 'kalau', 'tantrum', 'gitu', 'semua', 'hal', 'bawa', 'mana', 'tidak', 'yang', 'ingat', 'pula']

Pada tahap ini, proses *stemming* dilakukan untuk mengubah kata-kata dalam data teks berbahasa Indonesia menjadi bentuk dasarnya. Proses ini menggunakan bantuan library Sastrawi di *Python*, yang secara otomatis mengenali dan menghapus imbuhan seperti awalan, akhiran, sisipan, maupun imbuhan gabungan. Contohnya, kata "kualitasnya" diubah menjadi "kualitas" dan "dibawa" menjadi "bawa". Proses ini diterapkan secara konsisten pada seluruh data untuk memastikan teks lebih sederhana, terstruktur, dan siap untuk tahap analisis berikutnya.

4.2. *Transformation*

Sebelum dilakukan tahap *transformation*, data terlebih dahulu dibagi menggunakan metode split dengan perbandingan 80% untuk data training dan 20% untuk data testing. Proses transformation data dilakukan dengan menerapkan pembobotan kata menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF), yang diimplementasikan menggunakan modul TfidfVectorizer dari library Scikit-Learn. Teknik TF-IDF ini bertujuan untuk mengonversi data teks (komentar) menjadi format numerik agar dapat diproses oleh model *Support Vector Machine* (SVM). TfidfVectorizer menghitung bobot kata berdasarkan frekuensi kemunculannya di dokumen dan seberapa jarang kata tersebut muncul di keseluruhan dokumen, sehingga dapat menentukan tingkat kepentingan kata dalam konteks klasifikasi.

4.3. *Text Mining*

Proses text mining dalam penelitian ini dilakukan melalui beberapa tahapan, dimulai dengan klasifikasi data teks berupa komentar menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel Radial Basis Function (RBF). Pada tahap awal, model SVM dibangun dan dilatih menggunakan parameter default untuk menghasilkan baseline performa. Selanjutnya, untuk meningkatkan akurasi dan efisiensi model, dilakukan proses optimasi parameter dengan metode *Particle Swarm Optimization* (PSO). Metode ini diterapkan untuk mencari nilai parameter terbaik, seperti C yang mengontrol trade-off antara margin dan kesalahan klasifikasi, serta gamma yang menentukan pengaruh jarak dalam fungsi kernel, sehingga menghasilkan model yang lebih optimal dalam memprediksi data teks.

4.4. *Evaluation*

Pada tahap evaluasi, penelitian ini bertujuan untuk mengukur performa algoritma *Support Vector Machine* (SVM) dengan dan tanpa optimasi menggunakan metode *Particle Swarm Optimization* (PSO). SVM digunakan karena kemampuannya dalam mengklasifikasikan data secara efektif, terutama pada dataset dengan dimensi tinggi. Metrik presisi, akurasi, recall, dan F1-score digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan data secara menyeluruh. Evaluasi dilakukan dengan satu skenario pembagian data, yaitu 80% untuk data training dan 20% untuk data testing.

Table 7. Perbandingan evaluasi

	SVM	**SVM-PSO**
Akurasi	72,1%	75,9%
Presisi	72,3%	76,1%
Recall	72,1%	75,9%
F1-Score	71,4%	75,4%

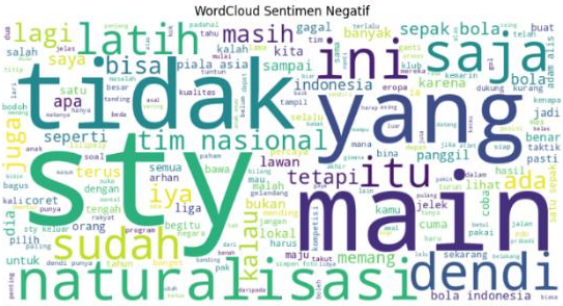
Tabel di atas menggambarkan perbandingan performa antara model SVM tanpa optimasi dan model SVM yang telah dioptimasi menggunakan PSO. Pada model awal, SVM tanpa optimasi menunjukkan performa dasar dengan parameter default,

menghasilkan nilai akurasi, presisi, recall, dan F1-score yang memadai namun masih memiliki ruang untuk peningkatan. Hasil ini mencerminkan bagaimana SVM bekerja dengan pengaturan standar, tetapi model masih dapat lebih ditingkatkan dalam hal keakuratan dan efisiensi. Setelah dilakukan optimasi menggunakan PSO, performa model meningkat secara signifikan di seluruh metrik evaluasi. Akurasi meningkat sekitar 3,8%, presisi naik 3,8%, recall juga mengalami peningkatan sebesar 3,8%, dan F1-score meningkat sekitar 4%.

Peningkatan ini menunjukkan bahwa PSO efektif dalam menemukan parameter optimal (C dan gamma) untuk SVM, sehingga model dapat memisahkan kelas data dengan lebih akurat dan andal. Rata-rata peningkatan keseluruhan mencapai sekitar 3,85%, yang menunjukkan bahwa optimasi dengan PSO secara signifikan meningkatkan kinerja model. Hasil ini membuktikan bahwa optimasi parameter menggunakan PSO tidak hanya memberikan peningkatan dalam akurasi, tetapi juga memberikan keseimbangan yang lebih baik antara presisi, recall, dan F1-score. Dengan demikian, SVM yang dioptimasi dengan PSO menjadi pilihan yang lebih unggul dalam menghadapi tugas klasifikasi yang lebih kompleks. Selain itu, metode ini juga menunjukkan fleksibilitasnya dalam berbagai jenis dataset, termasuk dataset dengan dimensi tinggi dan distribusi data yang tidak merata. PSO membantu dalam menghindari overfitting dengan menemukan parameter yang tepat, sehingga menghasilkan model yang lebih stabil dan dapat diandalkan. Hasil ini menegaskan pentingnya penggunaan metode optimasi seperti PSO dalam meningkatkan kinerja algoritma pembelajaran mesin pada masalah klasifikasi yang menantang.

Gambar 2. *Word Cloud Positif*

Word Cloud untuk sentimen positif menunjukkan kata-kata yang paling sering muncul dalam komentar yang mendukung Timnas Sepakbola Indonesia di bawah kepelatihan Shin Tae Yong. Kata-kata seperti "yang", "main", dan "sty" (singkatan dari Shin Tae Yong) menunjukkan bahwa komentar-komentar ini banyak berfokus pada pelatih dan gaya permainan tim. Selain itu, kata "naturalisasi" juga sering muncul, yang mengindikasikan adanya dukungan terhadap kebijakan naturalisasi pemain asing dalam skuad Timnas Indonesia. Kata-kata seperti "tim nasional", "latih", dan "banyak" menonjolkan optimisme terhadap perkembangan dan potensi tim, serta penekanan pada aspek pelatihan dan strategi yang diterapkan oleh pelatih. Secara keseluruhan, komentar positif ini mencerminkan harapan dan keyakinan banyak pendukung terhadap masa depan Timnas Indonesia, meskipun tantangan dan kekurangan masih ada.



Gambar 3. *Word Cloud negatif*

Word Cloud untuk sentimen negatif menggambarkan ketidakpuasan publik terhadap Timnas Sepakbola Indonesia dan pelatih Shin Tae Yong. Kata-kata seperti "tidak", "sty", dan "itu" menunjukkan kritik terhadap keputusan-keputusan yang diambil oleh pelatih serta kinerja Timnas Indonesia. Selain itu, kata "main" dan "latih" muncul dalam konteks ketidakpuasan terhadap permainan tim dan metode pelatihan yang diterapkan. Kritik terhadap kebijakan naturalisasi pemain juga terlihat dengan munculnya kata "naturalisasi". Kata-kata seperti "dendi" dan "sudah" menunjukkan kekecewaan terhadap performa beberapa pemain dan keputusan yang dianggap terlambat atau tidak efektif. Secara keseluruhan, komentar negatif ini mencerminkan frustasi penggemar terhadap kinerja tim dan kebijakan yang diambil selama kepemimpinan Shin Tae Yong.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini menunjukkan bahwa analisis sentimen terhadap kinerja Timnas Indonesia di era Shin Tae Yong menggunakan algoritma *Support Vector Machine* (SVM) yang dioptimasi dengan *Particle Swarm Optimization* (PSO) memberikan hasil yang lebih baik. Sebelum optimasi, model SVM mencapai akurasi 72,1%, sedangkan setelah optimasi dengan PSO, akurasi meningkat menjadi 75,9%. Penerapan PSO berhasil meningkatkan kinerja model dalam menganalisis sentimen publik terhadap Timnas Indonesia di media sosial X. Secara keseluruhan, optimasi menggunakan PSO terbukti efektif dalam meningkatkan akurasi dan kinerja model SVM, yang memiliki potensi besar untuk diterapkan pada analisis sentimen di berbagai konteks lainnya. Penelitian selanjutnya dapat mempertimbangkan penggunaan dataset yang lebih besar dan eksplorasi metode optimasi lainnya untuk perbandingan, serta penerapan model deep learning untuk hasil yang lebih akurat.

DAFTAR PUSTAKA

[1] T. Krisdiyanto, "Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, p. 32, 2021, doi: 10.24014/coreit.v7i1.12945.

[2] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, "Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 47–58, 2021, doi: 10.28932/jutisi.v7i1.3216.

[3] A. E. Pratama, A. Ariesta, and G. Gata, "Analisis Sentimen Masyarakat Terhadap Tim Nasional Indonesia Pada Piala AFF 2020 Menggunakan Algoritma K-Nearest Neighbors," *J. TICOM Technol. Inf. Commun.*, vol. 10, no. 3, pp. 187–196, 2022, [Online]. Available: https://jurnal-ticom.jakarta.aptikom.org/index.php/Ticom/article/view/33/47

[4] B. Pamungkas, A. Syaifuddin, and M. Muslimin, "Analisis Sentimen Twitter MenggunakanMetode Support Vector Machine (SVM) padaKasus Benih Lobster 2020," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 3, no. 2, pp. 10–20, 2021.

[5] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, 2021, doi: 10.34148/teknika.v10i1.311.

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

[7] R. Nurfitriana Handayani, "Optimasi Algoritma Support Vector Machine Untuk AnalisisSentimen Pada Ulasan Produk Tokopedia Menggunakan Pso," *Media Inform.*, vol. 20, no. 2, pp. 97–108, 2021.

[8] P. Arsi, R. Wahyudi, and R. Waluyo, "Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021, doi: 10.29207/resti.v5i2.2698.

[9] H. Sulastomo, Ramadiansyah, K. Gibran, E. Maryansyah, and A. Tegar, "Analisis Sentimen Pada Twitter @Ovo_Id dengan Metode Support Vectore Machine (SVM)," *J. Sains Komput. Inform. (J-SAKTI*, vol. 6, no. 2, pp. 1050–1056, 2022.

[10] I. Rossianah, "Pengaruh Motif Penggunaan Media Sosial Twitter / X Terhadap Kecemasan Dengan Kecenderungan Self Injury Pada Generasi Z," *Indones. J. Bus. Innov.*, vol. 1, no. 1, pp. 205–217, 2024.

[11] R. A. Siregar and D. Kusyani, "Tindak Tutur Ekspresif Dalam Meme Bu Tejo Tilik Di Twitter Sebagai Bahan Ajar Siswa Smp (Suatu Kajian Pragmatik)," *PRASASTI J. Linguist.*, vol. 6, no. 2, p. 227, 2021, doi: 10.20961/prasasti.v6i2.53492.

[12] T. Juniardi and C. A. Sugianto, "ANALISIS SENTIMEN TIM NASIONAL SEPAK BOLA INDONESIA DI TURNAMEN PIALA DUNIA U-17 INDONESIA PADA TWITTER (X) MENGGUNAKAN," vol. 12, no. 3, 2024.

[13] A. Nurian, T. N. Padilah, and G. Garno, "Analisis Sentimen Terhadap Pelayanan Disdukcapil Karawang Menggunakan Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4178.

[14] D. Alita, "Multiclass SVM Algorithm for Sarcasm Text in Twitter," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 118–128, 2021, doi: 10.35957/jatisi.v8i1.646.

[15] R. Ramlan, N. Satyahadewi, and W. Andani, "Analisis Sentimen Pengguna Twitter Menggunakan Support Vector Machine Pada Kasus Kenaikan Harga BBM," *Jambura J. Math.*, vol. 5, no. 2, pp. 431–445, 2023, doi: 10.34312/jjom.v5i2.20860.

[16] A. Herliana and S. S. Muawiyah, "Komparasi Optimasi Analisis Sentimen Cyberbullying Pada Instagram Berbasis Particle Swarm Optimization," *J. Responsif Ris. Sains dan Inform.*, vol. 6, no. 1, pp. 43–53, 2024, doi: 10.51977/jti.v6i1.1419.

[17] M. H. Ramdani, G. Pasek, S. Wijaya, and R. Dwiyansaputra, "Optimalisasi Pengenalan Wajah Berbasis Linear Discriminant Analysis dan K-Nearsert Neighbor Menggunakan Particle Swarm Optimization (Optimization Of Face Recognition Based On Linear Discriminant Analysis and K-Nearest Neighbor Using Particle Swarm Optimiz," vol. 4, no. 1, pp. 40–51, 2022, [Online]. Available: http://jtika.if.unram.ac.id/index.php/JTIKA/

[18] A. Taufik, "Optimasi Particle Swarm Optimization Sebagai Seleksi Fitur Pada Analisis Sentimen Review Hotel Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes," *J. Tek. Komput.*, vol. III, no. 2, pp. 40–47, 2017.

[19] M. Diki Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Mola Application Reviews on Google Play Store Using Support Vector Machine Algorithm," *J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, pp. 1–7, 2022.

[20] N. F. Az-haari, D. Juardi, and A. Jamaludin, "Analisis Sentimen Terhadap Boikot Brand Pro-Israel Pada Twitter Menggunakan Metode Naïve

Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 4256–4261, 2024, doi: 10.36040/jati.v8i3.9888.

[21] F. Meila, A. Sofyan, N. Sulistiyowati, and A. Voutama, "ANALISIS SENTIMEN TERHADAP RESPON PERUBAHAN NAMA TWITTER MENJADI ' X ' MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER," vol. 8, no. 5, pp. 10987–10994, 2024.

[22] L. Nurhidayati, Y. Umaidah, and U. Enri, "Analisis Sentimen Isu Childfree Di Media Sosial Twitter Menggunakan Algoritma Support Vector Machine," *J. Ilm. Wahana Pendidikan, Februari*, vol. 10, no. 4, pp. 422–430, 2024, [Online]. Available: https://doi.org/10.5281/zenodo.10521284

[23] N. Fibriyanti Arminda, N. Sulistiyowati, and T. Nur Padilah, "Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1817–1822, 2023, doi: 10.36040/jati.v7i3.7012.

[24] C. Joergensen E Munthe, N. Astuti Hasibuan, and H. Hutabarat, "Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace," *Resolusi Rekayasa Tek. Inform. dan Inf.*, vol. 2, no. 3, pp. 110–115, 2022, doi: 10.30865/resolusi.v2i3.309.

[25] A. Satria, "Implementasi Text Mining untuk Analisis Review pada Aplikasi Crowdfunding LX dan ST Menggunakan Metode Sentiment Analysis," *LANCAH J. Inov. Dan Tren*, vol. 2, no. 1, pp. 124-130file:///C:/Users/User/Pictures/PALEMBANG/JU, 2024.