

Optimasi Model Algoritma *Machine Learning Supervised* menggunakan Algoritma Genetika untuk Prediksi Kebakaran Hutan dan Lahan

Putri Yuli Utami^{1*}, Sucipto², Andrea Risli³, Tiara Aurilia Viona⁴

putriyuli@unmuhpnk.ac.id¹, sucipto@unmuhpnk.ac.id², 201230004@unmuhpnk.ac.id³,

tiaraviona007@gmail.com⁴

^{1,2,3,4} Universitas Muhammadiyah Pontianak

Informasi Artikel	Abstrak
Diterima : 6 Sep 2024 Direview : 7 Des 2024 Disetujui : 6 Apr 2025	Fenomena kebakaran hutan dan lahan sering terjadi di wilayah Indonesia terutama di Sumatera dan Kalimantan. Mitigasi kebakaran hutan dan lahan salah satunya dengan prediksi area rawan kebakaran hutan lahan. Pada penelitian ini optimasi Algoritma Genetika (GA) dilakukan untuk meningkatkan kinerja algoritma <i>Random Tree</i> dan <i>Hyper-SVM</i> dengan membandingkan hasil pengujian yang optimal. Dataset yang digunakan adalah data <i>hotspot</i> dan data iklim Kabupaten Ketapang tahun 2021-2022. Hasil dari pelatihan dan pengujian ini menunjukkan bahwa optimasi algoritma <i>Random Tree+GA</i> dengan nilai <i>PC</i> 0.6 dan parameter seleksi Boltzmann menghasilkan akurasi sebesar 99.77%, <i>Recall</i> sebesar 94.88%, <i>Precision</i> sebesar 95%, RMSE sebesar 0.015 dan Kappa sebesar 0.9 sedangkan untuk optimasi <i>Hyper-SVM+GA</i> dengan nilai <i>PC</i> 0.6 dan parameter seleksi Boltzmann menghasilkan akurasi sebesar 70.48%, <i>Recall</i> sebesar 52.14%, <i>Precision</i> sebesar 50.58, RMSE sebesar 0.493 dan Kappa sebesar 0.026. Hal ini menunjukkan bahwa optimasi model algoritma <i>Random Tree+GA</i> memiliki performa lebih baik dibandingkan optimasi <i>Hyper-SVM+GA</i> . Sehingga dapat disimpulkan bahwa Algoritma <i>Random Tree+GA</i> merupakan model klasifikasi terbaik dalam memprediksi kemungkinan terjadinya kebakaran hutan dan lahan di Kabupaten Ketapang.
Kata Kunci	
Algoritma Genetik, Hyper-SVM, Machine Learning, Random Tree, Supervised	

Keywords	Abstract
Genetic Algorithm, Hyper-SVM, Machine Learning, Random Tree, Supervised	Forest and land fires are a common occurrence in Indonesia, particularly in the provinces of Sumatra and Kalimantan. One strategy for mitigating the impact of forest and land fires is to predict areas that are prone to such incidents. In this study, genetic algorithm (GA) optimization was employed to enhance the efficacy of the random tree and hyper-SVM algorithms, with a view to identifying the most optimal test results. The dataset utilized in this study comprises hotspot data and climate data for Ketapang Regency during the 2021-2022 period. The results of the training and testing demonstrate that the Random Tree +GA algorithm optimization with a PC value of 0.6 and Boltzmann selection parameters yields an accuracy of 99.77%, a recall of 94.88%, a precision of 95%, an RMSE of 0.015, and a Kappa of 0.9. In contrast, the Hyper-SVM +GA optimization, with a PC value of 0.6 and Boltzmann selection parameters, yielded an accuracy of 70.48%, a recall of 52.14%, a precision of 50.58%, an RMSE of 0.493, and a Kappa of 0.026. The results demonstrate that the Random Tree +GA algorithm model optimization exhibits superior performance compared to Hyper-SVM +GA optimization. Consequently, it can be inferred that the Random Tree +GA algorithm represents the most effective classification model for predicting the likelihood of forest and land fires in Ketapang Regency.

A. Pendahuluan

Fenomena alam yang terus terjadi selama beberapa tahun terakhir ialah kebakaran hutan dan lahan. Bencana ini merupakan permasalahan yang serius karena berdampak pada keberlangsungan ekosistem makhluk hidup dan kesejahteraan masyarakat. Kebakaran hutan dan lahan cukup sering terjadi di wilayah Indonesia terutama di Sumatera dan Kalimantan[1]. Dampak kebakaran hutan dan lahan menimbulkan rusaknya ekosistem keanekaragaman hayati darat [2][3]. Beberapa faktor penyebab kebakaran hutan dan lahan diantaranya perubahan iklim, pembukaan lahan yang tidak berkelanjutan serta kondisi cuaca ekstrem [4].

Berdasarkan data kementerian Lingkungan Hidup dan Kehutanan luas kebakaran hutan dan lahan di Kalimantan Barat meningkat antara tahun 2014 hingga 2019, yakni mencapai 3.556,10 hektar menjadi 151.070,00 hektar pada tahun 2019 [5]. Luas kebakaran hutan dan lahan di Kalimantan Barat tahun 2023 mencapai 108.791 hektar [6]. Kabupaten Ketapang memiliki areal seluas 184 ribu hektar, yakni wilayah hutan gambut sebesar 28% dan hutan alam di wilayah perhutanan sosial sebesar 73% [7]. Luas lahan terbakar di Kabupaten Ketapang pada tahun 2023 sebesar 1.453,90 hektar [8][9]. Hal ini menyebabkan Kabupaten Ketapang memiliki risiko tinggi terjadi kebakaran hutan dan lahan.

Mitigasi kebakaran hutan dan lahan salah satunya dengan prediksi area rawan kebakaran hutan lahan merupakan salah satu cara mengatasi kebakaran hutan dan lahan yang terus berulang. Algoritma *machine learning* memiliki kemampuan dalam penanganan data yang kompleks dan *non-linier* termasuk dalam pengolahan data cuaca, kelembaban dan vegetasi [10]. Algoritma *Support vector machine* (SVM) dan *Random Tree* merupakan algoritma *supervised machine learning*. Algoritma SVM yang cukup sering digunakan adalah *hyper-SVM*, metode optimasi yang menggabungkan evolusi untuk mengolah data yang kompleks. Namun algoritma ini belum cukup efektif dalam pemilihan parameter yang optimal untuk melakukan prediksi [11]. Metode ini juga sangat sensitif dalam pemilihan parameter awal sehingga hasil klasifikasi tidak stabil [12]. Pendekatan *decision tree* juga cukup efektif dalam kasus klasifikasi kebakaran hutan dan lahan. Salah satunya adalah algoritma *Random Tree*, yang memiliki keunggulan dalam mengolah dataset jumlah besar, namun memiliki kelemahan terhadap sensitifitas data *noise* dan *outlier* serta kecenderungan untuk *overfitting* terhadap data yang kompleks menjadi kendala [13][14]. Hasil penelitian klasifikasi kebakaran hutan dan lahan menunjukkan bahwa keakuratan hasil prediksi memerlukan model prediksi yang lebih akurat [15].

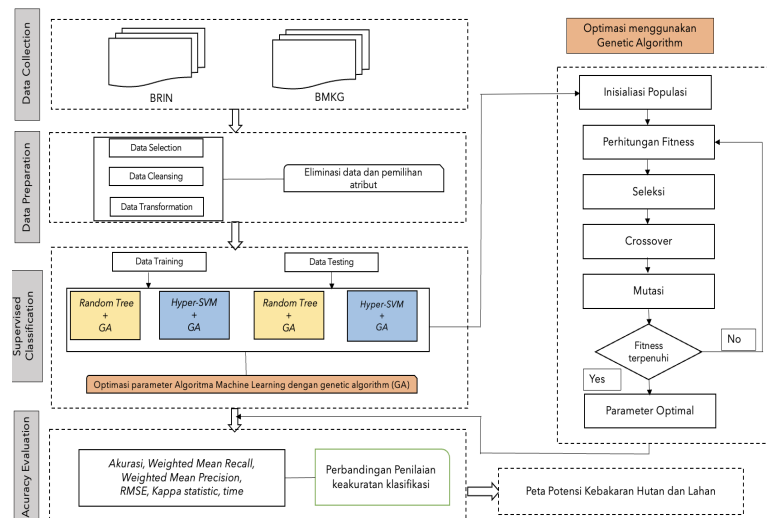
Algoritma evolusi untuk meningkatkan kinerja *hyper-SVM* dan *Random tree* dalam pencarian solusi optimasi yang efektif adalah Algoritma Genetika (GA). Kombinasi algoritma genetika dapat menghasilkan kombinasi parameter dengan kinerja yang optimal dan meningkatkan akurasi klasifikasi prediksi dataset yang kompleks [16]. Algoritma genetika mampu memberikan performa yang baik dalam klasifikasi hasil algoritma *Random Tree* [17]-[19]. Algoritma Genetika juga mampu meningkatkan performansi kinerja [20]. Untuk itu penelitian ini bertujuan meningkatkan performa kinerja algoritma *hyper-SVM* dan *Random Tree* dalam klasifikasi prediksi kebakaran hutan dan lahan melalui optimasi menggunakan algoritma genetika. Pengujian dilakukan melalui pengukuran akurasi, *kappa*, *RMSE*,

recall, *precision*, dan *time*. Hasil klasifikasi daerah rawan kebakaran hutan disajikan dalam visualisasi menggunakan *Google Data Studio*.

B. Metode Penelitian

1. Tahapan Penelitian

Penelitian ini mengimplementasikan algoritma *supervised machine learning* untuk klasifikasi kebakaran hutan dan lahan. Algoritma yang digunakan untuk memprediksi kebakaran hutan dan lahan adalah algoritma *hyper-SVM* dan algoritma *Random Tree*. Penelitian ini mengembangkan algoritma genetika untuk mengoptimasi kinerja algoritma *hyper-SVM* dan *Random Tree*. Tahapan penelitian terlihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

a. Data Collection

Tahapan *data collection* dilakukan pengumpulan data wilayah Kabupaten Ketapang, Kalimantan Barat melalui data citra satelit terkait data hotspot bersumber dari website BRIN <https://hotspot.brin.go.id/> dari tahun 2021 hingga 2022 serta data iklim dari website BMKG <https://dataonline.bmkg.go.id/home> terkait data curah hujan, kelembaban rata-rata dan suhu rata-rata. Sampel dataset yang digunakan berjumlah 4697 data terlihat pada Tabel 1.

Tabel 1. Database Penelitian

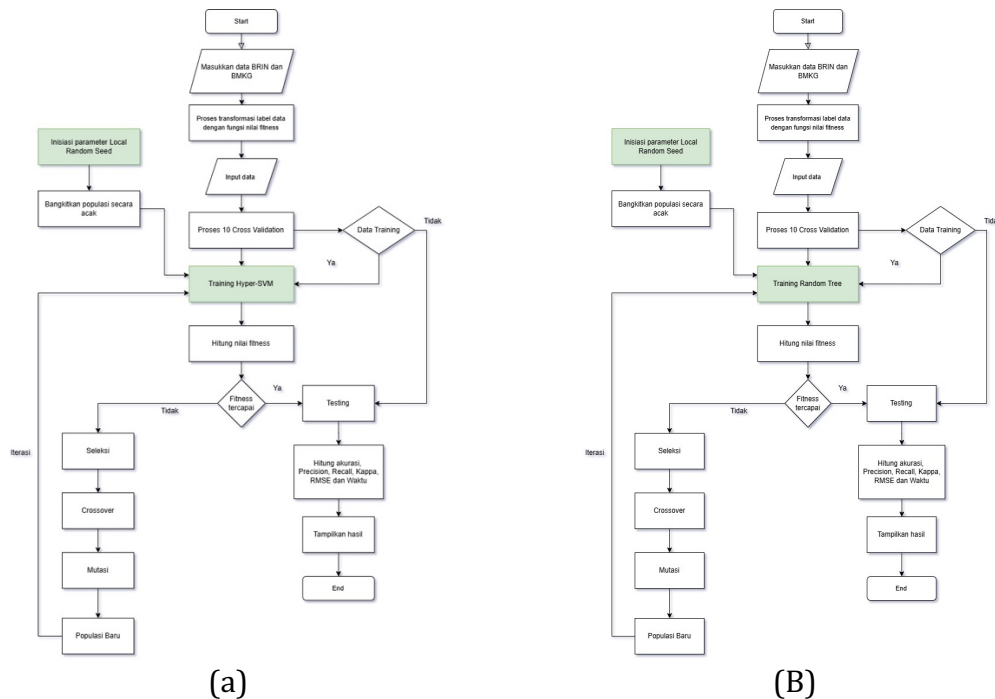
tingkat kepercayaan	Curah Hujan (mm)	temperatur rata-rata (C)	Kelembaban rata-rata (%)	Radius Kemungkinan	tanggal (WIB)	lintang	bujur	kecamatan	Status
0,0	0,0	-0,4	0,2	0,2	2021-01-17 00:00:00	2818864743333300,0	1103172225966600,0	Kendawangan	Rawan
0,0	0,0	-0,4	0,2	-0,3	2021-01-17 00:00:00	-264375353,0	1111241684,0	Manisamata	Rawan
0,0	0,0	-0,4	0,2	-0,3	2021-01-17 00:00:00	-272007442,0	11032154846,0	Kendawangan	Rawan
0,0	0,0	-0,4	-0,6	3,2	2021-01-20 00:00:00	-0,9	110493896484375,0	Sungai Laur	Rawan
0,0	0,0	-0,4	-0,6	0,0	2021-01-20 00:00:00	-0,8	11049296951,0	Sungai Laur	Rawan
0,0	0,0	-0,4	-0,6	-0,3	2021-01-21 00:00:00	-27247467,0	11055180359,0	Kendawangan	Rawan
0,0	0,0	-0,4	-0,6	-0,3	2021-01-21 00:00:00	-274842978,0	11037294769,0	Kendawangan	Rawan
0,0	0,0	-0,4	-0,6	-0,3	2021-01-21 00:00:00	-117965579,0	11050855255,0	Sandai	Rawan
0,0	0,0	-0,4	-0,6	-0,3	2021-01-21 00:00:00	-139322305,0	11027574158,0	Matan Hilir Utara	Rawan
0,0	0,0	-0,4	-0,6	-2,2	2021-01-21 00:00:00	2522098666666600,0	1106406810000000,0	Kendawangan	Rawan
0,0	0,0	-0,4	1,0	3,2	2021-01-25 00:00:00	-0,7	10994744873046800,0	Simpanghulu	Rawan
0,0	0,0	-0,4	1,0	3,2	2021-01-25 00:00:00	2224680185317990,0	11013546752929600,0	Matan Hilir Selatan	Rawan

b. Data Preparation

Tahapan *data preparation* dibagi dalam tiga tahapan yaitu *data selection*, *data cleansing*, dan *data transformation*. *Data selection* yaitu melakukan seleksi parameter data hotspot yaitu parameter id, tanggal (WIB), waktu (WIB), lintang, bujur, tingkat kepercayaan, radius kemungkinan, kecamatan dan kabupaten. Seleksi parameter yang digunakan dalam dataset iklim yaitu curah hujan, kelembaban, dan suhu. Tahap *data cleansing* dilakukan proses pembersihan data untuk menghindari *missing value*. Selanjutnya *data transformation* melakukan normalisasi data untuk diolah kedalam proses algoritma.

c. Supervised Classification

Pengembangan optimasi model algoritma *Hyper-SVM* dan *Random Tree* menggunakan algoritma genetika dilakukan menggunakan *AI Studio 2024*, optimasi parameter dilakukan inisialisasi awal menggunakan *parameter local random seed* untuk membangkitkan populasi awal secara acak. Dataset dibagi menjadi data latih dan data uji kemudian dilanjutkan proses optimasi algoritma genetika, menghitung nilai *fitness*, proses seleksi, proses *crossover*, proses mutasi kemudian menghasilkan hasil akhir yaitu solusi terbaik. Performa optimasi algoritma *Hyper-SVM* dan optimasi algoritma *Random Tree* akan dibandingkan untuk melihat kinerja terbaik. Tahapan penelitian ditunjukkan pada Gambar 2 dan Gambar 3.



(a) Tahapan optimasi Algoritma *Hyper-SVM+GA*
 (b) Tahapan optimasi Algoritma *Random Tree +GA*

d. Accuracy Evaluation

Pada tahap ini dilakukan pengujian performa model optimasi algoritma *Hyper-SVM* dan optimasi algoritma *Random Tree* terhadap nilai akurasi, *kappa*, *RMSE*, *recall*, *precision*, dan *time* [21] [22].

- Adapun perhitungan untuk nilai *RMSE* menggunakan persamaan (1) berikut.

$$RMSE = \frac{\sum_{t=1}^n (Y'_t - Y_t)^2}{n} \quad (1)$$

Keterangan:

Y'_t : nilai data aktual ke- t ($t = 1, \dots, n$),

Y_t : nilai aktual ke- t ($t = 1, \dots, n$),

n : banyaknya data yang diuji.

- Adapun perhitungan untuk nilai *Kappa* menggunakan persamaan (2) berikut.

$$Kappa = \frac{N \sum_{i=1}^n m_{i,i} - \sum_{i=1}^n (G_i C_i)}{N^2 - \sum_{i=1}^n (G_i C_i)} \quad (2)$$

Keterangan:

i : nomor kelas

N : total dari nilai yang diklasifikasi dibandingkan dengan nilai kebenaran

$m_{i,i}$: jumlah nilai milik kelas kebenaran i yang juga telah diklasifikasikan sebagai kelas i (yaitu, nilai yang ditemukan di sepanjang *confussion matrix*s)

C_i : jumlah total nilai prediksi yang termasuk dalam kelas i

G_i : jumlah total nilai kebenaran yang termasuk dalam kelas i

- Adapun perhitungan untuk nilai akurasi menggunakan persamaan (3) berikut.

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (3)$$

- Adapun perhitungan untuk nilai *Precision* menggunakan persamaan (4) berikut.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

- Adapun perhitungan untuk nilai *Recall* menggunakan persamaan (5) berikut.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

Keterangan:

TP: *True Positive*, merupakan total data positif yang diprediksi benar oleh model.

TN: *True Negative*, merupakan total data negatif yang diprediksi benar oleh model.

FN: *False Negative*, merupakan total data negatif namun diprediksi salah oleh model.

FP: *False Positive*, merupakan total data positif namun diprediksi salah oleh model

e. Peta Kerentanan Kebakaran Hutan dan Lahan

Pada tahap ini dilakukan identifikasi pemetaan area rawan kebakaran hutan dan lahan. Pemetaan dilakukan menggunakan *Google Data Studio*.

C. Hasil dan Pembahasan

1. Optimasi Algoritma *Random Tree* menggunakan Algoritma Genetika

a. Evaluasi Hasil Pengujian *Crossover*

Pengujian *Random Tree* dilakukan terhadap beberapa penetapan parameter algoritma genetika. Pengujian pertama dilakukan seleksi Boltzman dan mutasi Gaussian dengan beberapa nilai *probabilitas crossover (Pc)* dalam rentang 0,2 – 1. Hasil Pengujian Algoritma *Random Tree* diperlihatkan pada Tabel 3.

Tabel2. Parameter Algoritma Genetika

Tabel 3. Hasil Pengujian Pc pada Algoritma *Random Tree*

		<i>Pc</i>	Akurasi	<i>Kappa</i>	<i>RMSE</i>	<i>Time</i> (s)	<i>Recall</i>	<i>Precision</i>
Parameter	Nilai dan Metode							
		0.2	99.55	0.8	0.03	17	89.78	90
Pop size	15	0.3	99.76	0.9	0.014	11	94.88	95
MaxGen	50	0.4	99.77	0.9	0.015	9	94.88	95
Seleksi	Boltzman	0.5	99.53	0.8	0.03	7	89.77	90
Mutasi	Gaussian	0.6	99.79	0.9	0.014	11	94.89	95
		0.7	99.53	0.8	0.03	5	89.77	90
		0.8	99.77	0.9	0.015	9	94.88	95
		0.9	99.55	0.8	0.03	8	89.78	90
		1	99.77	0.9	0.015	9	94.88	95

Hasil pengujian Tabel 3 memperlihatkan pengujian nilai *probabilitas crossover (Pc)* parameter seleksi *Boltzman* berdasarkan perbandingan nilai *probabilitas crossover* terhadap hasil akurasi, *kappa*, *RMSE*, *time*, *recall*, *precision*.

Nilai akurasi sangat tinggi berada pada kisaran 99, 53% - 99,79%. Nilai kappa berada pada kisaran nilai 0,8 – 0,9 hal ini menunjukkan tingkat kesesuaian antara model klasifikasi dengan data referensi mendekati sempurna. Nilai *Root Mean Square Error* (RMSE) sangat rendah berada diantara kisaran 0,014 – 0,03, hal ini menunjukkan prediksi model sangat dekat dengan nilai sebenarnya. Waktu komputasi yang diperlukan berkisar diantara 5 – 17 detik, waktu yang cepat tidak secara signifikan mempengaruhi akurasi dan kinerja model prediksi. Nilai *recall* dan *precision* berada diantara 89,78% - 89,78% dan 90% - 95%. Hal ini memperlihatkan kemampuan model untuk mengidentifikasi positif yang benar dan presisi prediksi model lebih baik.

Berdasarkan beberapa kali hasil percobaan nilai *Pc* optimal 0,6 dengan nilai akurasi tertinggi 99,79%, nilai kappa berada di angka 0,9 tingkat kesesuaian *almost perfect agreement*, RMSE paling minimal dengan nilai 0,014, serta nilai *recall* dan *precision* tertinggi yaitu 94,89% dan 95%. Hal ini menunjukkan nilai *Pc* 0,6 memiliki kekuatan konsistensi diberbagai metrik kinerja. Perbedaan waktu dan kinerja dalam ini tidak memiliki pengaruh yang signifikan terhadap akurasi atau metrik lainnya karena waktu 11 detik cukup optimal untuk metrik akurasi 99,79%.

Pengujian berbeda pada Tabel 5 dengan parameter seleksi *Roulette wheel* berdasarkan perbandingan nilai *probabilitas crossover* terhadap hasil akurasi, *kappa*, *RMSE*, *time*, *recall*, *precision*. Beberapa kali percobaan memperlihatkan bahwa nilai *Pc* optimal berada pada nilai 0,6 dengan akurasi 99,77%, nilai kappa berada diangka 09, RMSE dengan nilai 0,015, times 3 detik serta nilai *recall* dan *precision* berada pada nilai 94,88% dan 95%. Berdasarkan percobaan kedua parameter seleksi kinerja optimasi algoritma *random tree* lebih optimal menggunakan parameter seleksi Boltzman.

Tabel 4. Parameter Algoritma Genetika

Parameter	Nilai dan Metode
Pop size	15
MaxGen	50
Seleksi	Roulette Wheel
Mutasi	Gaussian

Tabel 5. Hasil Pengujian *Pc* pada Algoritma *Random Tree*

<i>Pc</i>	Akurasi	Kappa	RMSE	Time (s)	Recall	Precision
0.2	99.55	0.8	0.03	15	89.78	90
0.3	99.55	0.8	0.03	8	89.78	90
0.4	99.55	0.8	0.03	4	89.78	90
0.5	99.74	0.9	0.016	4	94.87	95
0.6	99.77	0.9	0.015	3	94.88	95
0.7	99.77	0.9	0.015	9	94.88	95
0.8	99.77	0.9	0.015	8	94.88	95
0.9	99.53	0.8	0.03	10	89.77	90
1	99.77	0.9	0.015	8	94.88	95

b. Pengujian Algoritma *Random Tree* dengan Algoritma Genetika

Pengujian optimasi juga dilakukan perbandingan kinerja optimasi algoritma *Random Tree* dengan GA dan algoritma *Random Tree* tanpa GA.



Gambar 4. Grafik perbandingan metrik optimasi *Random Tree* dengan optimasi GA dan tanpa GA : (a)akurasi, *recall*, *precision* (b) Time dan Kappa (c) RMSE

Berdasarkan grafik (a) dapat diketahui bahwa sebelum dioptimasi akurasi *Algoritma Random Tree* sebesar 98.22%, *recall* sebesar 75% dan *precision* sebesar 74.41% sedangkan setelah dilakukan optimasi menggunakan algoritma genetika akurasi algoritma *Random Tree* menjadi 99.77%, nilai *recall* menjadi 94.88% dan nilai *Precision* sebesar 95% sehingga dapat disimpulkan bahwa terjadi peningkatan nilai akurasi, *recall* dan *precision*. Selanjutnya grafik (b) memperlihatkan lama proses *running* terjadi penambahan 2 detik dari waktu awal proses hanya 1 detik menjadi 3 detik. Nilai Kappa juga terjadi peningkatan, yakni sebelum dioptimasi nilai kappa sebesar 0.5 setelah dioptimasi nilai kappa menjadi 0.9. Kemudian grafik (c) memperlihatkan nilai RMSE terjadi penurunan, sebelum dioptimasi nilai RMSE sebesar 0.092 namun setelah dilakukan optimasi nilai RMSE terjadi penurunan yang signifikan menjadi 0.015. Sehingga berdasarkan hasil perbandingan akurasi, kappa, RMSE, *recall*, *precision* dan waktu proses dapat disimpulkan bahwa kinerja algoritma genetika memberikan hasil optimal dalam melakukan klasifikasi.

2. Optimasi Algoritma *Hyper-SVM* menggunakan Algoritma Genetika

a. Evaluasi Hasil Pengujian *Crossover*

Pengujian *Hyper-SVM* dilakukan terhadap beberapa penetapan parameter algoritma genetika. Pengujian pertama dilakukan seleksi Boltzman dan mutasi

Gaussian dengan beberapa nilai *probabilitas crossover* (P_c) dalam rentang 0,2 – 1. Hasil Pengujian Algoritma *Random Tree* diperlihatkan pada Tabel 7.

Tabel 6. Parameter Algoritma Genetika

Parameter	Nilai dan Metode
Pop size	15
MaxGen	50
Seleksi	Bolzman
Mutasi	Gaussian

Tabel 7. Hasil Pengujian P_c pada Algoritma *Hyper-SVM*

P_c	Akurasi	Kappa	RMSE	Time	Recall	Precision
0.2	68.46	-0	0.539	2	49.18	49.66
0.3	68.46	-0	0.539	2	49.18	49.66
0.4	68.46	-0	0.539	2	49.18	49.66
0.5	72.96	-0.001	0.507	5	50.18	49.98
0.6	70.48	0.026	0.493	3	52.14	50.58
0.7	70.48	0.026	0.493	3	52.14	50.58
0.8	70.48	0.026	0.493	3	52.14	50.58
0.9	70.48	0.026	0.493	4	52.14	50.58
1	68.46	-0	0.539	10	49.18	49.66

Hasil pengujian Tabel 7 memperlihatkan pengujian nilai *probabilitas crossover* (P_c) parameter seleksi *Bolzman* berdasarkan perbandingan nilai *probabilitas crossover* terhadap hasil akurasi, *kappa*, *RMSE*, *time*, *recall*, *precision*. Nilai akurasi sangat tinggi berada pada kisaran 99, 53% - 99,79%. Nilai *kappa* berada pada kisaran nilai 0,8 – 0,9 hal ini menunjukkan tingkat kesesuaian antara model klasifikasi dengan data referensi mendekati sempurna. Nilai *Root Mean Square Error* (RMSE) sangat rendah berada diantara kisaran 0,014 – 0,03, hal ini menunjukkan prediksi model sangat dekat dengan nilai sebenarnya. Waktu komputasi yang diperlukan berkisar diantara 5 – 17 detik, waktu yang cepat tidak secara signifikan mempengaruhi akurasi dan kinerja model prediksi. Nilai *recall* dan *precision* berada diantara 89,78% - 89,78% dan 90% - 95%. Hal ini memperlihatkan kemampuan model untuk mengidentifikasi positif yang benar dan presisi prediksi model lebih baik.

Berdasarkan beberapa kali hasil percobaan nilai P_c optimal berada pada nilai 0,6 dengan nilai akurasi tertinggi 99,79%, nilai *kappa* berada di angka 0,9 tingkat kesesuaian *almost perfect agreement*, RMSE paling minimal dengan nilai 0,014, serta nilai *recall* dan *precision* tertinggi yaitu 94,89% dan 95%. Hal ini menunjukkan nilai P_c 0,6 memiliki kekuatan konsistensi diberbagai metrik kinerja. Perbedaan waktu dan kinerja dalam hal ini tidak memiliki pengaruh yang signifikan terhadap akurasi atau metrik lainnya karena waktu 11 detik cukup optimal untuk metrik akurasi 99,79%.

Pengujian berbeda dilakukan pada Tabel 9 dengan parameter seleksi *Roulette wheel* berdasarkan perbandingan nilai *probabilitas crossover* terhadap hasil akurasi, *kappa*, *RMSE*, *time*, *recall*, *precision*. Beberapa kali percobaan memperlihatkan bahwa nilai P_c optimal berada pada nilai 0,6 dengan akurasi 99,77%, nilai *kappa* berada diangka 09, RMSE dengan nilai 0,015, times 3 detik serta nilai *recall* dan *precision* berada pada nilai 94,88% dan 95%. Berdasarkan percobaan kedua parameter seleksi kinerja optimasi algoritma *random tree* lebih optimal menggunakan parameter seleksi bolzman.

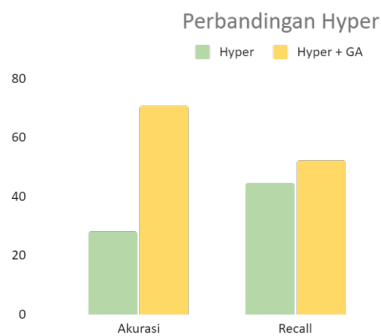
Tabel 8. Parameter Algoritma Genetika

Parameter	Nilai dan Metode
Pop size	15
MaxGen	50
Seleksi	Roulette Wheel
Mutasi	Gaussian

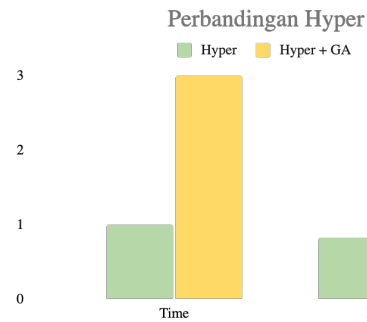
Tabel 9. Hasil Pengujian Pc pada Algoritma *Hyper-SVM*

Pc	Akurasi	Kappa	RMSE	Time	Recall	Precision
0.2	68.46	-0	0.539	3	49.18	49.66
0.3	68.46	-0	0.539	2	49.18	49.66
0.4	68.46	-0	0.539	2	49.18	49.66
0.5	72.96	-0.001	0.507	5	50.18	49.98
0.6	70.48	0.026	0.493	3	52.14	50.58
0.7	70.48	0.026	0.493	3	52.14	50.58
0.8	70.48	0.026	0.493	3	52.14	50.58
0.9	70.48	0.026	0.493	4	52.14	50.58
1	68.46	-0	0.539	4	49.18	49.66

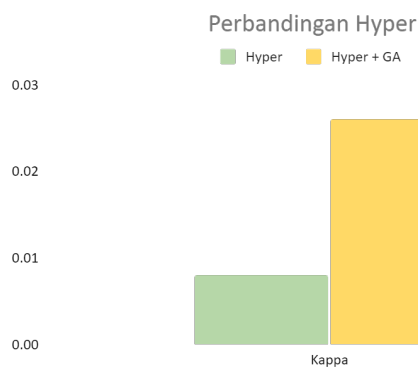
b. Pengujian *Algoritma Hyper-SVM* dengan Algoritma Genetika
 Pengujian optimasi juga dilakukan perbandingan kinerja algoritma *hyper-SVM* optimasi dengan Algoritma Genetika dan algoritma *hyper-svm* tanpa Algoritma Genetika.



(a)



(b)



(c)

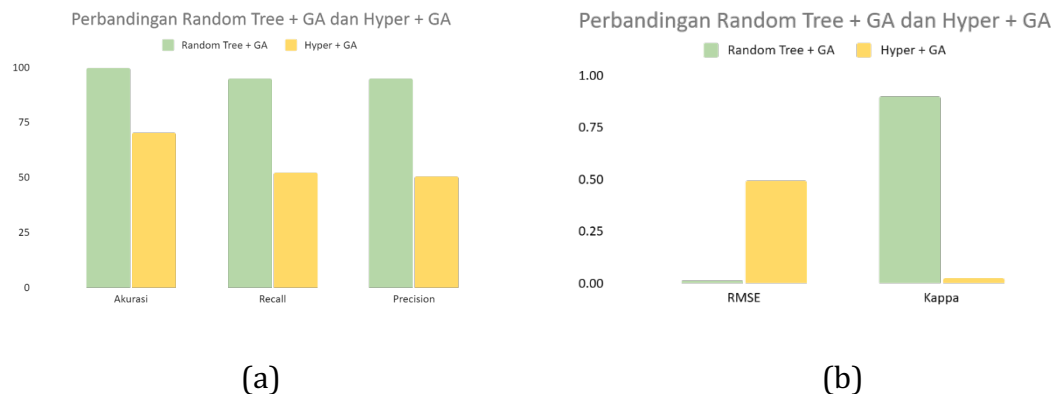
Gambar 5. Grafik perbandingan metrik optimasi *Hyper-SVM* dengan Optimasi GA dan tanpa GA : (a)akurasi, *recall*, *precision* (b) Time dan RMSE (c) Kappa

Berdasarkan Gambar 5 grafik (a) dapat diketahui bahwa sebelum dioptimasi akurasi algoritma *hyper-SVM* hanya sebesar 27.89%, *recall* sebesar 44.51% dan *precision* sebesar 48.38% sedangkan setelah dilakukannya optimasi GA akurasi algoritma *hyper-SVM* mengalami peningkatan menjadi 70.48%, *recall* sebesar

52.14% dan *precision* sebesar 50.58%. Grafik (b) memperlihatkan bahwa untuk proses *running* terjadi penambahan 2 detik dari waktu awal proses hanya 1 detik menjadi 3 detik. Sedangkan untuk nilai RMSE terjadi penurunan, sebelum dioptimasi nilai RMSE sebesar 0.816 namun setelah dilakukan optimasi nilai RMSE turun signifikan menjadi 0.493. Berikutnya grafik (c) memperlihatkan bahwa nilai kappa sebelum dioptimasi GA sebesar 0.008 setelah dioptimasi nilai kappa naik menjadi 0.026.

3. Perbandingan Algoritma *Random Tree* dan Hyper-SVM

Pengujian hasil performa optimasi *Random Tree+GA* dan Optimasi *Hyper-SVM+GA* diperlihatkan pada Gambar 6.

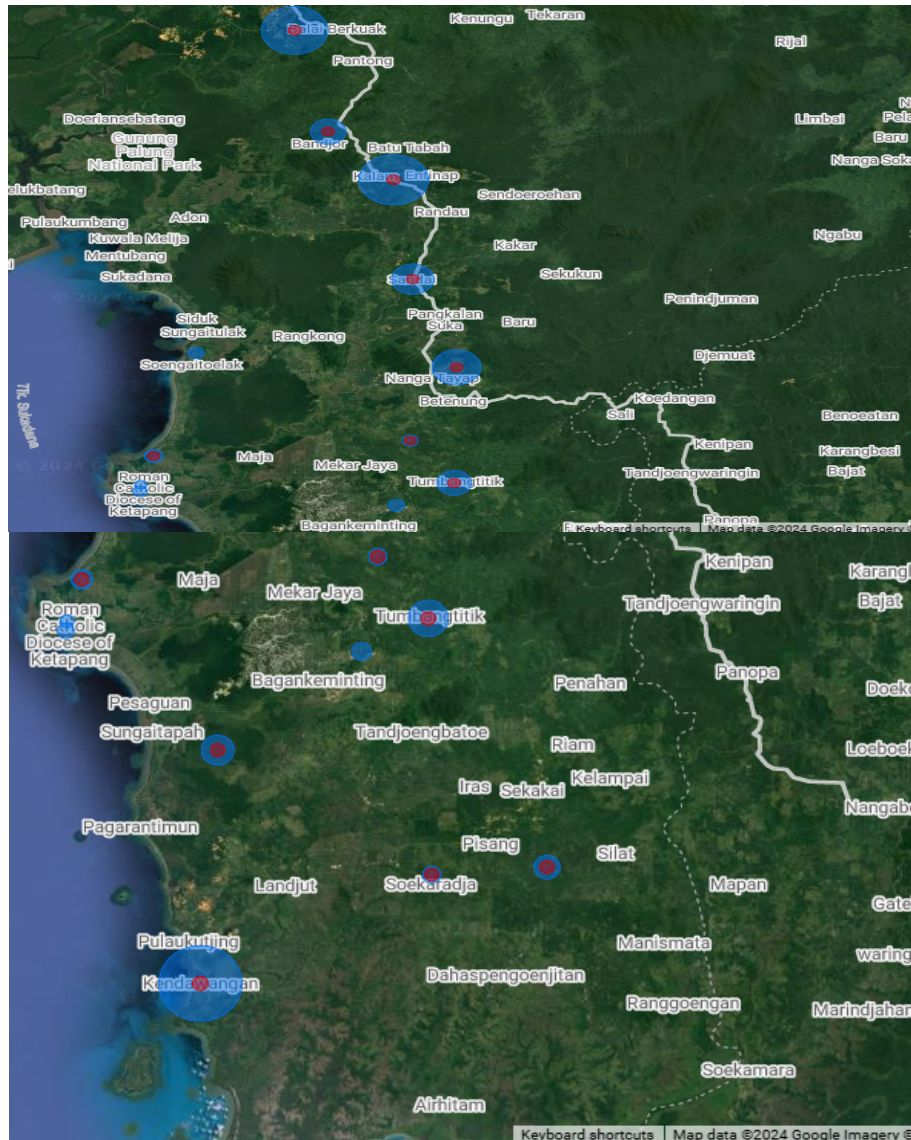


Gambar 6. Perbandingan Optimasi *RandomTree+GA* dan *Hyper-SVM+GA*: (a) akurasi, *recall*, *precision* (b) RMSE, kappa

Berdasarkan Gambar 6 memperlihatkan grafik pengujian perbandingan performa optimasi *Random Tree+GA* dan *Hyper-SVM+GA*, grafik (a) memperlihatkan optimasi algoritma *Random Tree+GA* memiliki akurasi sebesar 99.77%, *recall* sebesar 94.88%, *precision* sebesar 95%, dan RMSE sebesar 0.015 serta kappa sebesar 0.9. Sedangkan untuk optimasi algoritma *Hyper-SVM+GA* memperoleh hasil akurasi sebesar 70.48%, *recall* sebesar 52.14%, *precision* sebesar 50.58%, RMSE sebesar 0.493 dan kappa sebesar 0.026. Berdasarkan hasil tersebut teridentifikasi bahwa optimasi algoritma *Random Tree+GA* dapat melakukan klasifikasi lebih optimal dibandingkan algoritma *Hyper-SVM+GA*.

4. Visualisasi Hasil Prediksi Kebakaran Hutan dan Lahan

Klasifikasi hasil daerah rawan kebakaran hutan dan lahan menggunakan algoritma *Random Tree+GA* menghasilkan prediksi daerah rawan kebakaran hutan di Kabupaten Ketapang. Visualisasi hasil prediksi disajikan menggunakan *Google Data Studio* terlihat pada Gambar 7.



Gambar 7. Peta Daerah Rawan Kebakaran Hutan dan Lahan

Gambar 7 memperlihatkan klasifikasi daerah rawan kebakaran hutan dan lahan di Kabupaten Ketapang, daerah yang memiliki lingkaran berwarna biru berukuran besar mengindikasi daerah tersebut terdapat data *hotspot* yang banyak namun nilai *confidence hotspot* diantara range $30\% \leq C \leq 80\%$. Berikutnya untuk lingkaran berwarna biru berukuran kecil mengindikasikan daerah rawan yang memiliki data *hotspot* sedikit dan memiliki nilai *confidence hotspot* diantara range $30\% \leq C \leq 80\%$ sehingga daerah tersebut termasuk dalam kategori waspada dan perlu dilakukan pengawasan. Daerah yang termasuk kelas rawan berdasarkan visualisasi peta adalah Simpang Hulu, Kendawang, Marau, Nangatayap, Simpang Laur, Hulu Sungai, Simpang Dua, Matan Hilir Selatan, Benua Kayong, Singkup, Matan Hilir Utara, Manismata, Air Upas, Pemahan, Tumbang titi, Delta Pawan, Riam, Sandai, Muara Pawan dan Sungai Melayu Rayak. Lingkaran berwarna merah mengindikasikan bahwa daerah tersebut sangat rawan terjadi kebakaran hutan dan lahan dengan nilai *confidence hotspot* $80\% \leq C \leq 100\%$. Sehingga daerah tersebut termasuk dalam kategori tinggi dan perlu dilakukan penanggulangan. Berdasarkan visualisasi peta

hasil klasifikasi daerah yang termasuk kelas sangat rawan yaitu kecamatan AirUpas terdapat 1 titik, Hulu Sungai terdapat 1 titik, Kendawang terdapat 15 titik, Manismata terdapat 5 titik, Marau terdapat 11 titik, Matan Hilir Selatan terdapat 3 titik, Muara Pawan terdapat 3 titik, Nanga Tayap terdapat 7 titik, Pemahan terdapat 2 titik, Riam terdapat 5 titik, Sandai terdapat 6 titik, Simpang Dua terdapat 4 titik, Simpang Hulu terdapat 14 titik, Singkup terdapat 1 titik, Sungai Laur terdapat 14 titik dan Tumbang Titi terdapat 6 titik.

D. Simpulan

Berdasarkan hasil pelatihan dan pengujian algoritma *Random Tree* dan *Hyper-SVM* menggunakan optimasi algoritma genetika untuk prediksi daerah rawan kebakaran hutan dan lahan di Kabupaten Ketapang dapat disimpulkan bahwa kinerja model algoritma *Random Tree+GA* dapat melakukan klasifikasi lebih optimal yakni dengan nilai pengujian parameter *crossover PC* 0.6 dan parameter seleksi Boltzmann menghasilkan akurasi sebesar 99.77%, *recall* sebesar 94.88%, *precision* sebesar 95%, RMSE sebesar 0.015 dan Kappa sebesar 0.9.

E. Ucapan Terima Kasih

Ucapan terima kasih disampaikan kepada Direktorat Jendral Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan Kebudayaan, Riset dan Teknologi atas dukungan pendanaan melalui hibah skema Penelitian Dosen Pemula pada Tahun 2024 dengan nomor kontrak 108/II.3.AU.21/SP/2024. Terima kasih juga kami sampaikan kepada Rektor dan Ketua LPPM Universitas Muhammadiyah Pontianak atas dukungan kepada Tim Peneliti sehingga kegiatan Penelitian ini dapat terlaksana dengan baik.

F. Referensi

- [1] L. J. R. Nunes, C. I. R. Meireles, C. J. P. Gomes, and N. M. C. A. Ribeiro, "Forest contribution to climate change mitigation: Management oriented to carbon capture and storage," **Climate**, vol. 8, no. 2, 2020. [Online]. Available: <https://doi.org/10.3390/cli8020021>.
- [2] W. Osei-Wusu, J. Quaye-Ballard, T. Antwi, N. L. Quaye-Ballard, and A. Awotwi, "Forest Loss and Susceptible Area Prediction at Sefwi Wiawso District (SWD), Ghana," **International Journal of Forestry Research**, vol. 2020, 2020. [Online]. Available: <https://doi.org/10.1155/2020/8894639>.
- [3] R. Sousa-Silva, Q. Ponette, K. Verheyen, A. Van Herzele, and B. Muys, "Adaptation of forest management to climate change as perceived by forest owners and managers in Belgium," **Forest Ecosystems**, vol. 3, no. 1, 2016. [Online]. Available: <http://dx.doi.org/10.1186/s40663-016-0082-7>.
- [4] M. Osaki, D. Nursyamsi, and M. S. H. Noor, "Peatland in Indonesia," in **Tropical Peatland Ecosystems**, Springer, Tokyo, Japan, 2016. Available: [doi:10.1007/978-4-431-55681-7_3](https://doi.org/10.1007/978-4-431-55681-7_3).
- [5] J. T. Abatzoglou, A. P. Williams, and R. Barbero, "Global Emergence of Anthropogenic Climate Change in Fire Weather Indices," **Geophysical Research Letters**, vol. 46, no. 1, pp. 326-336, 2019, doi: [10.1029/2018GL080959](https://doi.org/10.1029/2018GL080959).

- [6] T. Preeti, S. Kanakaraddi, A. Beelagi, S. Malagi, and A. Sudi, "Forest Fire Prediction Using Machine Learning Techniques," in *2021 International Conference on Intelligent Technologies (CONIT 2021)*, KLE Technological University, Karnataka, Hubli, India, Institute of Electrical and Electronics Engineers Inc., 2021. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114858479&doi=10.1109%2FCONIT51480.2021.9498448&partnerID=40&md5=a0a63097e5fc5c8af336e391c201d10>.
- [7] Y. Cao, W. Feng, Y. Quan, W. Bao, G. Dauphin, A. Ren, *et al.*, "Forest Disaster Detection Method Based on Ensemble Spatial-Spectral Genetic Algorithm," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 7375–7390, 2022. doi: 10.1109/JSTARS.2022.3199539.
- [8] H. Nahlunnisa and S. Sopiyaandi, "Identifikasi Sebaran Titik Panas (Hotspot) Sebagai Penduga Kebakaran Hutan Dan Lahan Di Kabupaten Ketapang Dan Kayong Utara, Kalimantan Barat," *Jurnal Silva Samalas: Journal of Forestry and Plant Science*, vol. 6, no. 1, pp. 1–5, 2021. [Online]. Available: <https://doi.org/10.33394/jss.v6i1.8362>.
- [9] M. Naderpour, H. M. Rizeei, and F. Ramezani, "Forest fire risk prediction: A spatial deep neural network-based framework," *Remote Sensing*, vol. 13, no. 13, 2021. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85109217838&doi=10.3390%2Frs13132513&partnerID=40&md5=f7141b23418352e9bed86560bcf654a8>.
- [10] M. Mohajane, R. Costache, F. Karimi, Q. B. Pham, A. Essahlaoui, H. Nguyen, *et al.*, "Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area," *Ecological Indicators*, vol. 129, p. 107869, 2021. doi: 10.1016/j.ecolind.2021.107869
- [11] A. S. Nur, Y. J. Kim, J. Lee, and C. W. Lee, "Spatial Prediction of Wildfire Susceptibility Using Hybrid Machine Learning Models Based on Support Vector Regression in Sydney, Australia," *Remote Sensing*, vol. 15, no. 3, 2023, doi: 10.3390/rs15030760.
- [12] L. Gigović, H. R. Pourghasemi, S. Drobnjak, and S. Bai, "Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park," *Forests*, vol. 10, no. 5, 2019, doi: 10.3390/f10050408.
- [13] P. Cortez and A. Morais, "A Data Mining Approach to Predict Forest Fires using Meteorological Data," in *Proc. 13th Portugese Conf. Artificial Intelligence*, 2007, pp. 512-523. Available: <http://www.dsi.uminho.pt/~pcortez/fires.pdf>.
- [14] J. Franklin, "The elements of statistical learning: data mining, inference and prediction," *Mathematical Intelligencer*, vol. 27, no. 2, pp. 83-85, 2005, doi: 10.1007/BF02985802.
- [15] W. Ma, Z. Feng, Z. Cheng, S. Chen, and F. Wang, "Identifying forest fire driving factors and related impacts in China using random forest algorithm," *Forests*, vol. 11, no. 5, 2020, doi: 10.3390/F11050507.
- [16] Y. Jiang, R. Wei, J. Chen, and G. Wang, "Deep learning of qinling forest fire anomaly detection based on genetic algorithm optimization," *UPB Scientific

- Bulletin, Series C: Electrical Engineering and Computer Science*, vol. 83, no. 4, pp. 75-84, 2021.
- [17] T. D'heygere, P. L. M. Goethals, and N. De Pauw, "Genetic algorithms for optimization of predictive ecosystems models based on decision trees and neural networks," *Ecological Modelling*, vol. 195, no. 1-2, pp. 20-29, 2006, doi: 10.1016/j.ecolmodel.2005.11.005.
- [18] Z. Mohammed, C. Hanae, and S. Larbi, "Comparative study on machine learning algorithms for early fire forest detection system using geodata," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 5, pp. 5507-5513, 2020, doi: 10.11591/IJECE.V10I5.PP5507-5513.
- [19] A. R. Al-Shabeeb et al., "A Comparative Study of Genetic Algorithm-Based Ensemble Models and Knowledge-Based Models for Wildfire Susceptibility Mapping," *Sustainability (Switzerland)*, vol. 15, no. 21, 2023, doi: 10.3390/su152115598.
- [20] T. Taslim, D. Toresa, D. Jollyta, D. Suryani, and E. Sabna, "Optimasi K-Means dengan Algoritma Genetika untuk Target Pemanfaat Air Bersih Provinsi Riau," *Indonesian Journal of Computer Science*, vol. 10, no. 1, pp. 1-12, 2022, doi: 10.33022/ijcs.v10i1.3064.
- [21] C. Tan and Z. Feng, "Mapping Forest Fire Risk Zones Using Machine Learning Algorithms in Hunan Province, China," *Sustainability*, vol. 15, no. 7, 2023. [Online]. Available: <https://doi.org/10.3390/su15076292>.
- [22] C. Shi and F. Zhang, "A Forest Fire Susceptibility Modeling Approach Based on Integration Machine Learning Algorithm," *Forests*, vol. 14, no. 7, 2023.