

ANALISIS SENTIMEN ULASAN APLIKASI INVESTASI MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE BERBASIS PARTICLE SWARM OPTIMIZATION

Muhammad Samsul Ma'arif^{1*}, Jajam Haerul Jaman², Agung Susilo Yuda Irawan³

^{1,2,3} Universitas Singaperbangsa Karawang, Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361, Telp. (0267) 641177

Received: 18 Juni 2024

Accepted: 31 Juli 2024

Published: 7 Agustus 2024

Keywords:

SVM

Swarm

Sentimen

Correspondent Email:

2010631170022@student.unsika.ac.id

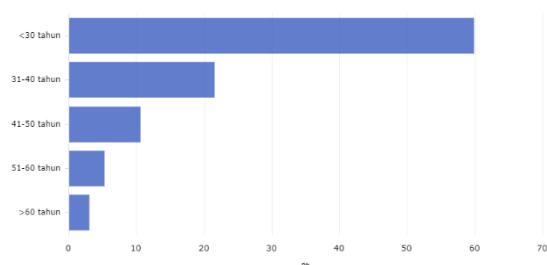
Abstrak. Perkembangan teknologi telah mengubah cara hidup manusia menjadi digital, memberikan dampak yang besar terhadap sektor ekonomi, bisnis, dan investasi. Namun, tindakan penipuan sering terjadi. Oleh karena itu, penting untuk mempertimbangkan ulasan pengguna sebelumnya dalam memilih aplikasi. Namun, jumlah ulasan yang terus meningkat membaca sebagian ulasan membuat hasilnya kurang representatif. Metodologi yang digunakan yaitu KDD dengan dua pemodelan, pertama Algoritma SVM dan yang kedua Algoritma SVM dioptimasi dengan PSO. Data dibagi menjadi dua, yaitu data pelatihan dan data uji, dengan empat skenario perbandingan 90:10, 80:20, 70:30, dan 60:40. Hasil yang didapat dari pemodelan pertama menyatakan bahwa Algoritma SVM mendapatkan hasil bahwa nilai *accuracy* tertinggi pada tiap aplikasi berbeda-beda. Sedangkan pada pemodelan kedua dengan Algoritma SVM dioptimasi oleh PSO mendapatkan bahwa nilai akurasi tertinggi pada semua aplikasi yaitu pada skenario ke-1 dengan pembagian data 90:10. Nilai evaluasi lainnya juga meningkat, termasuk akurasi, presisi, *recall*, dan *f1-score*. Aplikasi Ajaib berhasil mencapai hasil terbaik dengan akurasi tertinggi mencapai 93,33%, serta nilai presisi 100%, *recall* 90%, dan *f1-score* 95%. Maka dapat disimpulkan bahwa kombinasi analisis sentimen ulasan menggunakan algoritma SVM yang kemudian dioptimasi dengan PSO menunjukkan kinerja yang lebih unggul dibandingkan dengan penggunaan algoritma SVM saja.

Abstract. The development of technology has transformed human life into a digital era, significantly impacting the economic, business, and investment sectors. However, fraudulent activities frequently occur. Therefore, it is crucial to consider previous user reviews when selecting an application. The ever-increasing number of reviews makes reading only a portion of them less representative. The methodology used is KDD with two models: first, the SVM Algorithm, and second, the SVM Algorithm optimized with PSO. The data is divided into two sets: training data and test data, with four comparison scenarios of 90:10, 80:20, 70:30, and 60:40. The results from the first model indicate that the SVM Algorithm achieves the highest accuracy values differently for each application. In contrast, the second model with SVM optimized by PSO shows that the highest accuracy value for all applications is in the first scenario with a 90:10 data split. Other evaluation metrics also improved, including accuracy, precision, recall, and *f1-score*. The Ajaib application achieved the best results with the highest accuracy reaching 93.33%, along with a precision of 100%, recall of 90%, and *f1-score* of 95%. Thus, it can be concluded that the combination of sentiment analysis of reviews

using the SVM algorithm optimized with PSO demonstrates superior performance compared to using the SVM algorithm alone.

1. PENDAHULUAN

Dunia telah mengalami perkembangan teknologi yang pesat dan dinamis, mengubah gaya hidup manusia dari yang tradisional menjadi digital. Transformasi ini telah membawa dampak signifikan pada sektor ekonomi, bisnis, dan investasi. Menurut data yang dikumpulkan pada bulan November 2021 oleh Laporan Kustodian Sentral Efek Indonesia (KSEI, 2021) dalam artikel Katada Insight Center, terdapat sekitar 7,15 juta investor aktif di pasar modal Indonesia. Distribusi persentase investor di pasar modal berdasarkan kelompok usia dapat dilihat pada gambar 1.1 di bawah ini.



Gambar 1.1 Persentase Investor Pasar Modal Berdasarkan Kelompok Usia
(Sumber: Katada Insight Center, 2021)

Gambar 1.1 menunjukkan bahwa sekitar 59,81% dari investor pasar modal merupakan generasi milenial, yang terdiri dari individu yang berusia 30 tahun ke bawah. Sementara itu, investor pasar modal terbanyak berikutnya berasal dari kelompok usia 31-40 tahun, mencapai 21,48%. Jumlah investor pasar modal dari usia 41-50 tahun mencapai 10,51%. Selanjutnya, sebanyak 5,23% dari investor pasar modal berada dalam rentang usia 51-60 tahun, dan 2,97% lainnya berusia 60 tahun ke atas.

Dalam artikel IDN Times, menyebutkan bahwa terdapat tujuh aplikasi investasi populer di Indonesia yang cocok bagi pemula, antara lain Ajaib, Bareksa, Bibit, HSB Investasi, Landx, KoinWorks, dan Tanamduit. Namun, laporan artikel kompas.com dengan judul "Waspada, Ini Daftar 28 Investasi Bodong yang Disetop OJK" mengungkapkan bahwa beberapa aplikasi investasi online terlibat dalam

praktik penipuan. Oleh karena itu, dalam memilih aplikasi investasi online, sangat penting untuk mengumpulkan informasi dari pengalaman pengguna sebelumnya. Dalam proses pengambilan keputusan mengunduh aplikasi, ulasan dan penilaian pelanggan memiliki peran yang sangat penting. Ulasan memberikan wawasan berharga tentang pengalaman pengguna lain dengan aplikasi tersebut, termasuk kelebihan dan kekurangan yang pengguna sebelumnya alami. Analisis sentimen dapat mengungkap tingkat kepuasan dan kesetiaan pelanggan [1]. Meskipun begitu, membaca ulasan secara menyeluruh membutuhkan waktu yang lama karena jumlah ulasan pengguna terus meningkat setiap tahunnya. Membaca sebagian ulasan membuat hasilnya kurang representatif. Saat ini, penelitian mendalam mengenai sentimen pengguna terhadap aplikasi investasi online di Indonesia masih terbatas, sehingga menciptakan kebutuhan untuk melakukan analisis terhadap sentimen pengguna terkait berbagai platform investasi online yang ada. Analisis sentimen dapat memberikan wawasan yang lebih baik terkait perilaku pengguna terhadap aplikasi investasi online melalui ulasan yang terdapat di google play store.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Dalam industri *text mining*, analisis sentimen sering disebut sebagai *opinion mining* [2]. Prosesnya melibatkan pemahaman, ekstraksi, dan pemrosesan otomatis data teks untuk mendapatkan informasi mengenai sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen umumnya digunakan untuk memahami sikap seorang pembicara atau penulis terhadap suatu topik atau polaritas keseluruhan dokumen. Sikap ini bisa berasal dari pendapat, penilaian, atau evaluasi, juga dapat melibatkan keadaan afektif (emosi yang dialami penulis saat menulis) atau komunikasi emosional (emosi yang ingin disampaikan penulis kepada pembaca). Analisis sentimen dapat diterapkan pada banyak hal dan entitas, seperti barang dan layanan, peristiwa sosial dan

politik, dan kegiatan lainnya [3]. Salah satu manfaat utama analisis sentimen adalah membantu dalam memahami pendapat masyarakat atau orang lain mengenai suatu produk. Analisis sentimen dapat digunakan sebagai alat untuk melacak respon dan pandangan terhadap produk, merek, atau individu tertentu di platform web, sehingga bisnis dapat mengetahui apakah pandangan tersebut bersifat positif atau negatif. Analisis sentimen fokus pada klasifikasi ulasan berdasarkan polaritas, yang terbagi menjadi dua jenis: klasifikasi subjektivitas dan klasifikasi menjadi positif atau negatif [4].

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan suatu metode untuk melakukan prediksi pada kasus regresi atau klasifikasi salah satu jenis supervised learning [5]. Konsep SVM pertama kali diperkenalkan oleh Vapnik pada tahun 1992. Prinsip dasar dari algoritma SVM adalah menyelesaikan masalah klasifikasi, terutama dalam konteks pemisahan data linear. Hyperplane adalah garis yang digunakan sebagai pemisah optimal antara kelas -1 dan +1, dan optimalisasinya diukur dengan memaksimalkan jarak dari hyperplane ke titik data terdekat dari kedua kelas tersebut. Keunikan SVM, terletak pada kemampuannya untuk menangani pemisahan input yang bersifat non-linier [6]. Hal ini dapat dicapai melalui penggunaan berbagai jenis kernel, termasuk kernel linier, kernel polynomial, dan kernel radial basis function (RBF). Dengan memilih fungsi kernel yang sesuai, SVM mampu menangani masalah klasifikasi non-linier dengan efektif.

2.3. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) adalah algoritma optimasi yang terinspirasi oleh cara ikan dan burung mencari makanan [7]. Sebuah kelompok partikel berada di dalam ruang pencarian dengan lokasi dan ukuran yang tidak teratur. Posisi dan kecepatan partikel menjadi dua karakteristik utama dalam Particle Swarm Optimization (PSO). Salah satu aspek unik dari Particle Swarm Optimization (PSO) adalah bahwa jika suatu partikel menemukan jalur yang benar dan efisien menuju sumber makanan atau solusi terbaik, partikel tersebut akan berkomunikasi dengan partikel lain dalam

populasi untuk mengikuti jalur tersebut, meskipun mereka mungkin berada di lokasi yang jauh. Dalam penelitian [8]. Algoritma Particle Swarm Optimization memiliki kemiripan dengan Algoritma Genetik (GA) yang diawali dengan mencari sebuah individu (particle) menggunakan populasi (swarm) acak di dalam matriks. Baris dalam matriks biasa disebut *particle* atau kromosom jika dalam algoritma genetika. PSO mempunyai parameter seperti posisi, kecepatan, kecepatan maksimum, konstanta percepatan dan inersia. Algoritma PSO mempunyai kinerja yang lebih unggul dan cepat untuk masalah optimasi dibandingkan dengan algoritma genetika (GA) [8]. *Velocity* di dalam PSO untuk update masing-masing *particle* lalu dijumlahkan ke lokasi *particle*. Update *velocity* dapat dipengaruhi oleh dua solusi diantaranya global *best* dengan biaya rendah yang didapatkan dari sebuah *particle* dan solusi *local best* dengan biaya paling rendah dipopulasi awal. Apabila solusi *local best* kurang dari global *best* maka solusi *local best* dapat menggantikan solusi global *best*. Berikut istilah yang terdapat pada algoritma Particle Swarm Optimization diantaranya [9]:

- a. *Swarm* adalah populasi dalam suatu algoritma.
- b. *Particle* adalah anggota dalam sebuah *swarm*.
- c. *Pbest* (personal best) adalah posisi *particle* yang dioptimalkan untuk mendapatkan solusi terbaik.
- d. *Gbest* (*global best*) adalah posisi terbaik di antara semua *Pbest* yang ada.
- e. *Velocity* (*v*) digunakan untuk menggerakkan proses optimisasi, di mana *particle* harus berpindah tempat untuk memperbaiki posisinya.
- f. *Inertia weight* (θ) adalah parameter yang digunakan untuk mengatur efek dari *velocity* yang diberikan oleh suatu *particle*, disimbolkan dengan *w*.
- g. *Learning Rates* (*c1* dan *c2*) adalah konstanta yang menilai kemampuan individual *particle* (*c1*) dan kemampuan sosial *swarm* (*c2*), yang menunjukkan nilai *particle* terhadap memorinya.

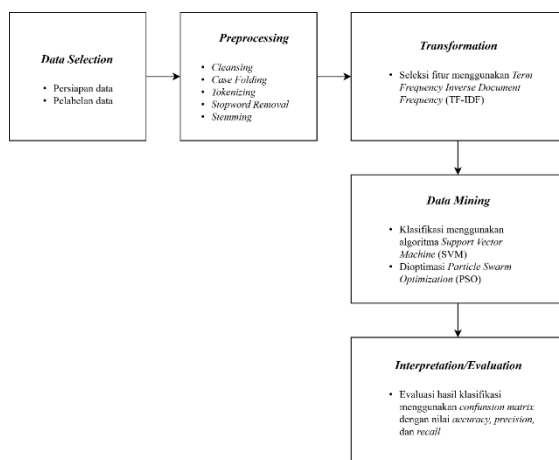
2.4. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah suatu proses analisis terstruktur yang digunakan untuk mendapatkan informasi yang akurat dan baru serta menemukan pola dalam kumpulan data yang besar dan kompleks [10]. *Data mining*, yang merupakan inti dari proses *Knowledge Discovery in Database (KDD)*, melibatkan penggunaan algoritma tertentu untuk menggali data, membentuk model, dan menemukan pola yang sebelumnya tidak diketahui. *Knowledge Discovery in Database (KDD)* adalah suatu proses berulang yang interaktif yang terdiri dari pemilihan, *preprocessing*, transformasi, *data mining*, dan evaluasi (interpretasi).

3. METODE PENELITIAN

3.1. Metodologi Penelitian

Penelitian ini menerapkan metode KDD. Berikut merupakan gambar 3.1 yang merupakan alur tahapan penelitian yang akan dilakukan.



Gambar 3.1 Alur Tahapan Penelitian

Gambar 3.1 menggambarkan proses penelitian yang meliputi beberapa tahapan. Tahap awal adalah pemilihan data (*data selection*), yang terdiri dari persiapan data dan pelabelan data. Setelah itu, dilanjutkan dengan tahap *preprocessing* yang mencakup *cleansing*, *case folding*, *stopword removal*, dan *stemming*. Selanjutnya, terdapat tahap transformasi menggunakan metode seleksi fitur TF-IDF. Tahap berikutnya adalah tahap data mining, yang mencakup pengujian SVM dan pengujian SVM dengan optimasi PSO. Tahap terakhir

adalah evaluasi menggunakan *confusion matrix* dengan memperhitungkan nilai *accuracy*, *precision*, *recall*, dan *f1-score*.

3.2. Rancangan Penelitian

Berikut ini adalah tahapan rancangan penelitian *Knowledge Discovery in Database (KDD)* yang akan digunakan dalam penelitian ini adalah sebagai berikut:

3.2.1. Pemilihan Data (*Data Selection*)

Pada tahap ini, persiapan dan pemilihan data akan dilakukan. Proses selanjutnya akan menggunakan data yang dikumpulkan dari google play store yang mencakup ulasan pengguna aplikasi investasi populer di Indonesia, yaitu Ajaib, Bareksa, Bibit, HSB Investasi, dan Tanamduit yang dikumpulkan dari bulan Januari 2022 hingga bulan Desember 2023 menggunakan teknik *web scraping* dengan memanfaatkan Google Colaboratory sebagai alat bantu.

3.2.2. Preprocessing (*Cleaning*)

Tahap *preprocessing* sangat krusial untuk mempersiapkan data sebelum masuk ke tahap klasifikasi. Proses *preprocessing* terdiri dari sejumlah langkah, termasuk membersihkan (*cleansing*), melipat kasus (*case folding*), *tokenizing*, penghapusan kata penghubung (*stopword removal*), dan pengembalian kata dasar (*stemming*) [11].

3.2.3. Transformasi

Dalam tahap transformasi, teknik seleksi fitur menggunakan TF-IDF (*Term Frequency Inverse Document Frequency*) digunakan untuk melakukan seleksi fitur. Ini dilakukan dengan memberikan bobot pada frekuensi kata-kata yang muncul dalam dokumen. Metode *Term Frequency-Inverse Document Frequency (TF-IDF)* ini umum dipakai untuk melihat seberapa jauh hubungan kata terhadap dokumen dengan memberikan bobot atau nilai dalam setiap kata. Pada penelitian yang dilakukan oleh Natasuwarna, 2020 menunjukkan bahwa *accuracy* tertinggi dalam penelitian tersebut menggunakan TF-IDF [12].

3.2.4. Data Mining

Pada tahap ini, dilakukan tahap klasifikasi terhadap ulasan pengguna aplikasi investasi. Klasifikasi ini melibatkan dua proses utama,

yaitu proses *training* dan proses testing. Pada proses *training*, digunakan *training set* yang sudah memiliki label-label yang diketahui untuk membangun model klasifikasi. Model ini akan didasarkan pada pola-pola yang ditemukan dalam data pelatihan. Setelah model terbentuk, langkah selanjutnya adalah melakukan proses testing. Proses testing bertujuan untuk mengevaluasi seberapa akurat model yang telah dibangun pada proses *training*. Dalam tahap ini, data yang sudah diubah bentuknya kemudian diuji dengan membaginya menjadi data latih dan data uji. Pembagian data dilakukan dengan menggunakan empat skenario yang berbeda. Berikut ini ilustrasi proses modelling dapat dilihat pada tabel 3.1 di bawah ini.

Tabel 3.1 Ilustrasi Splitting Data

Skenario	Persentase	Data Latih	Data Uji
1	90% dan 10%	540	60
2	80% dan 20%	480	120
3	70% dan 30%	420	180
4	60% dan 40%	360	240

3.2.5. Evaluasi (Interpretation)

Pada tahap ini, evaluasi dilakukan untuk mengukur hasil klasifikasi dengan memperhitungkan nilai akurasi (*accuracy*), presisi (*precision*), dan *recall*. Metode evaluasi ini memanfaatkan *confusion matrix*, yang terdiri dari empat nilai utama sebagai dasar perhitungan, yakni *true positive rate* (TP rate), *true negative rate* (TN rate), *false positive rate* (FP rate), dan *false negative rate* (FN rate) [13].

4. HASIL DAN PEMBAHASAN

Hasil penelitian yang telah dilakukan adalah menganalisis sentimen melalui klasifikasi ulasan pengguna aplikasi investasi, pada aplikasi Ajaib, Bareksa, Bibit, HSB Investasi, dan Tanamduit. Algoritma yang digunakan adalah *Support Vector Machine* (SVM) dalam mengolah data teks dari ulasan pengguna di google play store dan kemudian dioptimasi menggunakan *Particle Swarm Optimization* (PSO). Pengumpulan data dilakukan melalui teknik web *scraping* dengan memanfaatkan Google Colaboratory sebagai alat bantu. Gambar 4.1 merupakan contoh data yang telah diambil menggunakan teknik *scraping* dan disimpan dalam format file .xlsx. Hasil

scraping tersebut menunjukkan bahwa terdapat sebanyak 1194 data ulasan yang berhasil diekstrak dari aplikasi Ajaib.

Gambar 4.1 Hasil Scraping dalam Format .xlsx

Selain itu, aplikasi Bareksa berhasil mengekstrak 995 data ulasan, sementara aplikasi Bibit menghasilkan 1194 data ulasan yang berhasil diekstrak. Sebanyak 995 data ulasan berhasil diperoleh dari aplikasi HSB Investasi. Selain itu, aplikasi Tanamduit menghasilkan 1592 data ulasan yang berhasil diekstrak. Data tersebut merupakan data ulasan aplikasi pada situs web google play store dalam kategori data paling relevan selama bulan Januari 2022 hingga bulan Desember 2023.

Pada tahapan ini seluruh pengujian yang telah dilakukan akan dibandingkan untuk mengetahui model mana memiliki kinerja terbaik antara *Support Vector Machine* (SVM) dan *Support Vector Machine* (SVM) yang dioptimasi dengan *Particle Swarm Optimization* (PSO). Berikut adalah hasil perbandingan evaluasi untuk setiap skenario yang tercantum dalam Tabel 4.1 di bawah ini.

Tabel 4.1 Perbandingan Evaluasi Model SVM

Support Vector Machine								
Nama Apk	Sken ario	Accur acy	Presisi		Recall		F1-score	
			0	1	0	1	0	1
Ajaib	90:10	82,59 %	79 %	97 %	95 %	88 %	86 %	92 %
	80:20	83,33 %	81 %	90 %	86 %	86 %	83 %	88 %
	70:30	83,33 %	80 %	91 %	88 %	85 %	84 %	88 %
	60:40	86,11 %	77 %	84 %	77 %	84 %	77 %	84 %
Bareks a	90:10	83,15 %	68 %	87 %	60 %	91 %	64 %	90 %
	80:20	81,67 %	75 %	88 %	60 %	93 %	67 %	90 %

Nama Apk	Skenario	Support Vector Machine						
		Accuracy	Presisi		Recall		F1-score	
			0	1	0	1	0	1
	70:30	82,86 %	81 %	87 %	58 %	96 %	68 %	91 %
	60:40	80,28 %	71 %	89 %	58 %	94 %	64 %	91 %
Bibit	90:10	82,59 %	78 %	86 %	86 %	77 %	82 %	81 %
	80:20	82,08 %	88 %	84 %	84 %	88 %	86 %	86 %
	70:30	80,00 %	82 %	78 %	78 %	82 %	80 %	80 %
	60:40	81,11 %	79 %	76 %	74 %	81 %	76 %	79 %
HSB Investasi	90:10	90,93 %	56 %	92 %	56 %	92 %	56 %	92 %
	80:20	92,08 %	68 %	94 %	68 %	94 %	68 %	94 %
	70:30	92,38 %	69 %	94 %	69 %	94 %	69 %	94 %
	60:40	90,00 %	71 %	95 %	77 %	94 %	74 %	95 %
Tanamduit	90:10	86,30 %	67 %	92 %	82 %	84 %	74 %	88 %
	80:20	87,29 %	73 %	89 %	77 %	86 %	75 %	87 %
	70:30	87,14 %	73 %	91 %	84 %	84 %	78 %	87 %
	60:40	84,72 %	76 %	92 %	87 %	85 %	81 %	89 %

Berdasarkan tabel 4.8 nilai akurasi tertinggi pada aplikasi Ajaib adalah 86.11% pada skenario ke-4 dengan pembagian data 60:40. Sementara itu, nilai akurasi terendah pada aplikasi Ajaib adalah 82,59% pada skenario ke-1 dengan pembagian data 90:10. Untuk aplikasi Bareksa, nilai akurasi tertinggi adalah 83,15% pada skenario ke-1 dengan pembagian data 90:10, sedangkan nilai akurasi terendahnya adalah 80,28% pada skenario ke-4 dengan pembagian data 60:40. Pada aplikasi Bibit, nilai akurasi tertinggi adalah 82,59% pada skenario ke-1 dengan pembagian data 90:10, sementara nilai akurasi terendahnya adalah 80,00% pada skenario ke-3 dengan pembagian data 70:30. Untuk aplikasi HSB Investasi, nilai akurasi tertinggi adalah 92,38% pada skenario ke-3

dengan pembagian data 70:30, sedangkan nilai akurasi terendahnya adalah 90,00% pada skenario ke-4 dengan pembagian data 60:40. Terakhir, pada aplikasi Tanamduit, nilai akurasi tertinggi adalah 87,29% pada skenario ke-2 dengan pembagian data 80:20, sedangkan nilai akurasi terendahnya adalah 84,72% pada skenario ke-4 dengan pembagian data 60:40. Berikut hasil evaluasi dari 4 skenario dengan menggunakan *Support Vector Machine* (SVM) dioptimasi dengan algoritma *Particle Swarm Optimization* (PSO) dapat dilihat pada tabel 4.2 di bawah ini.

Tabel 4.2 Perbandingan Evaluasi Model SVM dan PSO

Nama Apk	Skenario	Support Vector Machine dan Particle Swarm Optimization						
		Accuracy	Precision		Recall		F1-score	
			0	1	0	1	0	1
Ajaib	90:10	93,33 %	83 %	10 0%	10 0%	90 %	91 %	95 %
	80:20	89,17 %	82 %	95 %	94 %	86 %	88 %	90 %
	70:30	88,89 %	81 %	96 %	95 %	85 %	87 %	90 %
	60:40	89,17 %	82 %	96 %	95 %	85 %	88 %	90 %
Bareksa	90:10	86,67 %	77 %	89 %	67 %	93 %	71 %	91 %
	80:20	85,83 %	78 %	88 %	60 %	94 %	68 %	91 %
	Skenario 3 70:30	84,44 %	76 %	86 %	56 %	94 %	64 %	90 %
	60:40	85,00 %	70 %	88 %	54 %	94 %	61 %	91 %
Bibit	90:10	90,00 %	83 %	10 0%	10 0%	81 %	91 %	89 %
	80:20	90,00 %	86 %	96 %	97 %	83 %	91 %	89 %
	70:30	86,21 %	83 %	92 %	93 %	81 %	88 %	86 %
	60:40	87,08 %	83 %	92 %	92 %	82 %	87 %	87 %
	90:10	91,67 %	83 %	93 %	56 %	98 %	67 %	95 %

Nama Apk	Skenario	Support Vector Machine dan Particle Swarm Optimization					
		Accuracy	Precision		Recall		F1-score
			0	1	0	1	
HSB Investasi	80:20	90,83 %	75 %	93 %	63 %	96 %	69 %
	70:30	90,56 %	73 %	94 %	66 %	95 %	69 %
	60:40	91,16 %	76 %	95 %	74 %	96 %	75 %
Tanamduit	90:10	88,33 %	75 %	95 %	88 %	88 %	81 %
	80:20	87,92 %	78 %	94 %	89 %	87 %	83 %
	70:30	87,22 %	78 %	93 %	87 %	87 %	82 %
	60:40	87,92 %	78 %	94 %	89 %	87 %	83 %

Hasil optimasi menggunakan *Particle Swarm Optimization* (PSO) menunjukkan bahwa nilai akurasi tertinggi untuk semua aplikasi terdapat pada skenario ke-1 dengan pembagian data 90:10. Aplikasi Ajaib mencapai akurasi sebesar 93,33%, Bareksa mencapai 86,67%, Bibit mencapai 90,00%, HSB Investasi mencapai 91,67%, dan Tanamduit mencapai 88,33%. Selain itu, terdapat peningkatan yang konsisten pada nilai evaluasi lainnya seperti *precision*, *recall*, dan *f1-score*. Sebaliknya, untuk beberapa aplikasi, nilai akurasi terendah ditemukan pada skenario ke-3 dengan pembagian data 70:30, di mana Ajaib mencapai 88,89%, Bareksa mencapai 84,44%, Bibit mencapai 86,21%, HSB Investasi mencapai 90,56%, dan Tanamduit mencapai 87,22%.

Berikut adalah hasil pencarian kata-kata yang sering muncul dalam kelas positif dan negatif, yang ditampilkan dalam visualisasi *word cloud* pada gambar-gambar di bawah ini.

4.1.1. Ajaib

Dari gambar 4.2 dapat diketahui bahwa kata yang sering muncul pada *class* positif yaitu kata “mula”, “muda”, “investasi”, “cocok” dan “bagus” sedangkan kata-kata pada *class* negatif seperti “bingung”, “error”, dan “uninstall” merupakan kata yang sering digunakan pengguna aplikasi Ajaib.



Gambar 4.2 Word cloud Positif Aplikasi dan Negatif Ajaib

4.1.2. Bareksa

Dari gambar 4.3 dapat diketahui bahwa kata yang sering muncul pada *class* positif yaitu kata “beli”, “mudah”, “baik”, dan “bagus” sedangkan kata-kata pada *class* negatif seperti “error”, “verifikasi”, dan “gagal” merupakan kata yang sering digunakan pengguna aplikasi Bareksa.



Gambar 4.3 Word cloud Positif Aplikasi dan Negatif Bareksa

4.1.3. Bibit

Dari gambar 4.4 dapat diketahui bahwa kata yang sering muncul pada *class* positif yaitu kata “baik”, “investasi”, “cepat”, dan “bagus” sedangkan pada *class* negatif seperti “kecewa”, “jual”, dan “tolong” merupakan kata yang sering digunakan pengguna aplikasi Bibit.



Gambar 4.4 Word cloud Positif Aplikasi dan Negatif Bibit

4.1.4. HSB Investasi

Dari gambar 4.4 dapat diketahui bahwa kata yang sering muncul pada *class* positif yaitu kata

“trading”, “mudah”, “cepat”, dan “bagus” sedangkan pada *class* negatif seperti “daftar”, “data”, dan “ribet” merupakan kata yang sering digunakan pengguna aplikasi HSB Investasi.



Gambar 4.5 *Word cloud* Positif Aplikasi dan Negatif HSB Investasi

4.1.5. *Tanamduit*

Dari gambar 4.6 dapat diketahui bahwa kata yang sering muncul pada *class* positif yaitu kata “reksadana”, “bagus”, “investasi”, dan “baik” sedangkan pada *class* negatif seperti “tipu”, “lambat”, dan “sulit” merupakan kata yang sering digunakan pengguna aplikasi Tanamduit.



Gambar 4.6 *Word cloud* Positif Aplikasi dan Negatif Tanamduit

5. KESIMPULAN

- a. Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk menganalisis sentimen terhadap aplikasi Ajaib, Bareksa, Bibit, HSB Investasi, dan Tanamduit, dengan menerapkan metodologi *Knowledge Discovery in Database* (KDD) yang meliputi tahapan pemilihan data (data selection), *preprocessing* (cleaning), transformasi, data mining, dan evaluasi (interpretation). Proses ini kemudian dioptimalkan menggunakan *Particle Swarm Optimization* (PSO). Aplikasi Ajaib mendapatkan ulasan positif dari pengguna, terutama karena cocok

untuk pemula. Bareksa juga menerima umpan balik positif, terutama terkait kemudahan layanan. Bibit disukai pengguna karena fitur edukasi dan layanannya yang cepat. HSB Investasi mendapat penilaian baik karena kemudahan dalam trading. Tanamduit mendapatkan ulasan positif untuk layanan asuransi dan reksadananya.

- b. 2. Pemodelan pertama menggunakan algoritma SVM menunjukkan variasi kinerja pada setiap aplikasi. Ajaib mencapai akurasi 86,11%, Bareksa 83,15%, Bibit 82,59%, HSB Investasi 92,38%, dan Tanamduit 87,29%. Aplikasi HSB Investasi menunjukkan hasil terbaik dengan akurasi tertinggi serta presisi, *recall*, dan *f1-score* masing-masing sebesar 94%. Pada pemodelan kedua, 4 skenario diuji menggunakan SVM yang dioptimasi dengan PSO, menunjukkan bahwa skenario ke-1 dengan pembagian data 90:10 memberikan nilai akurasi tertinggi untuk semua aplikasi. Nilai evaluasi lainnya juga meningkat, termasuk akurasi, presisi, *recall*, dan *f1-score*. Aplikasi Ajaib berhasil mencapai hasil terbaik dengan akurasi tertinggi mencapai 93,33%, serta nilai presisi 100%, *recall* 90%, dan *f1-score* 95%. Dengan demikian, kombinasi analisis sentimen ulasan pengguna pada aplikasi Ajaib, Bareksa, Bibit, HSB Investasi, dan Tanamduit menggunakan SVM yang dioptimasi dengan PSO menunjukkan kinerja yang lebih baik dibandingkan penggunaan SVM saja.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] R. Pertiwi, "Analisis Sentimen pada Data Survei Pelanggan untuk Mengukur Kepuasan dan Loyalti Pelanggan pada Industri Layanan," *duniabisnis.org*, vol. 3, no. 4, 2023.
- [2] F. V. Sari and A. Wibowo, "Analisis sentimen pelanggan toko online Jd. Id menggunakan metode Naïve Bayes Classifier berbasis konversi ikon emosi," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 10, no. 2, pp. 681–686, 2019.
- [3] A. P. R. Nababan, A. S. M. Lumenta, Y. D. Y. Rindengan, F. J. Pontoh, and Y. V. Akay, "Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode Lexicon Analysis," *Journal Unsrat.ac.id*, vol. Vol 15, pp. 33–34, 2020.
- [4] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, Jan. 2022, doi: 10.23960/jitet.v10i1.2262.
- [5] M. Muhathir, M. H. Santoso, and D. A. Larasati, "Wayang Image Classification Using SVM Method and GLCM Feature Extraction," *Journal Of Informatics And Telecommunication Engineering*, vol. 4, no. 2, pp. 373–382, Jan. 2021, doi: 10.31289/jite.v4i2.4524.
- [6] F. F. Irfani, "ANALISIS SENTIMEN REVIEW APLIKASI RUANGGURU MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," *JBMI (Jurnal Bisnis, Manajemen, dan Informatika)*, vol. 16, no. 3, pp. 258–266, Feb. 2020, doi: 10.26487/jbmi.v16i3.8607.
- [7] R. Abdulhakim, Y. Umaidah, and J. Haerul Jaman, "Optimasi *Support Vector Machine* Berbasis *Particle Swarm Optimization* Untuk Mendeteksi Hate Speech Pilkada Karawang," 2021. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [8] A. Agung Putra Adnyana, I. Made Widiartha, A. Muliantara, L. Gede Astuti, M. Agung Raharja, and I. Dewa Made Bayu Atmaja Darmawan, "Implementasi Metode Hybrid *Particle Swarm Optimization* dan Genetic Algorithm Pada Penjadwalan Job Shop Scheduling," *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 11, 2023.
- [9] H. Zulfia Zahro and F. Santi Wahyuni, "OPTIMASI PARTICEL SWARM OPTIMAZATION (PSO) UNTUK PENENTUAN BASE TRANCIVIER SYSTEM (BTS)," 2020.
- [10] K. Hamidah and A. Voutama, "Analisis Faktor Tingkat Kebahagiaan Negara Menggunakan Data World Happiness Report dengan Metode Regresi Linier.," *Explore IT: Jurnal Keilmuan dan Aplikasi Teknik Informatika*, vol. 15, no. 1, pp. 1–7, 2023.
- [11] N. Purwati and A. Dwi Januanti, "Aplikasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Tingkat Kelulusan Mahasiswa," *Jurnal Pepadun*, vol. 2, no. 1, pp. 123–137, Apr. 2021, doi: 10.23960/pepadun.v2i1.38.
- [12] A. P. Natasuwarna, "Seleksi Fitur *Support Vector Machine* pada Analisis Sentimen Keberlanjutan Pembelajaran Daring," *Techno. Com*, vol. 19, no. 4, pp. 437–448, 2020.
- [13] E. Sutoyo and A. Almaarif, "Data Mining untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme Naïve Bayes Classifier," *Jurnal Elektronik Ikatan Ahli Informatika Indonesia (IAII)*, vol. 1, no. 3, pp. 95–101, 2020.