



Optimasi Algoritma Random Forest dengan PSO untuk Peningkatan Akurasi Klasifikasi Kampanye Digital

Mochamad Fachri Ardiansyah¹, Hendra Maulana², Firza Prima Aditiawan³

¹²³Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya,
Indonesia

Email: ardiansyahfachri15@gmail.com^{*1}; hendra.maulana.if@upnjatim.ac.id²; firzaprima.if@upnjatim.ac.id³

Ardiansyah, M. F., Maulana, H., & Aditiawan, F. P. (2026). Optimasi Algoritma Random Forest dengan PSO untuk Peningkatan Akurasi Klasifikasi Kampanye Digital. *Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, 12(2), 291-299

DOI: <https://doi.org/10.33050/cerita.v12i2.3810>

ABSTRAK

Kampanye digital marketing semakin kompleks dengan volume data yang besar dan beragam. Penelitian ini mengusulkan metode inovatif untuk mengklasifikasikan keberhasilan kampanye digital marketing dengan mengoptimalkan model Random Forest dengan Particle Swarm Optimization (PSO). Jurnal ini bertujuan dengan peningkatan akurasi dan mengurangi overfitting dalam analisis data kampanye. Dengan menggunakan dataset digital marketing yang terdiri dari 8000 data dan 20 variabel, penelitian membandingkan performa Random Forest standar dengan Random Forest yang dioptimasi PSO. Hasil menunjukkan bahwa pendekatan PSO + Random Forest secara konsisten memberikan peningkatan kinerja, dengan akurasi tertinggi mencapai 90,19%, precision 89,36%, dan F1 Score 88% pada pembagian data 60%:40%. Metode ini efektif mengurangi kesalahan prediksi dan menghasilkan model klasifikasi yang lebih stabil. Penelitian memberikan kontribusi dalam pengembangan metode analisis data digital marketing dan menunjukkan potensi optimasi algoritma machine learning untuk pengambilan keputusan berbasis data yang lebih akurat.

Kata kunci: Digital Marketing, Random Forest, Particle Swarm Optimization, Klasifikasi, Optimasi Hyperparameter

ABSTRACT

Digital marketing campaigns are becoming increasingly complex with large and diverse data volumes. This research proposes an innovative method for classifying digital marketing campaign success by optimizing the Random Forest model with Particle Swarm Optimization (PSO). The purpose of this study to improve prediction accuracy and reduce overfitting in campaign data analysis. Using a digital marketing dataset consisting of 8000 data points and 20 variables, the research compares the performance of standard Random Forest with PSO-optimized Random Forest. Results show that the PSO + Random Forest approach consistently provides performance improvement, with the highest accuracy reaching 90.19%, precision of 89.36%, and F1 Score of 88% in the 60%:40% data split. This method effectively reduces prediction errors and produces a more stable classification model. The research contributes to the development of digital marketing data analysis methods and demonstrates the potential of machine learning algorithm optimization for more accurate data-driven decision-making.

Keywords: *Digital Marketing, Random Forest, Particle Swarm Optimization, Classification, Hyperparameter Optimization.*

I. PENDAHULUAN

Pemasaran digital menjadi salah satu solusi alternatif bagi para pelaku bisnis untuk menunjang aktivitas pemasaran produk dan jasa agar dapat mendistribusikan produk dan jasa dengan lebih cepat dan biaya yang relatif terjangkau (Aji & Suprianto, 2023). Kampanye pemasaran digital sekarang menjadi strategi utama untuk menjangkau konsumen secara lebih luas dan efektif. Dengan memanfaatkan platform digital seperti media sosial, situs web, dan pemasaran email, perusahaan dapat mengiklankan jasa atau layanan yang lebih personal dan interaktif. Indikator seperti jumlah klik, tingkat konversi, dan interaksi pengguna menjadi penting dalam menilai kesuksesan kampanye ini. Namun, di balik potensi besar kampanye digital marketing, tantangan utama yang dihadapi adalah bagaimana menganalisis dan mengklasifikasikan data yang dihasilkan untuk mengukur keberhasilannya. Data yang dihasilkan dari kampanye digital seringkali memiliki dimensi besar dan kompleks, mencakup berbagai metrik seperti performa iklan, perilaku pengguna, demografi, dan waktu kampanye. Kerumitan data ini memerlukan pendekatan analisis yang lebih canggih agar dapat dimanfaatkan secara efektif untuk menilai kesuksesan kampanye dan memprediksi hasilnya.

Klasifikasi merupakan proses untuk mengatur dan memberi arti informasi yang berguna untuk menentukan atau menetapkan kesesuaian peristiwa, gagasan, barang, dan orang. Model klasifikasi memperoleh beberapa fungsi pemetaan yang efektif dari data pelatihan

dan memprediksi kategori dengan bantuan fungsi pemetaan yang dibuat untuk data baru (Angela & Sukmono, 2023). Menurut tugas klasifikasinya, klasifikasi dibagi menjadi dua jenis, yaitu klasifikasi biner, yang hanya menghasilkan dua kemungkinan, seperti ramalan cuaca, dan klasifikasi multi-label, dengan lebih dari dua kemungkinan, seperti mengklasifikasikan kinerja siswa menjadi sangat baik, baik, cukup, atau buruk.

Salah satu algoritma populer dalam klasifikasi data adalah Random Forest, algoritma pembelajaran mesin berdasarkan pohon keputusan. Random Forest sangat tangguh saat menangani kumpulan data besar dan variabel kompleks. Algoritma ini beroperasi dengan mengonstruksi sejumlah besar pohon keputusan yang membentuk ensemble dan menghasilkan prediksi berdasarkan prediksi mayoritas dari seluruh pohon. Random Forest terkenal karena kemampuannya menghasilkan prediksi yang akurat dan tahan terhadap masalah overfitting.

Untuk mengatasi hal ini, Particle Swarm Optimization (PSO), algoritma pengoptimalan berbasis populasi terinspirasi oleh perilaku kawanan burung yang dapat diterapkan. PSO bekerja dengan mengoptimalkan solusi melalui kelompok partikel yang bergerak di dalam ruang pencarian, memperbarui posisinya berdasarkan pengalaman terbaik mereka dan seluruh kawanan. Algoritma ini telah terbukti efektif dalam mengoptimalkan hyperparameter pada model pembelajaran mesin, termasuk Random Forest.

Penggunaan PSO dalam optimasi hyperparameter di Random Forest bertujuan untuk meningkatkan akurasi klasifikasi, terutama dalam analisis data kampanye digital marketing

yang rumit. Dengan mengoptimalkan hyperparameter, PSO dapat menemukan konfigurasi model yang meningkatkan akurasi prediksi dan mengurangi overfitting. PSO juga dapat digunakan untuk pemilihan fitur, memilih fitur yang paling relevan dalam dataset agar model berjalan lebih efisien.

Penelitian ini bertujuan untuk memberikan kontribusi dalam pengembangan metode analisis data digital marketing dengan mengimplementasikan PSO untuk optimasi model Random Forest. Hasilnya diharapkan dapat membantu perusahaan memaksimalkan keberhasilan kampanye mereka melalui model klasifikasi agar lebih akurat dan efektif. Penelitian ini juga menambah pemahaman kita tentang penerapan metode optimasi dalam analisis data yang kompleks, serta memberikan

solusi praktis dalam pengambilan keputusan berbasis data.

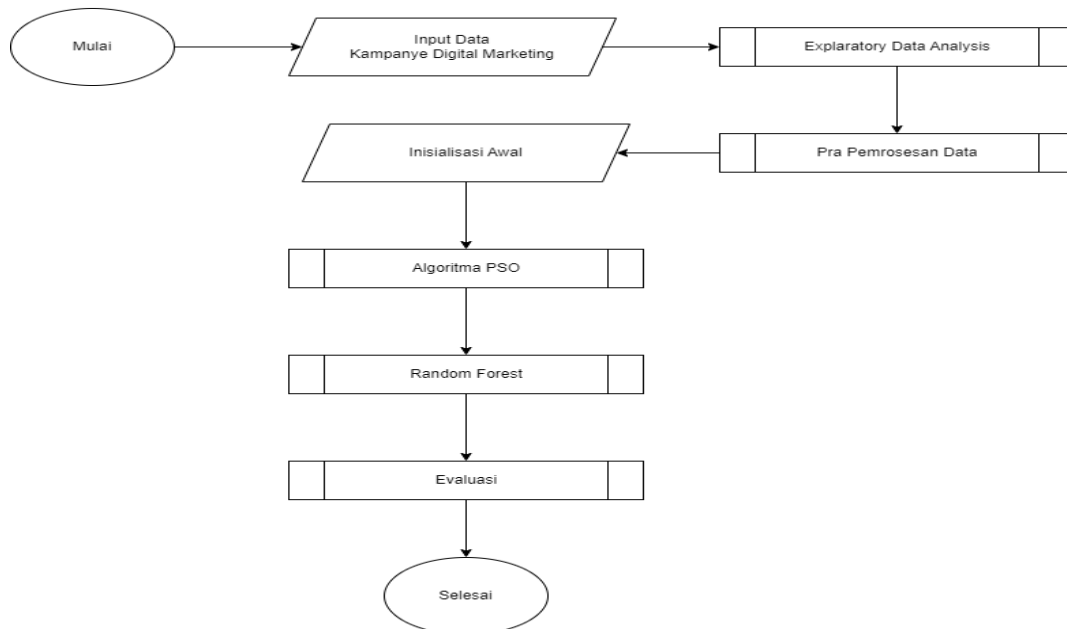
Secara keseluruhan, penelitian ini menggabungkan PSO dan Random Forest untuk mengatasi tantangan dalam analisis data kampanye digital marketing, dengan tujuan meningkatkan kualitas pengambilan keputusan berdasarkan hasil kampanye.

II. METODE PENELITIAN

A. Desain Sistem

Adapun desain rancangan dari sistem klasifikasi keberhasilan kampanye digital marketing menggunakan algoritma PSO dan Random Forest ditunjukkan di gambar 1.

Pada gambar 1 yang merupakan desain sistem penelitian, maka dapat dijelaskan tahapan – tahapan dari alur desain sistem tersebut.



Gambar 1. Desain Sistem

B. Input Data

Pengambilan data akan dilakukan di website. Dalam tahapan awal dilakukan input data csv yang sudah diambil dari platform Kaggle. Data yang digunakan adalah data “Digital Marketing Company” dengan terdapat 20 kolom yang berjumlah 8000 data. Tabel 1 mendeskripsikan 20 kolom yang terdapat pada dataset tersebut.

Tabel 1. Dataset Kampanye Digital Marketing

Nama Kolom	Jenis Data
Customer_ID	Integer
Age	Integer
Gender	String

Income	Integer
CampaignChannel	String
CampaignType	String
AdSpend	Float
ClickThroughRate (CTR)	Float
ConversionRate	Float
WebsiteVisits	Integer
PagesPerVisit	Float
TimeOnSite	Float
SocialShares	Integer
EmailOpens	Integer
EmailClicks	Integer
PreviousPurchases	Integer

<i>LoyaltyPoints</i>	Integer
<i>AdvertisingPlatform</i>	String
<i>AdvertisingTool</i>	String
<i>Conversion</i>	Boolean

C. Exploratory Data Analysis

Pada dataset yang digunakan terdapat cukup banyak kolom yang dimiliki. Untuk mendapatkan eksplorasi atau pemahaman yang lebih lanjut terhadap dataset maka dilakukan exploratory data analysis. Di mana dalam metode ini akan melibatkan beberapa teknik statistik dan visualisasi data yang diterapkan untuk mengenali pola, tren, serta karakteristik menarik dalam dataset. Adapun beberapa contoh dari penerapan exploratory data analysis ini sebagai berikut

- Membangun plot KDE (Distribusi Kernel Density), untuk melihat penyebaran distribusi variabel data
- Menampilkan korelasi atau hubungan antar variabel menggunakan heatmap

Tabel 2. Missing Value

kolom	hasil
Customer_ID	<i>null</i>
Age	<i>null</i>
Gender	<i>null</i>
Income	<i>null</i>
<u>CampaignChannel</u>	<i>null</i>
CampaignType	<i>null</i>
AdSpend	<i>null</i>
<u>ClickThroughRate (CTR)</u>	<i>null</i>
<u>ConversionRate</u>	<i>null</i>
<u>WebsiteVisits</u>	<i>null</i>
<u>PagesPerVisit</u>	<i>null</i>
TimeOnSite	<i>null</i>
<u>SocialShares</u>	<i>null</i>
<u>EmailOpens</u>	<i>null</i>
<u>EmailClicks</u>	<i>null</i>
<u>PreviousPurchases</u>	<i>null</i>
<u>LoyaltyPoints</u>	<i>null</i>
<u>AdvertisingPlatform</u>	<i>null</i>
<u>AdvertisingTool</u>	<i>null</i>
Conversion	<i>null</i>

Dalam tabel 2. menunjukkan bahwa hasil filter missing value tidak ditemukan pada seluruh

variabel. Selain itu, dalam pra-pemrosesan data, kumpulan data juga dibagi menjadi data pelatihan dan data uji untuk mengevaluasi model secara lebih akurat. Tahapan pra- pemrosesan ini sangat penting untuk meningkatkan kualitas data dan, pada akhirnya, meningkatkan performa model.

D. Pra-Pemrosesan Data

Setelah melakukan pemahaman dataset yang digunakan, dilanjutkan pra-pemrosesan data untuk mengolah data sebelum data digunakan sebagai pengujian. Metode ini berguna untuk memastikan data yang digunakan telah dipersiapkan dan berkualitas tinggi sebelum dimasukkan ke dalam model analitis atau algoritma pembelajaran mesin. Proses ini melibatkan beberapa langkah, dimulai dengan pembersihan data, yang melibatkan penghapusan atau pemrosesan data yang hilang, duplikat, atau tidak valid. Kemudian, dilakukan transformasi data, seperti normalisasi atau standarisasi, untuk menyelaraskan skala variabel sehingga tidak ada variabel yang mendominasi proses pembelajaran. Adapun tabel 3. menunjukkan contoh penerapan dalam memfilter missing value data.

E. Inisialisasi Awal

Setelah melakukan pembagian data latih dan data uji maka, maka selanjutnya menginisialisasikan parameter yang digunakan dalam melatih model. Dimana menentukan iterasi maksimal PSO, range, populasi dan parameter Random Forest.

F. Algoritma Particle Swarm Optimization (PSO)

Dalam penerapan algoritma PSO dapat berfungsi mengoptimalkan hyperparameter dari model Random Forest. Partikel PSO mempresentasikan set hyperparameter yang berbeda, dan setiap partikel dievaluasi berdasarkan kinerja klasifikasi model pada set pelatihan. Posisi partikel di update berdasarkan hasil evaluasi hingga ditemukan set hyperparameter yang optimal.

G. Random Forest

Selanjutnya, penerapan model klasifikasi dengan menggunakan random forest. Langkah awal dalam model klasifikasi ialah memisahkan antara fitur dan label. Kemudian membangun model random forest dengan pengklasifikasian yang berbentuk struktur pohon yang hampir sama dengan model decision tree dalam membangun pohon. Namun, model random forest membangun pohon lebih dari satu pohon. Pada dasarnya,

model Random Forest ini terdiri dari beberapa decision tree yang masing-masing dilatih pada data hasil resampling dengan teknik Bootstrap Sampling. Bootstrap sampling adalah metode pengambilan sampel dengan pengembalian (replacement) dari dataset asli. Teknik ini penting dilakukan untuk menciptakan variasi dalam pelatihan model tanpa harus menggunakan dataset tambahan. Ini membantu menghasilkan model yang lebih stabil, akurat, dan dapat menangani overfitting lebih baik dibandingkan model yang hanya dilatih pada satu set data yang sama.

H. Evaluasi Model

Setelah melakukan pengujian model dan mendapatkan hasil prediksi perlu dilakukan evaluasi untuk menilai kualitas dan performa model yang telah di uji. Evaluasi dilakukan dengan menganalisis hasil dari F1-Score dan dari grafik ROC agar lebih mudah untuk mengevaluasi kedua model.

Evaluasi menggunakan kurva ROC-AUC merupakan metode penting untuk menilai kinerja

model klasifikasi dalam penelitian ini. Kurva ROC (Receiver Operating Characteristic) memvisualisasikan kemampuan pada model untuk pembeda antara kelas positif dan kelas negatif melalui pengukuran True Positive Rate (TPR) dan False Positive Rate (FPR). Area Under the Curve (AUC) menjadi indikator utama dalam mengevaluasi performa model, dengan nilai mendekati 1 menunjukkan kemampuan diskriminasi yang sangat baik, sementara nilai 0.5 mengindikasikan model tidak lebih baik dari prediksi acak.

Proses evaluasi dimulai dengan menghitung probabilitas prediksi untuk setiap kelas dari model Random Forest baik hanya dengan random forest saja maupun yang dioptimasi dengan PSO. Selanjutnya, kurva ROC dibuat dengan memplot TPR terhadap FPR, yang memungkinkan visualisasi performa model pada berbagai ambang batas klasifikasi. Luas area di bawah kurva (AUC) dihitung untuk memberikan penilaian kuantitatif tentang kualitas model dalam membedakan kelas.

I. Skenario Pengujian

No.	Algoritma	Data Latih	Data Uji	Akurasi	Recall	Precision	F1 Score	ROC-AUC	Waktu Eksekusi
1.	Random Forest	50%	50%	-					
		60%	40%	-					
		70%	30%						
		80%	20%						
		90%	10%						
2.	Algoritma PSO + Random Forest	50%	50%	-					
		60%	40%	-					
		70%	30%						
		80%	20%						
		90%	10%						

Tabel 3. Skenario Pengujian

Pada tabel 3. skenario pengujian dilakukan dengan cara membagi rasio menjadi beberapa, guna untuk mengetahui apakah optimasi PSO bekerja dengan baik dan optimal.

III. HASIL DAN PEMBAHASAN

A. Persiapan dan Pengolahan Data

Pada studi ini, persiapan dan pengolahan data merupakan langkah fundamental untuk memastikan kualitas dan keandalan model yang akan dibuat. Proses ini mencakup serangkaian langkah sistematis mulai dari pemahaman

karakteristik dataset sampai transformasi data ke dalam format yang sesuai dengan model yang akan digunakan. Tahap ini sangat krusial karena kualitas data sangat berpengaruh dalam performa model machine learning yang akan dikembangkan.

B. Deskripsi Dataset

Dataset merupakan data kampanye digital marketing yang kurang lebih sebanyak 8000 sampel data dengan 20 atribut, termasuk didalamnya 1 atribut target yaitu 'Conversion'. Dataset ini mencakup berbagai aspek dai kampanye digital marketing seperti informasi pelanggan yang terdiri dari (Age, Gender, Income), metrics kampanye terdiri (AdSpend, CTR, ConversionRate), dan metrics engagement berupa (WebsiteVisits, PagesPerVisit, TimeOnSite, dan SocialShares).

Atribut-atribut dalam dataset dapat dikategorikan menjadi beberapa kelompok utama :

1. Data Demografis: CustomerID, Age, Gender, Income
2. Informasi Kampanye: CampaignChannel, CampaignType, AdSpend
3. Metrics Performa: ClickThroughRate, ConversionRate
4. Metrics Engagement: WebsiteVisits, PagesPerVisit, TimeOnSite
5. Metrics Sosial dan Email: SocialShares, EmailOpens, EmailClicks
6. Data Loyalitas: PreviousPurchases, LoyaltyPoints
7. Platform Advertising: AdvertisingPlatform, AdvertisingTool
8. Target Variable: Conversion

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Customer	Age	Gender	Income	Campaign	Campaign	AdSpend	ClickThro	Conversio	WebsiteVi	PagesPerV	TimeOnSit	SocialShar	EmailOpen	EmailClick	PreviousP	LoyaltyPoi	Advertisin	Advertisin	Conversion
2	8000	56	Female	136912	Social Me	Awarenes	6497.87	0.043919	0.088031	0	2.399017	7.396803	19	6	9	4	688	IsConfid	ToolConf	1
3	8001	69	Male	41760	Email	Retention	3898.669	0.155725	0.182725	42	2.917138	5.352549	5	2	7	2	3459	IsConfid	ToolConf	1
4	8002	46	Female	88456	PPC	Awarenes	1546.43	0.27749	0.076423	2	8.223619	13.7949	0	11	2	8	2337	IsConfid	ToolConf	1
5	8003	32	Female	44085	PPC	Conversio	539.5259	0.137611	0.088004	47	4.540939	14.68836	89	2	2	0	2463	IsConfid	ToolConf	1
6	8004	60	Female	83964	PPC	Conversio	1678.044	0.252851	0.10994	0	2.046847	13.99337	6	6	6	8	4345	IsConfid	ToolConf	1
7	8005	25	Female	42925	Social Me	Awarenes	9579.388	0.153795	0.161316	6	2.12585	7.752831	95	5	8	0	3316	IsConfid	ToolConf	1
8	8006	38	Female	25615	Referral	Awarenes	7302.9	0.040975	0.060977	42	1.753995	10.69867	54	14	3	6	930	IsConfid	ToolConf	1
9	8007	56	Female	57083	Social Me	Conversio	5324.284	0.052878	0.188946	48	2.626015	2.987817	96	9	3	0	2983	IsConfid	ToolConf	1
10	8008	36	Female	140788	Email	Retention	9421.251	0.023536	0.112585	13	5.472843	14.28742	73	4	8	5	460	IsConfid	ToolConf	1
11	8009	40	Male	130764	Social Me	Awarenes	6229.193	0.066641	0.169786	22	1.135665	4.613312	14	8	4	8	3789	IsConfid	ToolConf	1
12	8010	28	Male	88319	PPC	Retention	7283.887	0.269148	0.109868	47	2.054181	3.29472	94	16	5	6	1987	IsConfid	ToolConf	1
13	8011	28	Male	29605	SEO	Awarenes	2898.253	0.198557	0.156679	16	9.164109	1.372299	23	18	6	5	2403	IsConfid	ToolConf	1
14	8012	41	Male	131093	Email	Conversio	5930.92	0.224269	0.01977	13	8.889232	8.480384	28	16	5	4	2946	IsConfid	ToolConf	1
15	8013	53	Male	64608	SEO	Awarenes	1520.425	0.190282	0.16185	8	8.497723	14.25272	81	8	1	2	2818	IsConfid	ToolConf	1
16	8014	57	Male	148788	Email	Awarenes	4104.383	0.122096	0.109552	20	7.356535	2.46603	70	16	8	1	3247	IsConfid	ToolConf	1
17	8015	41	Female	39192	Email	Awarenes	429.0792	0.07289	0.095389	44	6.953695	2.252039	33	0	1	1	4474	IsConfid	ToolConf	0
18	8016	20	Female	23331	PPC	Conversio	2244.618	0.278428	0.031206	1	3.479642	4.452666	77	7	8	8	3252	IsConfid	ToolConf	1
19	8017	39	Female	71330	Social Me	Conversio	841.1994	0.159301	0.150897	31	4.014951	10.14821	83	18	5	2	3463	IsConfid	ToolConf	1
20	8018	19	Female	117392	PPC	Conversio	3939.01	0.149519	0.034921	7	3.12222	8.443296	92	8	4	6	2222	IsConfid	ToolConf	1

Gambar 2. Tampilan Dataset dalam Bentuk Excel

Dataset ini memiliki karakteristik yang seimbang dalam hal distribusi kelas target, yang memungkinkan model untuk mempelajari pola dari kedua subset yaitu data-test dan data-train. Selain itu, dataset ini mencakup beragam aspek penting dalam digital marketing, yang mendukung analisis menyeluruh.

Pada tahap awal, dataset ini menunjukkan kualitas data yang cukup baik, ditandai dengan tidak adanya nilai yang hilang maupun data duplikat. Dengan adanya kondisi tersebut dataset ini memiliki fondasi yang kokoh untuk mendukung proses analisis data dan pengembangan model yang akan dilakukan. Namun beberapa variabel kategorikal, seperti

Gender, CampaignChannel, CampaignType, AdvertisingPlatform, dan AdvertisingTool, memerlukan langkah preprocessing lanjutan agar dapat diolah secara efektif dalam algoritma machine learning.

C. Evaluasi Skenario Pengujian

Setelah dilakukan pengujian pada kedua model Random Forest dan Random Forest dengan optimasi PSO, dilakukan evaluasi sesuai skenario pengujian yang telah ditentukan sebelumnya dengan menguji semua model yang telah dilatih, seperti yang telah disajikan dalam Tabel 4.

Tabel 4. Hasil Skenario Pengujian

No.	Algoritma	Data Latih	Data Uji	Akurasi	Recall	Precision	F1 Score	AUC	Waktu Eksekusi
1.	Random Forest	50%	50%	80%	80%	87.5%	82%	78.21%	21.51 detik
		60%	40%	80.34%	80.34%	87.38%	83%	79.12%	23.34 detik
		70%	30%	81.75%	81.75%	87.85%	84%	80.38%	25.04 detik
		80%	20%	80.50%	80.50%	87.89%	83%	80.95%	27.08 detik
		90%	10%	80.25%	80.25%	87.17%	83%	77.95%	30.63 detik
2.	Algoritma PSO + Random Forest	50%	50%	89.70%	89.70%	89.72%	86%	81.19%	0.48 detik
		60%	40%	89.66%	89.66%	89.69%	86%	81.27%	0.73 detik
		70%	30%	89.83%	89.83%	89.70%	87%	81.03%	0.60 detik
		80%	20%	89.31%	89.31%	88.66%	86%	82.43%	1.66 detik
		90%	10%	89%	89%	87.40%	86%	81.04%	0.97 detik

D. Performa Algoritma Random Forest

Algoritma Random Forest diuji pada berbagai distribusi rasio dari pembagian secara 50% dan 50% hingga 90%:10%. Berdasarkan hasil pengujian, algoritma ini menghasilkan akurasi yang cukup stabil berkisar antara 80% hingga 80.25%. Nilai recall konsisten dengan nilai akurasi pada setiap pembagian data. Precision algoritma Random Forest berkisar antara 87.17% hingga 87.89%, sedangkan F1 Score berada pada rentang 82% hingga 84%.

Pada pembagian data 70%:30%, algoritma Random Forest mencapai performa terbaik dengan akurasi dan recall 81.75%, precision 87.85%, dan F1 Score 84%. Dengan nilai AUC 80.38%, mengindikasikan kemampuan model dalam membedakan antara kelas positif dan negatif.

E. Performa Algoritma PSO + Random Forest

Algoritma PSO + Random Forest menghasilkan performa unggul dibandingkan algoritma Random Forest standar. Akurasi algoritma ini berkisar antara 89% hingga 89.83%, dengan nilai tertinggi dicapai pada pembagian data 70%:30%. Nilai recall konsisten dengan nilai akurasi pada semua distribusi data. Precision algoritma PSO + Random Forest berkisar antara 89.72% hingga 87.40%, sedangkan F1 Score berada pada rentang 86% hingga 87%.

Performa terbaik algoritma PSO + Random Forest dicapai pada pembagian data 70%:30%

dengan akurasi dan recall 89.83%, precision 89.70%, dan F1 Score 87%.

Dengan nilai AUC 81.03%, mengindikasikan peningkatan sebanyak 0.65% dimana model dapat membedakan antara kelas positif dan negatif.

IV. KESIMPULAN

Algoritma PSO + Random Forest secara konsisten menampilkan performa jauh lebih unggul dibanding algoritma Random Forest murni pada semua matrik evaluasi, termasuk akurasi, recall, precision, dan F1 Score. Peningkatan kinerja yang dicapai dengan PSO menegaskan bahwa optimasi parameter menggunakan algoritma ini efektif dalam meningkatkan performa Random Forest. Hasil terbaik diperoleh pada pembagian data 60%:40% dan 70%:30%, menunjukkan bahwa kombinasi ini bekerja optimal dalam skenario tersebut. Kedua algoritma menunjukkan ketahanan terhadap variasi pembagian data, tetapi PSO + Random Forest lebih stabil dalam hal performa. Secara keseluruhan, penggunaan PSO sebagai metode optimasi pada Random Forest memberikan peningkatan yang signifikan, terutama dalam aspek akurasi dan precision.

Pemanfaatan Data yang Lebih Luas dan Variatif Guna memperkuat kemampuan generalisasi model, dianjurkan untuk menggunakan kumpulan data yang lebih komprehensif dan meliputi beragam tipe kampanye pemasaran digital dari bermacam sektor industri.

Selain PSO, dapat dilakukan eksplorasi penggunaan algoritma optimasi lainnya seperti Genetic Algorithm (GA) atau Bayesian Optimization untuk membandingkan efektivitasnya dalam meningkatkan kinerja model. Implementasi model dalam sistem atau platform digital marketing secara real-time dapat memberikan wawasan yang lebih mendalam mengenai efektivitas kampanye dan memungkinkan perbaikan strategi pemasaran secara dinamis.

Secara keseluruhan, optimasi parameter menggunakan PSO terbukti efektif dalam meningkatkan kinerja Random Forest. Peningkatan ini terutama terlihat dalam aspek akurasi dan precision, yang menjadi indikator penting dalam menilai performa model. Dengan hasil yang lebih unggul dan stabil, PSO + Random Forest dapat menjadi alternatif yang lebih handal dibandingkan Random Forest standar, terutama dalam aplikasi yang membutuhkan model dengan tingkat ketepatan tinggi

DAFTAR PUSTAKA

- Aji, P. W. S., & Suprianto, S. (2023). Stroke disease prediction using random forest method. Universitas Muhammadiyah Sidoarjo. <https://doi.org/10.21070/ups.2643>
- Angela, F. C., & Sukmono, T. (2023). Forecasting sales of clothing products with a combination support vector regression and particle swarm optimization algorithm. Universitas Muhammadiyah Sidoarjo. <https://doi.org/10.21070/ups.1832>
- Arifin, T., & Herliana, A. (2018). Optimasi metode klasifikasi dengan menggunakan particle swarm optimization untuk identifikasi penyakit diabetes retinopathy. *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, 4(2), 77–81. <https://doi.org/10.23917/khif.v4i2.6825>
- Ary Prandika Siregar, Dwi Priyadi Purba, Jojo Putri Pasaribu, & Khairul Reza Bakara. (2023). Implementasi algoritma random forest dalam klasifikasi diagnosis penyakit stroke. *Jurnal Penelitian Rumpun Ilmu Teknik*, 2(4), 155–164. <https://doi.org/10.55606/juprit.v2i4.3039>
- Fauzi, A., & Yunial, A. H. (2022). Optimasi Algoritma Klasifikasi Naive Bayes, Decision Tree, K – Nearest Neighbor, dan Random Forest menggunakan Algoritma Particle Swarm Optimization pada Diabetes Dataset. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 8(3), 470. <https://doi.org/10.26418/jp.v8i3.56656>
- Irawan, I., Qisthiano, M. R., Syahril, M., & Jakak, P. M. (2023a). Optimasi Prediksi Kelulusan Tepat Waktu: Studi Perbandingan Algoritma Random Forest dan Algoritma K-NN Berbasis PSO. *Jurnal Pengembangan Sistem Informasi Dan Informatika*, 4(4), 26–35. <https://doi.org/10.47747/jpsii.v4i4.1374>
- Irawan, I., Qisthiano, M. R., Syahril, M., & Jakak, P. M. (2023b). Optimasi Prediksi Kelulusan Tepat Waktu: Studi Perbandingan Algoritma Random Forest dan Algoritma K-NN Berbasis PSO. *Jurnal Pengembangan Sistem Informasi Dan Informatika*, 4(4), 26–35. <https://doi.org/10.47747/jpsii.v4i4.1374>
- Kurniadi, D., Nuraeni, F., Faturrohman, N., & Mulyani, A. (2024). Klasifikasi Perputaran Karyawan Perusahaan Menggunakan Algoritma Random Forest dan Random Over-sampling. *Edu Komputika Journal*, 10(2), 93–103. <https://doi.org/10.15294/edukomputika.v10i2.73782>
- Rahman, B., Fauzi, F., & Amri, S. (2023). Perbandingan Hasil Klasifikasi Data Iris menggunakan Algoritma K-Nearest Neighbor dan Random Forest. *Journal Of Data Insights*, 1(1), 19–26. <https://doi.org/10.26714/jodi.v1i1.135>
- Ramadhan, I. (2022). Prediction of 6 Personality Characters (HEXACO) from Social Media using Random Forest Classifier and Particle Swarm Optimization. *Jurnal Komputasi*, 10(2), 112–125. <https://doi.org/10.23960/komputasi.v10i2.3171>
- Rochman, B. F., Rahim, A., & Siswa, T. A. Y. (2024). Optimasi Algoritma KNN dengan Parameter K dan PSO Untuk Klasifikasi Status Gizi Balita. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(3), 1609. <https://doi.org/10.30865/mib.v8i3.7841>

- Salwa Alexita, A. C., Kusumaningtyas, P., & Rofi'i, M. (2025). OPTIMASI ALGORITMA RANDOM FOREST MENGGUNAKAN PSO UNTUK KLASIFIKASI KANKER PAYUDARA DENGAN CITRA MAMMOGRAMS. *Teknika STTKD: Jurnal Teknik, Elektronik, Engine*, 11(1), 47–54. <https://doi.org/10.56521/teknika.v11i1.1346>
- Santika, A. A., Saragih, T. H., & Muliadi, M. (2023). Penerapan Skala Likert pada Klasifikasi Tingkat Kepuasan Pelanggan Agen Brilink Menggunakan Random Forest. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, 11(3), 405. <https://doi.org/10.26418/justin.v11i3.62086>
- Santoso, F., & Hartati, E. (2022). PENGGUNAAN ALGORITMA RANDOM FOREST DALAM KLASIFIKASI BUAH SEGAR DAN BUSUK. *Jurnal Algoritme*, 3(1), 133–140. <https://doi.org/10.35957/algoritme.v3i1.3404>
- Saud Abd, N., Salim Atiyah, O., Taher Ahmed, M., & Bakhit, A. (2024). Digital marketing data classification by using machine learning algorithms. *Iraqi Journal for Electrical and Electronic Engineering*, 20(1), 245–256. <https://doi.org/10.37917/ijeee.20.1.23>
- Syukron, A., & Subekti, A. (2018). Penerapan Metode Random Over-Under Sampling dan Random Forest Untuk Klasifikasi Penilaian Kredit. *Jurnal Informatika*, 5(2), 175–185. <https://doi.org/10.31294/ji.v5i2.4158>
- Wang, H., Lu, Q., & Chen, C. (2024). Prediction of tennis match based on random forest and PSO-SVM. *Proceedings of the 2024 International Conference on Smart City and Information System*, 556–560. <https://doi.org/10.1145/3685088.3685185>