

Pemetaan Daerah Rawan Longsor di Kabupaten Bandung Barat menggunakan Metode *Machine Learning* dengan Teknik SVM

Afif Faadhilah, Hary Nugroho

Program Studi Teknik Geodesi, Institut Teknologi Nasional Bandung

Email: afiffaadhilah3@gmail.com

Received 12 Mei 2024 | Revised 20 Mei 2024 | Accepted 28 Mei 2024

ABSTRAK

Kabupaten Bandung Barat memiliki kerawanan tinggi terhadap bencana tanah longsor karena karakteristik topografinya yang berupa perbukitan dan pegunungan. Kerawanan ini semakin meningkat akibat perubahan fungsi lahan dari kawasan vegetasi menjadi area permukiman. Penelitian ini bertujuan untuk memberikan informasi wilayah yang memiliki kerawanan tanah longsor di Kabupaten Bandung Barat. Informasi ini disajikan dalam bentuk peta yang diperoleh melalui proses klasifikasi data yang berkontribusi terhadap terjadinya tanah longsor dengan menggunakan metode Support Vector Machine (SVM). Metode SVM dipilih karena keunggulannya dalam mengolah data yang kompleks dan menghasilkan prediksi yang akurat. Penelitian ini menggunakan enam parameter: curah hujan, jenis batuan, jenis tanah, kemiringan lereng, kerapatan vegetasi, dan penggunaan lahan, serta data kejadian longsor dari tahun 2015 hingga 2023. Dataset terdiri atas 550 titik, dengan 326 titik longsor dan 214 titik nonlongsor, yang dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian. Model divalidasi menggunakan 128 titik kejadian longsor dari tahun 2022 hingga 2023. Hasil prediksi menunjukkan klasifikasi wilayah dengan potensi longsor rendah sebesar 42% (nilai 0), potensi menengah sebesar 28% (nilai 0,50), dan potensi tinggi sebesar 30% (nilai 1). Wilayah dengan potensi longsor rendah meliputi Kecamatan Cipeundeuy dan Batujajar. Kecamatan Rongga sebagian besar memiliki potensi menengah, sedangkan Kecamatan Cililin dan Cisarua sebagian besar memiliki potensi tinggi. Berdasarkan perhitungan metode feature importance, ditemukan bahwa kemiringan lereng adalah variabel yang paling berpengaruh dengan nilai kepentingan sebesar 0,35. Model ini menunjukkan performa yang cukup baik dan tidak mengalami overfitting karena memiliki akurasi pelatihan sebesar 79% yang hampir sepadan dengan akurasi validasi sebesar 76%.

Kata kunci: Longsor, Machine Learning, Pemetaan, dan Support Vector Machine.

ABSTRACT

West Bandung Regency is highly susceptible to landslides due to its topographic characteristics, which consist of hills and mountains. This susceptibility is further increased by land use changes from vegetated areas to residential areas. This study aims to provide information on areas prone to landslides in West Bandung Regency. This information is presented as a map obtained through the classification process of data contributing to landslides using the Support Vector Machine (SVM) method. The SVM method was chosen for its advantages in handling complex data and providing accurate predictions. This study used six parameters: rainfall, rock type, soil type, slope gradient, vegetation density, and land use, as well as landslide occurrence data from 2015 to 2023. The dataset consists of 550 points, with 326 landslide points and 214 non-landslide points, divided into 70% for training and 30% for testing. The model was validated using 128 landslide occurrence points from 2022 to 2023. The prediction results showed the classification of areas with low landslide potential at 42% (value 0), medium potential at 28% (value 0.50), and high potential at 30% (value 1). Areas with low landslide potential include Cipeundeuy and Batujajar Districts. Rongga District is mainly characterised by medium potential, while

Cililin and Cisarua Districts are primarily characterised by high potential. Based on the feature importance calculation, it was found that the slope gradient is the most influential variable, with an importance value of 0.35. The model performed well and did not overfit, as it had a training accuracy of 79%, nearly equal to the validation accuracy of 76%.

Keywords: *Landslide, Machine Learning, Mapping, and Support Vector Machine.*

1. PENDAHULUAN

Tanah longsor sering terjadi di daerah dengan topografi curam dan curah hujan tinggi. Di Indonesia, fenomena ini diperburuk oleh aktivitas manusia seperti pengalihfungsian hutan menjadi perkebunan dan pemukiman, yang mengurangi kapasitas tanah untuk menyerap air, sehingga meningkatkan risiko longsor. Penggunaan sumber daya alam yang tidak teratur dan melampaui daya dukungnya juga dapat memicu bencana ini. Data dari Badan Nasional Penanggulangan Bencana (BNPB) menunjukkan bahwa pada tahun 2023, terdapat 579 kejadian tanah longsor di Indonesia, menjadikannya bencana dengan intensitas tertinggi keempat setelah kebakaran hutan, cuaca ekstrem, dan banjir [1].

Kabupaten Bandung Barat di Provinsi Jawa Barat merupakan daerah dengan potensi tinggi terhadap tanah longsor karena topografinya yang berupa perbukitan dan pegunungan. Peningkatan jumlah penduduk menyebabkan lahan pemukiman semakin terbatas, memaksa banyak orang tinggal di daerah rawan longsor. Pada lima bulan pertama tahun 2023, Kabupaten Bandung Barat telah mengalami 133 kejadian bencana alam, dimana 65 kejadian diantaranya adalah tanah longsor [2].

Kondisi ini memerlukan perhatian khusus untuk mitigasi dan pemetaan daerah rawan longsor guna mengurangi kerugian baik material maupun nonmaterial, serta membantu pemerintah meningkatkan kesiapsiagaan terhadap bencana. Pembelajaran Mesin (*Machine Learning*), sebagai bagian dari kecerdasan buatan, dapat digunakan untuk memprediksi kejadian tanah longsor berdasarkan data historis. Algoritma *Support Vector Machine* (SVM) sering digunakan untuk pemetaan kerawanan tanah longsor karena kemampuannya dalam menangani data yang kompleks dan memberikan prediksi akurat [3], [4].

Penelitian ini bertujuan untuk memberikan informasi wilayah yang memiliki kerawanan tanah longsor di Kabupaten Bandung Barat. Informasi ini disajikan dalam bentuk peta yang diperoleh melalui proses klasifikasi data yang berkontribusi terhadap terjadinya tanah longsor dengan menggunakan algoritma *machine learning* dengan teknik SVM. Adapun data yang dipergunakan meliputi kemiringan lereng, curah hujan, jenis tanah, jenis batuan, kerapatan vegetasi, dan sejarah kejadian longsor.

2. METODOLOGI

2.1 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang terdiri atas data spasial dalam format *vektor* dan *raster*, serta data atribut yang relevan untuk analisis kerawanan tanah longsor di Kabupaten Bandung Barat. Data ini dikumpulkan dari berbagai instansi terkait. Tabel 1 memberikan rincian lengkap mengenai jenis data yang digunakan, format data, sumber data, dan tahun pengumpulan data.

Tabel 1. Data Penelitian Kerawanan Tanah Longsor

No	Data	Sumber Data	Tahun
1	Data Historis Kejadian Longsor	BPBD Provinsi Jawa Barat	2015 – 2023
2	Data Curah Hujan Skala 1:25.000	Stasiun Klimatologi Jawa Barat	2022
3	Peta Jenis Batuan Skala 1:25.000	Pusat Survei Geologi	2010
4	Peta Jenis Tanah Skala 1:25.000	Pusat Survei Geologi	2010
5	Peta Batas Administrasi Skala 1:25.000	BIG	2021
6	Peta Penggunaan Lahan Skala 1:25.000	BAPPEDA	2021
7	DEM Nasional Skala 1:25.000	BIG	2018

No	Data	Sumber Data	Tahun
8	Citra Sentinel 2A 20231220 Resolusi 20 m	Copernicus	2023

2.2 Machine Learning SVM

SVM adalah teknik klasifikasi dalam *supervised learning* yang menggabungkan teori-teori komputasi yang telah ada selama puluhan tahun, seperti konsep margin dan *hyperplane* yang diperkenalkan oleh Aronszajn pada tahun 1950, serta konsep-konsep pendukung lainnya. Metode ini memerlukan dua jenis training set, yaitu positif dan negatif, untuk membuat keputusan terbaik dalam memisahkan data positif dari data negatif di ruang n-dimensi menggunakan *hyperplane*. Konsep kerjanya adalah SVM berusaha untuk mencari *hyperplane* pemisah atau *discrimination boundaries*, yang optimal dan memaksimalkan *margin* antara dua kelas [5]. Proses klasifikasi dalam SVM dapat diartikan sebagai usaha untuk menemukan *hyperplane* yang memisahkan dua kelompok data tersebut. Berbagai alternatif *hyperplane* ditunjukkan pada Gambar 2. *Hyperplane* terbaik yang memisahkan kedua kelas ditemukan dengan cara mengukur *margin hyperplane* dan mencari nilai maksimumnya. *Margin* ini mengacu pada jarak antara *hyperplane* dan pola terdekat dari masing-masing kelas, yang dikenal sebagai *support vector*. Gambar 2 di sebelah kanan menunjukkan *hyperplane* terbaik yang terletak di tengah-tengah antara kedua kelas, sedangkan titik merah dan kuning dalam lingkaran hitam menunjukkan *support vector*. Proses untuk menentukan posisi *hyperplane* ini adalah inti dari proses pelatihan pada SVM. Setiap data (contoh) diberi notasi $x_i \in \mathbb{R}^D$, $i = 1, 2, \dots, N$, dimana N adalah jumlah total data. Kelas positif diberi notasi +1 dan kelas negatif diberi notasi -1. Oleh karena itu, setiap data dan label kelasnya diberi notasi sebagai $y_i \in \{-1, +1\}$. Diasumsikan bahwa kedua kelas dapat dipisahkan dengan sempurna oleh *hyperplane* dalam ruang fitur berdimensi D (*D-dimensional feature space*).

Hyperplane didefinisikan sebagai berikut [6]:

$$w \cdot x_i + b = 0 \tag{2.1}$$

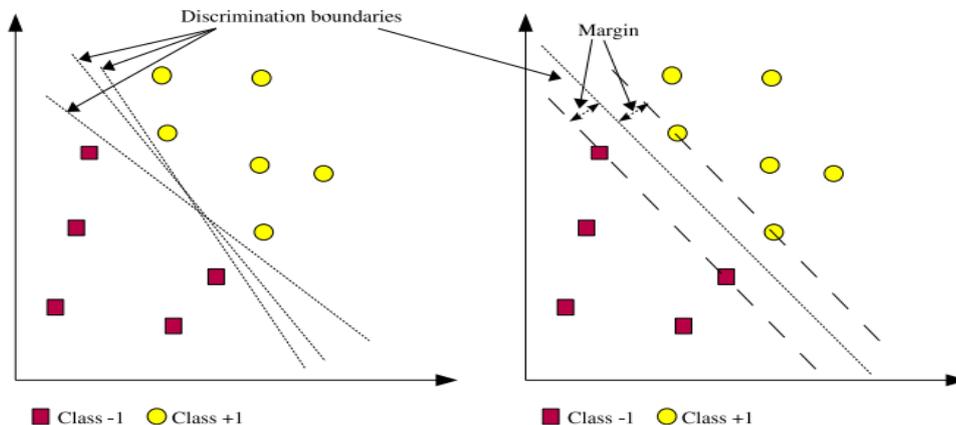
Data x_i digolongkan kedalam kelas negatif karena memenuhi pertidaksamaan berikut:

$$w \cdot x_i + b \leq -1 \tag{2.2}$$

Adapun data x_i yang digolongkan kedalam kelas positif karena memenuhi pertidaksamaan:

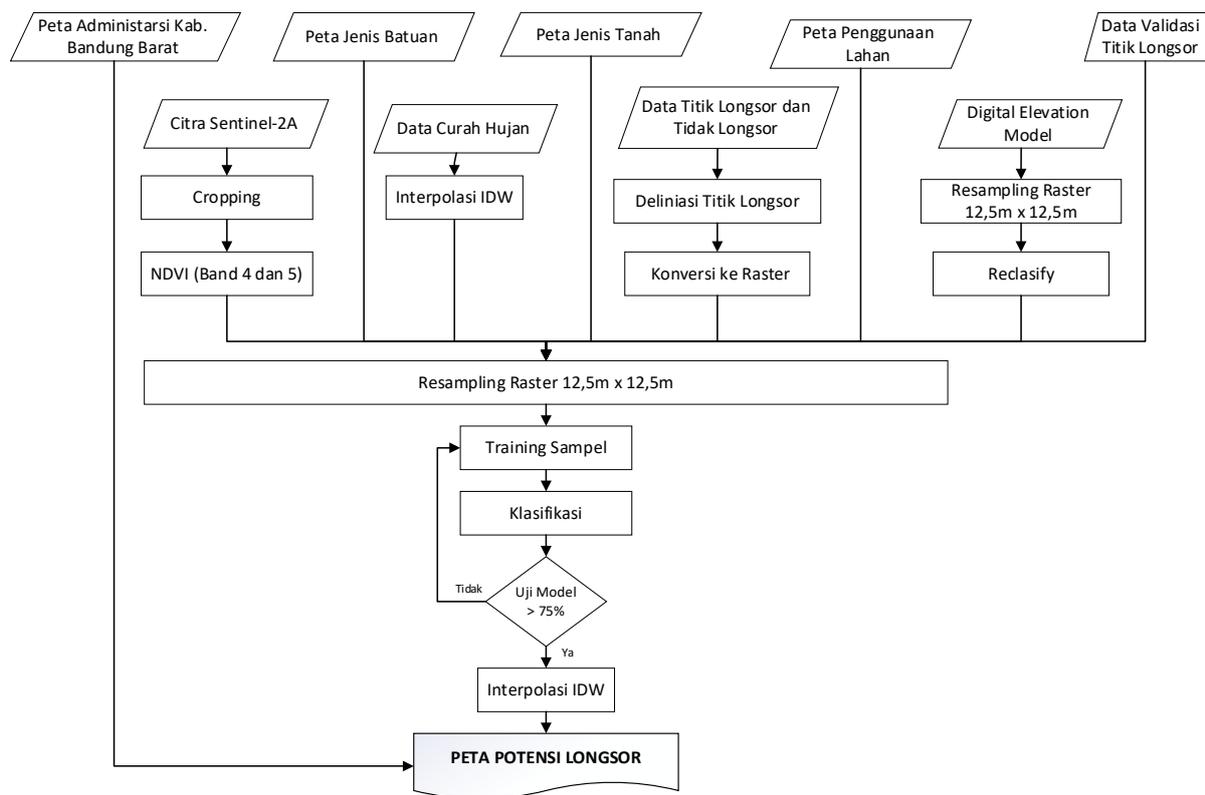
$$w \cdot x_i + b \geq 1 \tag{2.3}$$

Margin yang optimal dihitung dengan memaksimalkan jarak antara *hyperplane* dan *pattern* terdekat.



Gambar 2. Konsep SVM dalam menemukan *hyperplane* terbaik antara kelas positif dan negatif [7]

2.3 Diagram Alir Penelitian



Gambar 3. Diagram Alir Penelitian

2.4 Kelebihan dan Kekurangan *Support Vector Machine*

Menurut Nugroho [7], algoritma SVM memiliki berbagai kelebihan dan juga kekurangan. Hal ini dapat dijelaskan sebagai berikut:

Kelebihan:

- SVM unggul dalam mengklasifikasikan pola data, terutama melalui generalisasi yang optimal dengan meminimalkan kesalahan pada data pelatihan. Dengan memilih *hyperplane* yang memiliki margin terbesar, SVM dapat menentukan batas keputusan yang optimal dan stabil, sehingga memberikan prediksi yang akurat bahkan pada data yang kompleks dan tidak terlihat selama proses pelatihan. Penelitian empiris menunjukkan bahwa SVM umumnya menghasilkan error generalisasi yang lebih rendah dibandingkan dengan metode pembelajaran mesin lainnya, menjadikannya pilihan yang efektif untuk berbagai masalah klasifikasi.
- SVM relatif mudah diimplementasikan karena tahapan penentuan *support vector* dapat dikategorikan sebagai masalah optimasi kuadrat (*QP problem*). Dengan adanya pustaka yang mampu menyelesaikan masalah optimasi kuadrat, implementasi SVM menjadi lebih sederhana. Selain itu, SVM juga dapat dipecahkan dengan metode sekuensial, yang mempermudah proses implementasi dan pemrosesan data secara berurutan, menjadikannya pilihan praktis dan efisien untuk berbagai aplikasi.

Kekurangan:

- Secara teori, SVM dikembangkan untuk masalah klasifikasi dengan dua kelas. Meskipun telah dimodifikasi untuk menangani masalah dengan lebih dari dua kelas menggunakan strategi seperti *One versus Rest* dan *Tree Structure*, setiap strategi memiliki kelemahan masing-masing. Oleh karena itu, penelitian dan pengembangan SVM untuk masalah multikelas masih merupakan area yang terus berkembang.
- SVM dapat menjadi sulit diterapkan pada masalah dengan skala besar, yang dalam hal ini berarti jumlah sampel yang sangat besar.

2.5 Inverse Distance Weight (IDW)

Dalam penelitian ini metode IDW digunakan untuk melakukan estimasi kerawanan longsor pada wilayah yang berada di antara wilayah rawan dan tidak rawan. Metode ini diterapkan setelah berakhirnya proses klasifikasi menggunakan SVM. Dengan diterapkannya metode ini maka akan diketahui wilayah-wilayah yang memiliki tingkat kerawanan rendah, menengah dan tinggi yang dihitung berdasarkan jarak dari daerah yang dinyatakan rawan longsor dan tidak longsor.

Metode IDW adalah teknik estimasi yang menggunakan model blok sederhana dengan mempertimbangkan titik-titik terdekat di sekitarnya. Metode interpolasi ini didasarkan pada asumsi bahwa pengaruh titik terhadap estimasi nilai semakin besar jika jaraknya lebih dekat dengan titik yang tidak diketahui nilainya. Dalam metode IDW, nilai yang terukur pada titik-titik di sekitar lokasi digunakan untuk memperkirakan nilai variabel pada lokasi yang ditargetkan. Asumsi utama IDW adalah bahwa titik yang lebih dekat ke lokasi estimasi akan memberikan pengaruh yang lebih besar dibandingkan dengan titik yang lebih jauh [8][9].

Menurut Pasaribu dkk. [10], fungsi pembobotan umum yang digunakan dalam metode ini adalah invers dari kuadrat jarak, yang dirumuskan dalam formula berikut:

$$Z^* = \sum_{i=1}^N \omega_i Z_i \tag{2.4}$$

dimana,

$$Z_i \ (i = 1,2,3, \dots, N)$$

merupakan nilai ketinggian data yang ingin diinterpolasi pada sejumlah N titik, dan bobot (*weight*) ω_i yang dirumuskan sebagai:

$$\omega_i = \frac{h_j^{-P}}{\sum_{j=0}^n h_i^{-P}} \tag{2.5}$$

P adalah nilai positif, disebut sebagai *parameter power* (biasanya bernilai 2). h_i merupakan jarak dari titik interpolasi ke sebaran titik. Hal ini dinyatakan dalam persamaan berikut:

$$h_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{2.6}$$

Koordinat titik interpolasi dinyatakan sebagai (x, y), sedangkan (x_i, y_i) mewakili koordinat dari setiap titik yang tersebar. Fungsi perubahan bobot bervariasi untuk seluruh data titik sebaran hingga mendekati nol seiring dengan bertambahnya jarak dari titik-titik tersebut.

Keunggulan dari metode interpolasi IDW adalah kemampuannya untuk mengontrol karakteristik interpolasi dengan membatasi titik-titik masukan yang digunakan dalam proses interpolasi. Titik-titik yang jauh dari titik sampel dan diperkirakan memiliki korelasi spasial yang rendah atau bahkan tidak ada korelasi spasial dapat diabaikan dalam perhitungan. Titik-titik yang digunakan dalam interpolasi dapat ditentukan secara langsung atau berdasarkan jarak yang ingin diinterpolasi [10].

2.5 Matriks Konfusi (*Confusion Matrix*)

Matriks konfusi adalah metode yang digunakan untuk menghitung akurasi dalam konsep *data mining*. Untuk mengukur performa model klasifikasi SVM, digunakan matriks konfusi yang memberikan skor terhadap prediksi model berdasarkan data pengujian. Matriks konfusi menampilkan performa model dalam bentuk tabel. Setiap baris pada matriks menggambarkan klasifikasi aktual dari data, sedangkan setiap kolom menunjukkan klasifikasi prediksi dari data, atau sebaliknya. Presisi (*precision*) adalah proporsi kasus yang diprediksi sebagai positif dan benar-benar positif dalam data sebenarnya. Sementara itu, *recall* (*sensitivity*) adalah proporsi kasus positif yang sebenarnya dan benar-benar diprediksi sebagai positif [11].

Tabel 2. *Confusion Matrix*

<i>Actual</i>	<i>Classified Matrix</i>	
	+	-
+	<i>True Positives (A)</i>	<i>False Negatives (B)</i>
-	<i>False Positives (C)</i>	<i>Trie Negatives (D)</i>

Sumber: [11]

2.6 *Feature Importance*

Feature Importance adalah langkah dalam membangun model *machine learning* yang melibatkan perhitungan skor untuk semua fitur input dalam sebuah model guna menentukan seberapa penting setiap fitur dalam proses pengambilan keputusan. Semakin tinggi skor untuk sebuah fitur, semakin besar pengaruhnya terhadap model dalam memprediksi suatu variabel tertentu [12].

2.7 Pengolahan Data

Pengolahan data terdiri atas pembuatan dataset dengan memproses setiap parameter menjadi data raster yang kemudian dilakukan *extract multi values point* ke data titik kejadian longsor tahun 2015 – 2021 yang berjumlah 336 titik lalu tambahan titik tidak longsor dengan 214 titik. Dataset yang telah dibuat, kemudian dilakukan pembangunan model algoritma untuk klasifikasi dengan metode *machine learning* SVM dengan membagi dataset menjadi 70% pelatihan untuk pembuatan model dan 30% untuk pengujian atau evaluasi model. Model algoritma yang diperoleh selanjutnya digunakan untuk melakukan klasifikasi terhadap data baru. Tahap akhir adalah melakukan visualisasi sebuah peta dan divalidasi menggunakan 128 titik terjadinya longsor tahun 2022-2023. Hasil klasifikasi selanjutnya diolah menggunakan metode IDW untuk memperoleh strata kerawanan terhadap longsor. Klasifikasi potensi bencana tanah longsor ini dapat dilihat pada Tabel 3.

Tabel 3. Klasifikasi Bencana Tanah Longsor

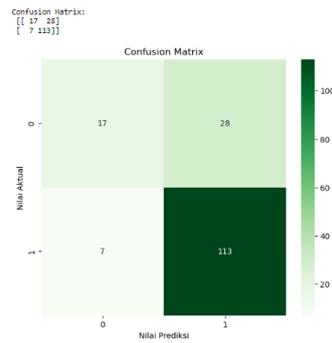
No	Klasifikasi Potensi Longsor	Indeks Nilai Potensi Longsor
1	Rendah	0
2	Menengah	0,50
3	Tinggi	1

(Sumber: [13])

3. HASIL DAN PEMBAHASAN

3.1 Evaluasi Model

Evaluasi model ini dilakukan menggunakan 165 titik longsor acak dari total 550 kejadian longsor. Dari evaluasi tersebut, diperoleh konfusi matriks yang kemudian digunakan untuk menghitung nilai akurasi, presisi, *recall*, dan nilai *F1-score* dari model algoritma SVM yang telah diuji. Gambar 4 menunjukkan matriks konfusi yang menjelaskan bahwa model memberikan prediksi sebagai berikut: 17 titik lokasi diprediksi tidak mengalami longsor dan memang tidak terjadi longsor, 7 titik lokasi diprediksi mengalami longsor tetapi tidak terjadi longsor, 28 titik diprediksi tidak mengalami longsor tetapi ternyata terjadi longsor, dan 113 titik diprediksi mengalami longsor dan memang terjadi longsor.



Gambar 4. Confusion Matrix

Hasil akurasi dari evaluasi model memiliki nilai 79% atau 0,7878 yang artinya model mempunyai kinerja yang baik karena semakin tinggi tingkat akurasi model, semakin baik pula kinerjanya dalam mengklasifikasikan data yang belum pernah dilihat sebelumnya.

```
Accuracy: 0.7878787878787878
Confusion Matrix:
[[ 17 28]
 [ 7 113]]
Classification Report:
      precision    recall  f1-score   support

     0       0.71      0.38      0.49         45
     1       0.80      0.94      0.87        120

 accuracy          0.79         165
 macro avg         0.75         165
 weighted avg         0.78         165
```

Gambar 5. Classification Report

Nilai akurasi pelatihan sebesar 0.79 menunjukkan bahwa model ini telah mampu memprediksi dengan benar sebesar 79% dari data pelatihan, baik wilayah longsor maupun tidak longsor. Namun, akurasi pelatihan yang tinggi tidak selalu berarti model akan berkinerja baik pada data yang tidak terlihat (data testing). *Precision* 0.75 menunjukkan bahwa dari seluruh wilayah yang diprediksi oleh model sebagai wilayah longsor, sebanyak 75% benar-benar merupakan wilayah longsor. Hal ini menunjukkan bahwa model cukup akurat dalam mengidentifikasi wilayah yang benar-benar mengalami longsor, namun model ini tidak dijamin dapat mengidentifikasi semua kasus longsor. *Recall* 0.66 menunjukkan bahwa dari seluruh wilayah yang sebenarnya mengalami longsor, model yang dibangun hanya berhasil mengidentifikasi 66% dari keseluruhan wilayah. Hal ini memiliki arti bahwa terdapat 34% wilayah longsor yang tidak terdeteksi oleh model (*false negatives*). Perlu diperhatikan pula bahwa nilai *recall* yang lebih rendah dibandingkan dengan *precision* menunjukkan bahwa model mungkin telah salah

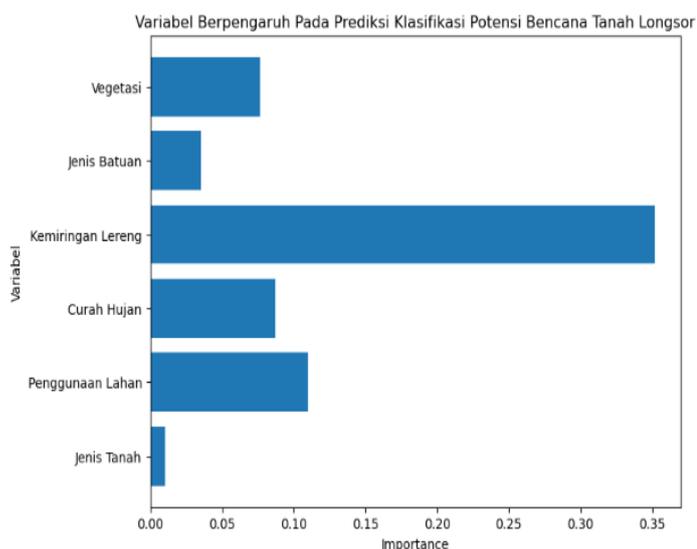
mengidentifikasi beberapa kasus positif atau longsor sebenarnya, yang berarti terdapat banyak titik kejadian longsor yang tidak terdeteksi oleh model.

F1 score yang mencakup *precision* dan *recall* memberikan gambaran tentang seberapa baik model dalam mendeteksi wilayah longsor. Nilai 0.68 menunjukkan bahwa ada keseimbangan yang cukup baik antara *precision* dan *recall*, meskipun masih ada kesalahan yang tentunya dapat diperbaiki. Hal ini berarti bahwa model yang telah dibangun memiliki performa yang wajar dalam mendeteksi wilayah longsor, namun mungkin perlu beberapa perbaikan untuk mengurangi kesalahan baik dalam *false positives* maupun *false negatives*.

3.2 Feature Importance

Penentuan variabel yang berpengaruh terhadap tanah longsor dilakukan menggunakan fungsi "*feature_importance*" dari library '*scikit-learn*'. *Feature importance* menunjukkan nilai signifikansi dari masing-masing variabel terhadap klasifikasi bencana tanah longsor. Variabel yang paling berpengaruh dalam pengklasifikasian bencana tanah longsor adalah kemiringan lereng, dengan nilai 0,35. Urutan variabel lain yang berpengaruh termasuk penggunaan lahan (0,1), curah hujan (0,08), kerapatan vegetasi (0,07), jenis batuan (0,03), dan jenis tanah (0,01).

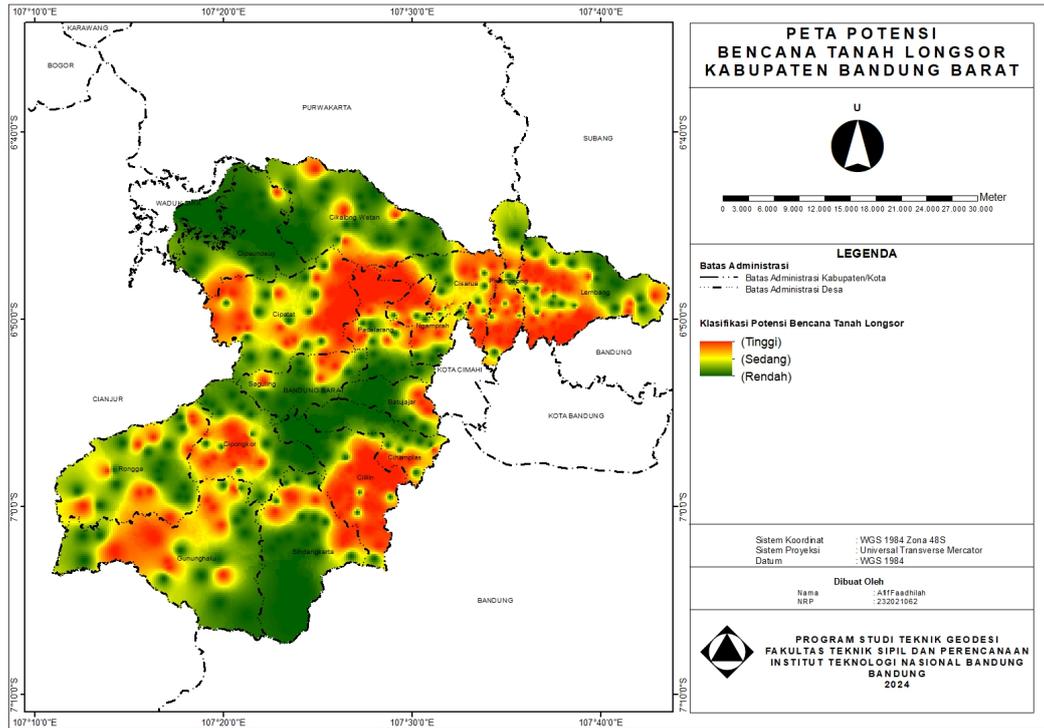
Kemiringan lereng menjadi variabel utama karena signifikannya pengaruh terhadap stabilitas fisik lereng. Lereng yang curam cenderung memiliki stabilitas yang lebih rendah karena gaya gravitasi yang lebih besar dapat memicu pergeseran massa tanah. Hal ini diperkuat oleh penelitian [4], yang menjelaskan bahwa lereng curam memiliki risiko lebih tinggi untuk mengalami tanah longsor, terutama jika terdapat aktivitas manusia seperti pertanian, pembebanan, pemotongan lereng, dan penambangan yang mempengaruhi morfologi alam. Akibatnya, risiko terjadinya longsor meningkat dan volume longSORan yang dihasilkan dapat lebih besar.



Gambar 6. Grafik *Feature Importance*

3.3 Visualisasi dan Luasan Klasifikasi Potensi Bencana Tanah Longsor

Visualisasi dan luasan hasil klasifikasi Potensi Bencana Tanah Longsor dengan menggunakan algoritma SVM dapat dilihat pada Gambar 7 berikut ini.



Gambar 7. Visualisasi Klasifikasi Potensi Bencana Tanah Longsor Kabupaten Bandung Barat

Pada Gambar 7 telah ditunjukkan hasil klasifikasi potensi bencana tanah longsor di Kabupaten Bandung Barat yang dilakukan menggunakan teknik Support Vector Machine. Klasifikasi potensi rendah memiliki nilai 0, mencakup luas 53.470 ha atau 42% dari total luas wilayah Kabupaten Bandung Barat. Selanjutnya, klasifikasi potensi menengah dengan nilai 0,50 mencakup luas 35.919 ha atau 28% dari total luas wilayah. Terakhir, klasifikasi tinggi dengan nilai 1 mencakup luas 38.020 ha atau 30% dari total luas wilayah kabupaten. Meskipun klasifikasi potensi longsor rendah mendominasi hasil prediksi pengolahan Machine Learning di Kabupaten Bandung Barat, potensi bencana tanah longsor di daerah tersebut tetap signifikan. Hal ini disebabkan oleh persentase luas klasifikasi potensi bencana longsor menengah dan tinggi yang cukup besar. Oleh karena itu, risiko terjadinya bencana tanah longsor di Kabupaten Bandung Barat memerlukan perhatian khusus dan tindakan preventif yang signifikan. Hasil perhitungan tercantum pada Tabel 4.

Tabel 4. Luasan Klasifikasi Potensi Bencana Tanah Longsor

No	Klasifikasi Potensi Bencana Tanah Longsor	Luasan (ha)	Persentase Luasan
1	Rendah (0)	53.470	42%
2	Menengah (0,50)	35.919	28%
3	Tinggi (1)	38.020	30%
Jumlah		127.410	100%

Tabel 5 menyajikan data luasan beberapa kecamatan yang terklasifikasi sebagai wilayah dengan potensi longsor rendah di Kabupaten Bandung Barat. Kecamatan Cipeundeuy menonjol sebagai wilayah dengan kategori potensi bencana tanah longsor rendah. Dari total luas wilayah kecamatan yang mencapai 8.553 ha, sebanyak 7.167 ha atau 84% masuk dalam kategori potensi rendah, menunjukkan bahwa sebagian besar wilayah Kecamatan Cipeundeuy memiliki risiko rendah terhadap bencana tanah longsor. Kondisi serupa terlihat di Kecamatan Batujajar, yang memiliki total luas wilayah 3.220 ha, dengan sekitar 2.162 ha atau 67% dari luas wilayah tergolong dalam kategori risiko

rendah terhadap tanah longsor. Kecamatan Cikalong Wetan dan Sindangkerta juga menunjukkan dominasi pada kategori potensi longsor rendah, dengan persentase luas wilayah yang tergolong dalam kategori risiko rendah masing-masing sebesar 67% dan 66%. Meskipun beberapa kecamatan di Kabupaten Bandung Barat, seperti Cipeundeuy, Batujajar, Cikalong Wetan, dan Sindangkerta, memiliki dominasi wilayah dengan potensi bencana tanah longsor yang rendah, tetap penting untuk selalu waspada dan melakukan upaya mitigasi. Meskipun risiko bencana tanah longsor rendah, kondisi alam yang dinamis dan perubahan cuaca ekstrem dapat meningkatkan potensi bencana. Langkah-langkah preventif seperti pemantauan kondisi tanah, pengendalian erosi, serta pendidikan dan kesadaran masyarakat terhadap bencana harus terus dilakukan untuk memastikan keselamatan dan ketahanan wilayah terhadap kemungkinan bencana di masa mendatang.

Tabel 5. Luasan Dominasi Beberapa Kecamatan Pada Klasifikasi Potensi Longsor Rendah

No	Kecamatan	Total Luasan (ha)	Luasan Potensi Rendah	Persentase
1	Cipeundeuy	8.553	7.167	84%
2	Batujajar	3.220	2.162	67%
3	Cikalong Wetan	11.509	7.560	66%
4	Sindangkerta	10.814	7.148	66%

Tabel 6 menyajikan data luasan beberapa kecamatan yang terklasifikasi sebagai potensi longsor menengah di Kabupaten Bandung Barat. Kecamatan Rongga menonjol sebagai wilayah dengan kategori potensi bencana tanah longsor menengah. Dari total luas wilayah kecamatan yang mencapai 11.641 ha, sebanyak 6.108 ha atau 52% masuk dalam kategori potensi menengah, menunjukkan bahwa sebagian besar wilayah Kecamatan Rongga memiliki risiko menengah terhadap bencana tanah longsor. Oleh karena itu, langkah-langkah preventif seperti perbaikan struktur tanah, pembangunan sistem drainase yang baik, serta peningkatan kesadaran dan kesiapsiagaan masyarakat sangat diperlukan untuk mengurangi risiko dan dampak bencana. Selain itu, pemantauan yang lebih ketat dan respons cepat terhadap perubahan lingkungan juga penting untuk meminimalkan potensi kerugian dan ancaman terhadap keselamatan penduduk.

Tabel 6. Luasan Dominasi Beberapa Kecamatan Pada Klasifikasi Potensi Longsor Menengah

No	Kecamatan	Total Luasan (ha)	Luasan Potensi Menengah	Persentase
1	Rongga	11.641	6.108	52%

Tabel 7 menyajikan data luasan beberapa kecamatan yang terklasifikasi sebagai potensi longsor tinggi di Kabupaten Bandung Barat. Kecamatan Cililin menonjol sebagai wilayah dengan kategori potensi bencana tanah longsor tinggi, di mana sekitar 4.727 ha atau 61% dari total luas wilayah kecamatan ini tergolong dalam kategori risiko tinggi. Demikian pula, Kecamatan Cisarua, Kecamatan Ngamprah, dan Kecamatan Cipatat juga menunjukkan dominasi dalam kategori potensi bencana tanah longsor tinggi, dengan persentase luas wilayah yang tergolong dalam kategori risiko tinggi masing-masing sekitar 52%, 52%, dan 49%.

Mengingat besarnya risiko ini, langkah-langkah mitigasi yang dapat diambil untuk mengurangi risiko dan dampak bencana tanah longsor mencakup peningkatan kerapatan vegetasi guna mencegah erosi dan meningkatkan stabilitas tanah. Penerapan sistem drainase yang efektif juga diperlukan untuk mengurangi akumulasi air yang dapat memicu longsor, sementara pembangunan tembok penahan tanah di daerah dengan kemiringan curam menjadi prioritas. Edukasi kepada masyarakat setempat tentang tanda-tanda awal longsor dan langkah-langkah evakuasi yang aman sangatlah penting untuk

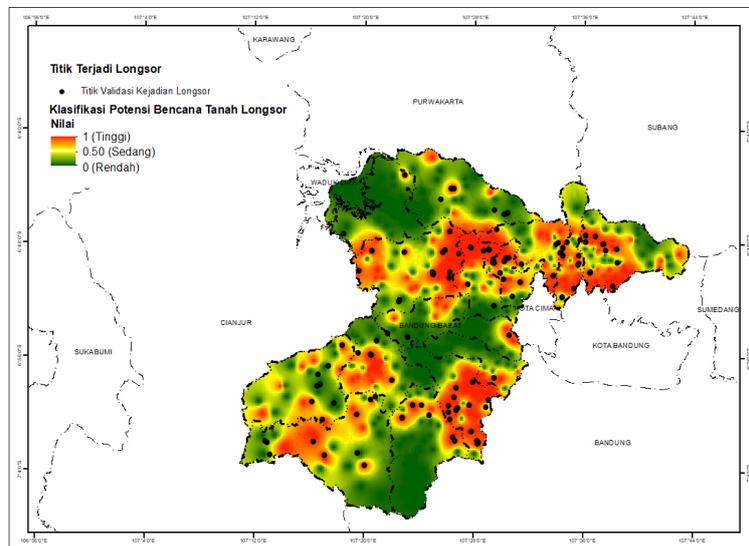
meningkatkan kewaspadaan dan kesiapsiagaan. Dengan upaya mitigasi yang tepat, risiko dan dampak bencana tanah longsor di kecamatan-kecamatan ini dapat diminimalkan secara signifikan.

Tabel 7. Luasan Dominasi Beberapa Kecamatan Pada Klasifikasi Potensi Longsor Tinggi

No	Kecamatan	Total Luasan (Ha)	Luasan Potensi Tinggi (Ha)	Persentase
1	Cililin	4.727	7.795	61%
2	Cisarua	2.786	5.407	52%
3	Ngamprah	1.662	3.193	52%
4	Cipatat	11.990	5.854	49%

3.4 Uji Validasi

Dari 120 titik terjadinya longsor tahun 2022 di Kabupaten Bandung divisualisasikan pada Gambar 8, terdapat 97 titik tervalidasi akurat dari hasil pemodelan yang dihasilkan.



Gambar 8. Visualisasi Sebaran Titik Validasi Potensi Bencana Tanah Longsor Kabupaten Bandung Barat

Perhitungan akurasi yang dilakukan mendapatkan nilai 75,78%, artinya antara hasil evaluasi pemodelan machine learning dengan pemodelan yang sudah dilakukan validasi memiliki informasi yang cukup akurat tentang klasifikasi potensi bencana tanah longsor di Kabupaten Bandung Barat. Untuk perhitungan akurasinya sebagai berikut:

$$\begin{aligned}
 \text{Akurasi Keseluruhan} &= \frac{\text{Jumlah titik validasi yang benar}}{\text{Total titik validasi}} \times 100\% \\
 &= \frac{97}{128} \times 100\% = 75,78\%
 \end{aligned}$$

Hasil perbandingan antara akurasi training (79%) dan nilai validasi (76%) menunjukkan hasil yang cukup berimbang yang menunjukkan bahwa model algoritma yang dibangun memiliki hasil yang tidak *overfitting*. Hal ini memiliki arti bahwa algoritma yang dibangun dapat melakukan klasifikasi dengan hasil yang konsisten, baik pada data training maupun data yang baru atau data yang belum pernah dilihat sebelumnya. Namun hasil klasifikasi dengan nilai akurasi dan validasi demikian perlu untuk diketahui penyebabnya. Terdapat beberapa hal yang dipandang boleh jadi menjadi penyebabnya.

1. Keterbatasan dalam menangani data yang tidak berimbang. Pada studi ini data longsor yang digunakan berjumlah 336 dan titik tidak longsor 214. Walaupun perbedaan jumlah data training ini tidaklah terlalu besar sehingga membuat data tidak longsor menjadi langka, namun adanya perbedaan keduanya membuat terjadinya ketidakseimbangan data (*imbalance data*). SVM mungkin sedikit terpengaruh dalam mengidentifikasi data minoritas, sebab SVM cenderung sensitif terhadap ketidakseimbangan data dibandingkan metode lain seperti *Random Forest*.
2. Optimalisasi *hyperparameter*. Pada penelitian ini *hyperparameter* menggunakan nilai-nilai *default* yang telah ditetapkan dari sumbernya, seperti kernel, nilai C, gamma, degree [14]. Disinyalir *hyperparameter default* yang dimiliki SVM tidak cukup dapat memberikan hasil optimal dalam penelitian ini.
3. Karakteristik Data. Berdasarkan beberapa penelitian diketahui bahwa beberapa tipe dataset memiliki pola yang lebih mudah ditangkap oleh *machine learning* dengan tipe pohon seperti *Random Forest* dibandingkan SVM [15].
4. Regularisasi dan *margin*. SVM berusaha untuk memaksimalkan margin antara kelas-kelas, yang dapat menyebabkan model menjadi terlalu kaku jika data tidak dapat dipisahkan dengan jelas. Regularisasi yang tidak tepat dapat membuat model menjadi terlalu ketat atau terlalu longgar. Data yang digunakan dalam penelitian ini boleh jadi tidak terlalu mudah untuk dipisahkan [16].
5. *Preprocessing*. Dalam tahap *preprocessing*, penelitian ini tidak melakukan normalisasi. Walaupun beberapa ahli menyatakan hal ini tidak terlalu berpengaruh, namun perbedaan nilai diantara fitur yang terlalu besar sedikit banyak akan menyebabkan perbedaan dalam proses klasifikasi [17].

4. KESIMPULAN

Hasil prediksi klasifikasi potensi longsor di Kabupaten Bandung Barat terbagi menjadi tiga kategori. Potensi bencana tanah longsor rendah ditandai dengan nilai 0, mencakup luas 53.470 Ha atau 42% dari total luas wilayah. Potensi menengah memiliki nilai 0,50 dengan luas 35.919 Ha atau 28%, sementara potensi tinggi memiliki nilai 1 dengan luas 38.020 Ha atau 30%. Kecamatan yang dominan dalam masing-masing kategori potensi longsor adalah Kecamatan Cipeundeuy untuk potensi rendah dengan persentase 84% dan luas area 7.167 hektar, Kecamatan Rongga untuk potensi menengah dengan persentase 52% dan luas 6.108 hektar, serta Kecamatan Cililin untuk potensi tinggi dengan persentase 61% dan luas area 7.795 hektar. Analisis variabel yang mempengaruhi terjadinya longsor menggunakan fungsi '*feature Importance*' menunjukkan bahwa variabel yang paling berpengaruh adalah kemiringan lereng dengan nilai 0,35. Variabel lain yang berpengaruh adalah penggunaan lahan (0,1), curah hujan (0,08), jenis kerapatan vegetasi (0,07), jenis batuan (0,03), dan jenis tanah (0,01), sesuai urutan signifikansinya terhadap klasifikasi bencana tanah longsor. Evaluasi model *machine learning* menggunakan teknik *Support Vector Machine* menunjukkan akurasi sebesar 78,78%. Hasil validasi terhadap 128 titik kejadian longsor menunjukkan akurasi sebesar 75,78%, menegaskan bahwa lebih dari 76% prediksi klasifikasi potensi bencana tanah longsor di Kabupaten Bandung Barat sesuai dengan data validasi yang tersedia. Ini mencerminkan ketepatan model dalam memetakan potensi bencana dengan akurasi yang signifikan. Model yang dibangun masih harus diperbaiki untuk dapat meningkatkan kemampuan identifikasi dan klasifikasi. Beberapa strategi yang dapat dilakukan adalah:

- i. **Penyesuaian *Threshold***: Mengubah *threshold* keputusan model untuk meningkatkan *recall*, meskipun hal ini mungkin akan dapat menurunkan *precision*.

- ii. **Peningkatan Data:** Menambah data pelatihan, khususnya data tentang wilayah longsor yang tidak terdeteksi, dapat membantu model belajar lebih baik.
- iii. **Teknik Resampling:** Menggunakan teknik oversampling pada data longsor atau undersampling pada data tidak longsor untuk menyeimbangkan dataset.
- iv. **Tuning Hyperparameter:** Melakukan tuning parameter model untuk menemukan konfigurasi yang optimal untuk mendeteksi wilayah longsor.

UCAPAN TERIMA KASIH

Penulis berterima kasih yang sebesar-besarnya kepada seluruh pihak instansi terkait yang telah membantu dalam penyusunan penelitian ini, yaitu BPBD Provinsi Jawa Barat, Stasiun Klimatologi Provinsi Jawa Barat, BAPPEDA Kabupaten Bandung Barat serta pihak yang tidak dapat penulis sebutkan satu persatu.

DAFTAR PUSTAKA

- [1] B. Mulia, “BNPB Catat 4.940 Bencana Terjadi Sepanjang 2023, Korban Jiwa 267 Orang,” *Detik.com*, Jakarta, Indonesia, hal. 1, 12 Januari 2024. [Daring]. Tersedia pada: <https://news.detik.com/berita/d-7137719/bnpb-catat-4-940-bencana-terjadi-sepanjang-2023-korban-jiwa-267-orang>
- [2] Himan Kamaludin, “Selama 5 Bulan, KBB Diterjang 133 Bencana Alam, 17 Orang Luka-Luka dan 2 Meninggal Dunia,” *Tribun.Jabar*, Bandung, Indonesia, hal. 1, 2024. [Daring]. Tersedia pada: <https://jabar.tribunnews.com/2023/05/10/selama-5-bulan-kbb-diterjang-133-bencana-alam-17-orang-luka-luka-dan-2-meninggal-dunia>
- [3] Y. Huang dan L. Zhao, “Review on landslide susceptibility mapping using support vector machines,” *Catena*, vol. 165, no. January, hal. 520–522, 2018, doi: 10.1016/j.catena.2018.03.003.
- [4] Y. Song *et al.*, “Evaluating Landslide Susceptibility Using Sampling Methodology and Multiple Machine Learning Models,” *ISPRS Int. J. Geo-Information*, vol. 12, no. 5, 2023, doi: 10.3390/ijgi12050197.
- [5] J. Big, R. C. Chen, C. Dewi, S. W. Huang, dan R. E. Caraka, “Selecting critical features for data classification based on machine learning methods,” *J. Big Data*, vol. 7, no. 52, 2020, doi: 10.1186/s40537-020-00327-4.
- [6] D. Kumar, M. Thakur, C. S. Dubey, dan D. P. Shukla, “Landslide susceptibility mapping & prediction using Support Vector Machine for Mandakini River Basin, Garhwal Himalaya, India,” *Geomorphology*, vol. 295, hal. 115–125, 2017, doi: 10.1016/j.geomorph.2017.06.013.
- [7] A. S. Nugroho, “Pengantar Support Vector Machine,” *J. Data Mining, Jakarta*, hal. 3, 2007.
- [8] P. B. Santoso, Yanto, A. Apriyono, dan R. Suryani, “Inverse distance weighting interpolated soil properties and their related landslide occurrences,” *MATEC Web Conf.*, vol. 195, 2018, doi: 10.1051/mateconf/201819503013.
- [9] P. S. Ananda, “Analisa Potensi Bencana Tanah Longsor menggunakan Interpolasi Inverse Distance Weighted (IDW),” *J. BATIRSI*, vol. 6, no. 1, hal. 6–9, 2022.
- [10] J. M. Pasaribu, N. Suryo, H. P. Pemanfaatan, dan P. Jauh, “Perbandingan Teknik Interpolasi Dem Srtm Dengan Metode Inverse Distance Weighted (Idw), Natural Neighbor Dan Spline (Comparison of Dem Srtm Interpolation Techniques Using Inverse Distance Weighted (Idw), Natural Neighbor and Spline Method),” *J. Penginderaan Jauh*, vol. 9, no. 2, hal. 126–139, 2012.
- [11] T. Rosandy, “Naive Bayes Vs C4.5 Ke Kelancaran Biaya Tetek Bengkek,” *2016*, vol. 2, no. 01,

- hal. 52–62, 2016.
- [12] D. Rengasamy *et al.*, “Feature importance in machine learning models: A fuzzy information fusion approach,” *Neurocomputing*, vol. 511, no. September, hal. 163–174, 2022, doi: 10.1016/j.neucom.2022.09.053.
 - [13] N. Nurwatik, M. H. Ummah, A. B. Cahyono, M. R. Darminto, dan J. H. Hong, “A Comparison Study of Landslide Susceptibility Spatial Modeling Using Machine Learning,” *ISPRS Int. J. Geo-Information*, vol. 11, no. 12, 2022, doi: 10.3390/ijgi11120602.
 - [14] T. Agrawal, *Hyperparameter Optimization in Machine Learning*. Bangalore, Karnataka, India: Apress, 2021. doi: 10.1007/978-1-4842-6579-6 ISBN-13.
 - [15] H. Shirmard, E. Farahbakhsh, dan D. Muller, “A review of machine learning in processing remote sensing data for mineral exploration,” 2021.
 - [16] A. Fernández, S. García, M. Galar, dan R. C. Prati, *Learning from Imbalanced Data Sets*. Springer Nature Switzerland, 2018. doi: 10.1007/978-3-319-98074-4.
 - [17] H. Nugroho, K. Wikantika, S. Bijaksana, dan A. Saepuloh, “Handling imbalanced data in supervised machine learning for lithological mapping using remote sensing and airborne geophysical data,” *Open Geosci.*, vol. 15, no. 1, 2023, doi: 10.1515/geo-2022-0487.