



# Analisa Slope Wilayah Kebakaran Hutan Menggunakan Metode Naïve Bayes

Lisda<sup>\*1</sup>, Nenen Isnaeni<sup>2</sup>, Muhammad Raafi'u Firmansyah<sup>3</sup>

<sup>\*1,2,3</sup>Institut Teknologi Telkom Purwokerto

E-mail: <sup>\*1</sup>lisdaalis@ittelkom-pwt.ac.id, <sup>2</sup>isnaeni@ittelkom-pwt.ac.id,

<sup>3</sup>muhammadraafiu@ittelkom-pwt.ac.id

## Abstract

Forest and land fires are an economic and environmental problem that can cause serious damage. We can predict what factors cause forest fires. It is undeniable that topographical conditions affect the triggering and propagation of fires. The topographical condition itself is in the form of a slope, where fire propagation will be faster when going up the slope than going down the slope. This study aims to match whether the slope and display locations that are prone to the spread of certain fires with high fire intensity actually have a high fire potential and report the magnitude of the influence of the slope in the prediction of fire potential. One of the common approaches to classifying data is to use data mining. So in this study the researchers used the Naive Bayes Classifier as a classification method by getting the highest accuracy value of 0.99%.

**Keywords** : Analysis, Slope, Forest fire, Naive Bayes.

## Abstrak

Kebakaran hutan dan lahan menjadi masalah ekonomi dan lingkungan yang dapat menyebabkan kerusakan serius. Kita dapat memprediksi apa saja faktor yang mengakibatkan kebakaran hutan. Tidak dipungkiri bahwa keadaan topografi berpengaruh pada pemicuan dan penjalaran kobaran api. Kondisi topografi itu sendiri berbentuk kemiringan lereng (slope), dimana penjalaran api akan lebih cepat bila menaiki lereng daripada menuruni lereng. Penelitian ini bertujuan untuk mencocokkan apakah kemiringan lereng (slope) dan menampilkan lokasi yang rawan penyebaran api tertentu dengan intensitas kebakaran tinggi benar-benar memiliki potensi kebakaran yang tinggi dan melaporkan besarnya pengaruh slope didalam prediksi potensi kebakaran. Salah satu pendekatan yang umum dilakukan untuk mengklasifikasikan data adalah dengan menggunakan data mining. Sehingga pada penelitian ini peneliti menggunakan Naive Bayes Classifier sebagai metode klasifikasi dengan mendapatkan nilai akurasi tertinggi sebesar 0.99%.

**Kata Kunci** : Analisa Slope, Kebakaran hutan, Naive Bayes

## I. PENDAHULUAN

Suatu permukaan tanah yang miring dengan sudut tertentu terhadap bidang horizontal disebut lereng. Abramson membedakan lereng menjadi lereng buatan (engineered slope) dan lereng alam (natural slope). Lereng buatan dibuat dari tanah asli dengan memotong tanah tersebut untuk pembuatan jalan atau saluran air untuk irigasi. Lereng alam terbentuk karena

proses alam, lereng-lereng alam yang telah ada selama bertahun-tahun dapat tiba-tiba runtuh dikarenakan adanya gempa, aliran air tanah, hilangnya kuat geser, perubahan tegangan, cuaca dan perubahan topografi [1]. Topografi merupakan gambaran permukaan bumi yang meliputi relief dan posisi alamnya serta ciri-ciri yang merupakan hasil dari buatan manusia dan salah satu dari faktor kebakaran hutan.



Tidak dipungkiri bahwa keadaan topografi berpengaruh pada kemiringan. Kemiringan (slope) salah satu faktor terpenting yang mempengaruhi bahaya kebakaran hutan dengan asumsi tidak ada angin, api dan panas meningkat tergantung pada kemiringan dan juga menjadi pemicu, penjalaran kobaran api pada saat terjadi kebakaran. Dimana penjalaran api akan lebih cepat bila menaiki lereng daripada menuruni lereng. Dan semakin curam lereng atau sudut kemiringan semakin besar, maka kecepatan penjalaran api akan semakin besar pula.

Kebakaran hutan yang terjadi di Indonesia disebabkan oleh manusia dan didukung oleh kekeringan yang panjang sebagai akibat dari faktor rendahnya nilai curah hujan yang terjadi di beberapa daerah dalam waktu yang cukup lama. Kebakaran hutan dapat dipengaruhi oleh keadaan cuaca yang meliputi keadaan suhu, kecepatan angin, dan curah hujan[2].

Data Mining adalah proses ekstraksi suatu data yang bersifat implisit dan tidak berguna menjadi informasi dari data yang jumlahnya besar. Data mining merupakan proses untuk menemukan korelasi, pola, dan tren baru yang bermakna dengan memilah-milah data dalam jumlah besar yang disimpan di dalam repositori, menggunakan teknologi pengenalan

pola serta teknik statistik dan matematika[3].

*Website* adalah sebuah sarana informasi yang dimanfaatkan dalam layanan *internet* [4].

Pada penelitian terdahulu yang berjudul *Analysis of The Stability Variation of a Slope Crossed by Forest Fire oleh Romina Secci*[5], membahas tentang variasi stabilitas lereng bukit yang dilintasi kebakaran hutan. Untuk melakukan analisis stabilitas, peneliti menggunakan dua model geoteknik diturunkan dengan integrasi hasil survei geoteknik (Pengukuran NSPT dan uji geser langsung) dengan tomografi resistivitas listrik. Pendekatan klasik berdasarkan Limit Metode kesetimbangan digunakan untuk menentukan koefisien keamanan. Selanjutnya, efek kebakaran hutan di lereng stabilitas telah dibahas memodifikasi model 2D dan memperkenalkan lapisan tipis ultra dangkal dengan kekuatan geser parameter yang ditentukan pada sampel tanah yang terbakar, dikumpulkan setelah penyeberangan api.

Penelitian lain yang dilakukan Ugur Baltaci dan Feriha Yildirm, menyajikan penelitian mengenai Effect of Slope on the Analysis of Forest Fire Risk [6]. Penelitian ini mengevaluasi pengaruh kemiringan pada risiko kebakaran hutan dengan sistem informasi geografis (GIS) dan titik-titik kebakaran hutan yang telah



terjadi di Turki dalam 10 tahun terakhir. Dengan demikian, penilaian yang objektif coba dilakukan tentang pengaruh kemiringan lereng terhadap risiko kebakaran hutan. Menurut hasil analisis, 40,53% kebakaran hutan terjadi di wilayah Turki kemiringan 0-10%. Selain itu, 87,16% kebakaran ini terjadi pada kemiringan 0-30%.

Mikael Lundbark dkk dalam penelitiannya membahas tentang Global Analysis of the Slope of Forest Land [7]. Pada penelitian ini memprediksi kemungkinan regional untuk peralatan panen yang berbeda, hasilnya dapat digunakan untuk menghubungkan variabel hutan geografis dengan kemiringan. Dan juga dapat digunakan dalam pemadaman kebakaran hutan strategis dan perencanaan konservasi dan pengelolaan hutan skala besar.

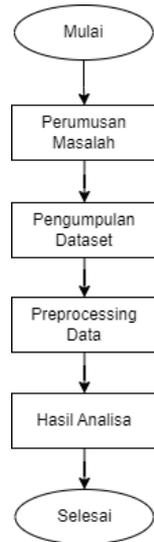
Penelitian yang juga dilakukan oleh Mohammad Reza Alizadeh dkk dimana melakukan penelitian mengenai pemanasan yang memungkinkan terjadinya kenaikan lereng (slope) akibat kebakaran hutan. Dengan judul penelitian Warming Enabled Upslope Advance in Western US Forest Fire [8]. Peneliti fokus pada distribusi elevasi hutan kebakaran di ekoregion pegunungan di Amerika Serikat bagian barat dan menunjukkan tingkat kenaikan terbesar di area terbakar di atas 2.500 m selama 1984 hingga 2017. Selanjutnya, peneliti menunjukkan bahwa elevasi

tinggi menembakkan lereng ke atas dengan perubahan kumulatif rata-rata 252 m (-107 hingga 656 m; 95% CI) dalam 34 tahun di seluruh ekoregion yang diteliti. Peneliti juga mendokumentasikan hubungan antar tahunan yang kuat antara dataran tinggi kebakaran dan defisit tekanan uap musim panas (VPD) lereng atas kemajuan kebakaran konsisten dengan pemanasan yang diamati yang dipantulkan oleh penyimpangan lereng rata-rata dari isoline VPD 295 m (59 hingga 704 m; 95% CI) selama 1984 hingga 2017. Salah satu pendekatan yang umum dilakukan untuk mengklasifikasikan data adalah dengan menggunakan data mining.

Pada penelitian ini peneliti menggunakan Naive Bayes Classifier sebagai metode klasifikasi. Penelitian ini bertujuan untuk mencocokkan apakah kemiringan lereng (slope) dan menampilkan lokasi yang rawan penyebaran api tertentu dengan intensitas kebakaran tinggi benar-benar memiliki potensi kebakaran yang tinggi dan melaporkan besarnya pengaruh slope didalam prediksi potensi kebakaran dalam sebuah website.

## **II. METODE PENELITIAN**

Alur penelitian dalam analisa slope wilayah kebakaran hutan ditunjukkan pada Gambar 1.



Gambar 1. Metode penelitian

### 2.1. Perumusan Masalah

Perumusan masalah merupakan langkah utama dalam metode investigasi. Dalam penelitian, rumusan masalah adalah metode prediksi daerah rawan kebakaran dengan menggunakan algoritma naive bayes, sehingga diharapkan akan tercipta model yang akurat.

### 2.2. Pengumpulan Dataset

Pada penelitian ini dataset diperoleh dari peta digital elevation map (DEM) dari <https://tanahair.indonesia.go.id/demnas/> [9]. Dengan menggunakan tool ArcGis untuk mengekstraksi fitur dan dataset ini belum pernah digunakan pada penelitian sebelumnya. Peta DEM yang kami gunakan terbatas hanya pada daerah kecamatan tanah Siang Kalimantan Tengah. Hasil dari ekstraksi fitur memperoleh dataset sebanyak 40204 data. Dataset yang diperoleh belum pernah digunakan pada

penelitian sebelumnya dapat dilihat pada Gambar 2.

OBJECTID	LOCATIONID	X	Y	Elevation	Slope_Degree	Slope_Percent	NEAR_FID	RIVER_DIST	ROAD_DIST
0	1	506615.108120	9.988430e+06	506.728485	2.781732	4.858856	302	2.236306	9462.735725
1	2	502526.215	9.988430e+06	502.526215	1.979101	3.455567	302	51.417961	9590.133321
2	3	501268.280	9.988430e+06	501.268280	2.493582	4.354872	302	10.058789	9719.387342
3	4	504655.731	9.988430e+06	504.655731	3.427884	5.989934	302	92.557718	9850.424710
4	5	498521.057	9.988430e+06	498.521057	5.475194	9.885211	302	176.871090	9983.175202
...	...	...	...	...	...	...	...	...	...
40199	40200	561705.36	9.925672e+06	56.170536	0.132954	0.232048	107	230.618039	689.011386
40200	40201	475345.46	9.925672e+06	47.534546	0.531040	0.920807	35	64.540076	43.839876
40201	40202	507397.65	9.925672e+06	50.739765	0.871722	1.521568	7	146.579560	48.861224
40202	40203	540215.91	9.925672e+06	54.021591	0.720402	1.257406	107	216.661311	181.120547
40203	40204	560371.40	9.925672e+06	56.037140	0.491555	0.857947	107	319.405148	292.089247

Gambar 2. Dataset

### 2.3 Preprocessing Data

Pada tahapan ini, data yang telah diperoleh dan dikumpulkan kemudian diproses kembali. Data diproses menggunakan algoritma k-means clustering dengan dua cluster yaitu 0 dan 1. Untuk mendapatkan hasil yang sesuai, kami menggunakan 4 atribut yang digunakan dalam clustering yaitu elevation, slope degree, river distance dan road distance. Hasil dari cluster tersebut akan digunakan sebagai pengelompokan data potensi penyebaran api di suatu area, dimana nilai 0 sebagai potensi penyebaran api rendah sedangkan 1 sebagai potensi penyebaran api tinggi. Keterangan dataset dapat dilihat pada Gambar 3.

	Elevation	Slope_Degree	RIVER_DIST	ROAD_DIST	Class
0	506.728485	2.781732	2.236306	9462.735725	1
1	502.526215	1.979101	51.417961	9590.133321	1
2	501.268280	2.493582	10.058789	9719.387342	1
3	504.655731	3.427884	92.557718	9850.424710	1
4	498.521057	5.475194	176.871090	9983.175202	1
...	...	...	...	...	...
40199	56.170536	0.132954	230.618039	689.011386	0
40200	47.534546	0.531040	64.540076	43.839876	0
40201	50.739765	0.871722	146.579560	48.861224	0
40202	54.021591	0.720402	216.661311	181.120547	0
40203	56.037140	0.491555	319.405148	292.089247	0

Gambar 3. Hasil proses data dengan K-mean clustering



### 2.3 Analisa

Analisa merupakan tahapan yang dilakukan setelah mengumpulkan data dari metode penelitian. Analisis adalah cara khusus untuk menganalisis suatu masalah. Pada tahap analisis proses, peneliti menelusuri langkah-langkah menemukan prediksi menggunakan algoritma dalam data mining.

1) Data Selection : Pada fase ini dilakukan langkah-langkah pemilihan data. Seleksi data dalam proses seleksi menggunakan 4 atribut yaitu elevation, slope degree, river dist, dan road dist. Data yang tidak disertakan pada variabel penelitian yang digunakan akan dibiarkan tidak terpakai.

2) K-Means Clustering : K-Means clustering merupakan algoritma yang memiliki kemampuan untuk mengelompokkan data menjadi data cluster. Algoritma ini dapat menerima data tanpa kategorisasi, juga mengelompokkan beberapa data ke dalam kelompok-kelompok yang menyatakan bahwa data dalam satu kelompok memiliki karakteristik yang sama dan berbeda dengan data pada kelompok lainnya.

Langkah-langkah algoritma K-Means adalah sebagai berikut[10]:

- a. Menentukan nilai k atau jumlah Cluster pada data set.
- b. Menentukan nilai pusat (centroid). Penentuan nilai centroid pada tahap

awal dilakukan secara random, sedangkan pada tahap iterasi digunakan rumus seperti:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (1)$$

Keterangan:

$V_{ij}$  = Centroid rata-rata cluster ke- $i$  untuk variabel ke- $j$

$N_i$  = Jumlah anggota cluster ke- $i$

$k$  = Indeks dari cluster

$j$  = Indeks dari variabel

$X_{kj}$  = nilai data ke- $k$  variabel ke- $j$  untuk cluster tersebut.

c. Menghitung jarak antara titik centroid dengan titik tiap objek menggunakan Euclidean Distance. Euclidean Distance merupakan jarak garis lurus biasa antara dua titik dalam ruang Euclidean, dengan rumus seperti dibawah ini:

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2)$$

Keterangan:

$D_e$  = Euclidean Distance

$i$  = Banyaknya objek

$(x, y)$  = Koordinat objek

$(s, t)$  = Koordinat centroid

3) Naive Bayes Classifier : Algoritma naive bayes classifier adalah salah satu statistik classifier yang dapat memprediksi probabilitas keanggotaan kelas data yang classifier ini diklasifikasikan ke dalam kelas tertentu sesuai dengan perhitungan probabilitas. Naive Bayes dapat disebut juga dengan



simple bayes dan independence bayes [11]. Formula dari teorema naive bayes:

$$P(H|X) = \frac{p(H|X)p(H)}{p(X)} \quad (3)$$

Keterangan :

$P(H|X)$  = Probabilitas akhir bersyarat (conditional probability) suatu hipotesis  $H$  terjadi jika diberikan bukti (evidence)  $X$  terjadi.

$P(X|H)$  = Probabilitas sebuah bukti  $X$  terjadi akan mempengaruhi hipotesis  $H$ .

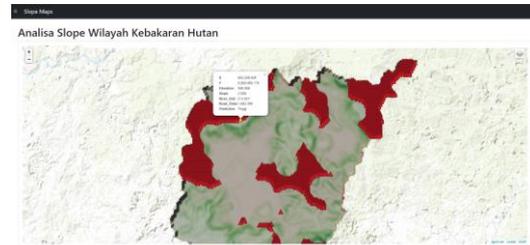
$P(H)$  = Probabilitas awal (priori) hipotesis  $H$  terjadi tanpa memandang bukti apapun.

$P(X)$  = Probabilitas awal (priori) bukti  $X$  terjadi tanpa memandang hipotesis atau bukti yang lain.

### III. HASIL DAN PEMBAHASAN

Hasil dari implementasi website disini peneliti menggunakan tools dari aplikasi QGis. Sebelum peta dapat ditampilkan pada website peneliti melakukan pra-pemrosesan data terhadap data set yang sudah diekstraksi dari peta Digital Elevation Model (DEM). Pada tahap pra-pemrosesan data menggunakan algoritma K-Means Clustering, atribut yang digunakan ada empat antara lain elevation, slope degree, river distance, dan road distance. Akan tetapi atribut  $x$  dan  $y$  tetap disimpan agar dapat menampilkan hasil pengelompokan

data pada peta di website. Berikut merupakan hasil analisa slope pada website, ditampilkan pada gambar 4.



gambar 4. Hasil implementasi website analisa slope

Setelah dilakukan pengklasifikasian dengan menggunakan algoritma naive bayes. Akurasi yang diperoleh adalah sebesar 0.99 yang dihitung dengan menggunakan Confusion Matrix. Gambar di bawah ini memaparkan perbandingan hasil precision, recall, dan F1-score, antara potensi penyebaran api rendah sama dengan 0 dan penyebaran api tinggi sama dengan 1. Untuk potensi penyebaran api rendah sama dengan 0 memiliki nilai precision, recall, dan f1-score sebesar 0.99, 0.99, 0.99, sedangkan untuk potensi penyebaran api tinggi sama dengan 1 memiliki nilai sebesar 0.97, 0.97, dan 0.97. Untuk mempermudah dalam pembacaan data, keterangan dapat dilihat pada gambar 5.

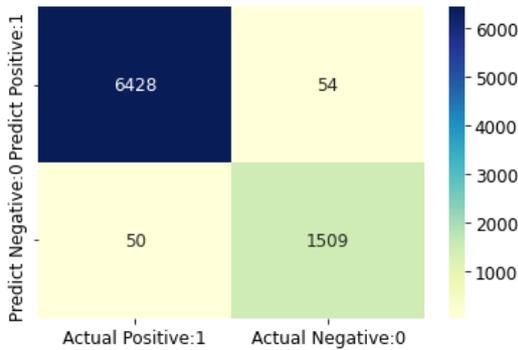
	precision	recall	f1-score	support
0	0.99	0.99	0.99	6482
1	0.97	0.97	0.97	1559
accuracy			0.99	8041
macro avg	0.98	0.98	0.98	8041
weighted avg	0.99	0.99	0.99	8041

Gambar 5. Hasil akurasi

Pada gambar 6 merupakan hasil visualisasi confusion matrix dari Naive bayes dengan perhitungan accuracy.

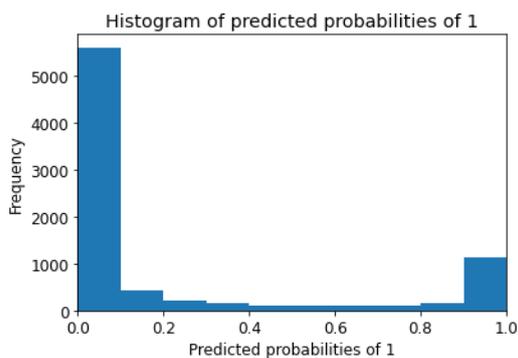


Dari gambar dapat dilihat bahwa terdapat 6428 nilai true positif dan 54 false positif. Kemudian 50 nilai false negative dan ada 1509 true negative.



Gambar 6. Hasil confusion matrix

Prediksi probabilitas rendah sama dengan 0 dan tinggi sama dengan 1. Berdasarkan hasil histogram area dengan potensi penyebaran api rendah lebih besar dari area potensi penyebaran api tinggi. Hasil nilai prediksi probabilitas dari potensi penyebaran api rendah dan tinggi dapat dilihat pada Gambar 7.



Gambar 7. Histogram Predict Probabilities

#### IV. KESIMPULAN

Pada penelitian sebelumnya yang pernah dilakukan di Turki dengan

judul “Effect of Slope on the Analysis of Forest Fire Risk” menggunakan GIS dalam menganalisis data slope mendapatkan hasil 40,53% kebakaran hutan dimulai di daerah dengan kemiringan 0-10%. Juga, 23,57% kebakaran terjadi di daerah dengan kemiringan 11-20% dan 7,09% kebakaran terjadi di daerah dengan kemiringan 21-30%. Dengan kata lain, 87,16% kebakaran dimulai di daerah dengan kemiringan 0-30%. Dibandingkan dengan penelitian yang dilakukan saat ini bertujuan untuk mencocokkan apakah kemiringan lereng (slope) dan lokasi yang rawan penyebaran api tertentu dengan intensitas kebakaran tinggi benar-benar memiliki potensi kebakaran yang tinggi dan melaporkan besarnya pengaruh slope di dalam prediksi potensi kebakaran. Menggunakan algoritma k-means clustering untuk pemrosesan dan pengumpulan data dengan dua cluster yaitu 0 dan 1. Setelah itu, dilakukan pengklasifikasian dengan menggunakan algoritma naive bayes dengan akurasi yang diperoleh sebanyak 99%. Akurasi 99% disini didapat dengan kondisi dataset dari peta digital elevation model (DEM) kecamatan Tanah Siang Kalimantan Tengah, bukan dari data yang pernah terjadi kebakaran. Berikut detail hasil perbandingan precision, recall dan f1-score dapat dilihat pada Tabel 1.



Tabel 1. Perbandingan hasil

	Precision	Recall	F1-Score
Potensi Penyebaran Api Rendah	0,99	0,99	0,99
Potensi Penyebaran Api Tinggi	0,97	0,97	0,97
Accuracy	0,99		

*Informasi Galuh*, 1(2), 52–58. <https://doi.org/10.25157/jsig.v1i2.3203>

[5] Baltacı Ugur and Yıldırım Feriha, "Effect of Slope on the Analysis of Forest Fire Risk," *Hacettepe Journal of Biology and Chemistry*, vol. 48, pp. 373–379, 2020.

[6] Lundbäck Mikael, Persson Henrik, Haggström Carola, and Nordfjell Tomas, "Global analysis of the slope of forest land," *Forestry An International Journal of Forest Research*, Jun. 2020.

[7] Alizadeh Mohammad Reza, Abatzoglou John T., Luce Charles H., Adamowski Jan F., Farid Arvin, and Sadegh Mojtaba, "Warming enabled upslope advance in western US forest fires," Jun. 2022.

[8] Sutoyo Muh.Nurtanzis, "Algoritma K-Means."

[9] <https://tanahair.indonesia.go.id/demnas/>.

[10] Fitriyani, "Prediksi Diabetes Menggunakan Algoritma Naïve Bayes dan Greedy Forward Selection," *Jurnal Nasional Teknologi dan Sistem Informasi*, 2021.

[11] Janssens A Cecile J W and Martens Forike K, "Reflection on modern methods: Revisiting the area under the ROC Curve," *International Journal of Epidemiology*, 2020.

DAFTAR PUSTAKA

[1] Asyrowi Hamam, Saharjo Bambang Hero, and Putra Erianto Indra, "Pola Sebaran Hotspot di Taman Hutan Raya Raden Soerjo (Hotspot Distribution Patterns in Raden Soerjo Grand Forest Park)," *Jurnal Penelitian Hutan dan Konservasi Alam*, 2021.

[2] Rudiyan Ari, Dzulkifli Akhmad Erik, and Munazar Khabib, "Klasifikasi Kebakaran Hutan Menggunakan Metode K-Nearest Neighbor: Studi Kasus Hutan Provinsi Kalimantan Barat," *Jurnal Teknologi Informasi dan Multimedia*, vol. Vol. 3, No. 4, pp. 195–202, Feb. 2022.

[3] Pratiwi Trya Ayu, Irsyad Muhammad, and Kurniawan Rahmad, "Klasifikasi Kebakaran Hutan dan Lahan Menggunakan Algoritma Naïve Bayes (Studi Kasus: Provinsi Riau)," *Jurnal Sistem Dan Teknologi Informasi*, vol. Vol.9 No.2, Apr. 2021.

[4] Usep Abdul Rosid, Sidiq, M., Mulyana, D., & Yudi Permana, N. (2023). Website Design with Waterfall Method in Ciamis Regency (Case Study in Galuh Ciamis Nature and Environment Care Community) . *Jurnal Sistem*