

Sistem Informasi Geografis Menggunakan Algoritma C4.5 untuk Prediksi Hasil Panen Padi di Kabupaten Simalungun

Fitri Annisa Purba¹, Ali Ikhwan^{1*}

¹Program Studi Sistem Informasi, Universitas Islam Negeri Sumatera Utara, Indonesia

Email: aliikhwan@uinsu.ac.id*

Info Artikel

Kata Kunci :

pertanian padi, sistem informasi geografis, algoritma C4.5

Keywords :

rice farming, geographic information system, C4.5 algorithm

Tanggal Artikel

Dikirim : 4 November 2024

Direvisi : 16 November 2024

Diterima : 17 November 2024

Abstrak

Kabupaten Simalungun merupakan salah satu daerah penghasil padi utama di Indonesia, yang kadang menghadapi fluktuasi hasil panen. Rendahnya hasil panen padi dapat menurunkan nilai jual dan keuntungan petani, sehingga diperlukan analisis data yang mampu memprediksi hasil panen secara akurat agar lebih efektif dan berkelanjutan. Penelitian ini mengembangkan sistem informasi geografis (SIG) berbasis web untuk memetakan hasil panen padi di Kabupaten Simalungun, menggunakan Algoritma C4.5 sebagai metode prediksi. Data yang diolah mencakup hasil panen padi tahun 2023-2024 dengan 1.000 sampel dan lima atribut utama, seperti luas lahan dan curah hujan, yang menghasilkan tingkat akurasi prediksi sebesar 85%. Sistem ini dirancang menggunakan *XAMPP*, *MySQL*, dan *Framework Laravel*, serta visualisasi peta dilakukan melalui *Leaflet.js*. Kontribusi utama penelitian ini adalah memberikan alat berbasis SIG yang tidak hanya mempermudah pemetaan hasil panen secara visual, tetapi juga meningkatkan efektivitas pengambilan keputusan Dinas Pertanian, seperti mengidentifikasi kecamatan dengan hasil panen rendah untuk prioritas bantuan dan membantu petani memantau perkembangan hasil panen secara spasial. Dengan prediksi yang lebih akurat, sistem ini berpotensi mendukung peningkatan efisiensi alokasi sumber daya dan kebijakan pertanian berbasis data.

Abstract

Simalungun Regency is one of Indonesia's major rice-producing regions, often facing fluctuations in harvest yields. Low rice yields can reduce sales value and farmers' profits, necessitating a data-driven analysis to accurately predict harvests for more effective and sustainable agricultural planning. This study developed a web-based Geographic Information System (GIS) to map rice harvest yields in Simalungun Regency, utilizing the C4.5 algorithm as a prediction method. The processed data included rice harvest results for 2023-2024, consisting of 1,000 samples and five key attributes, such as land size and rainfall, achieving a prediction accuracy rate of 85%. The system was built using XAMPP, MySQL, and the Laravel Framework, with harvest mapping visualized through Leaflet.js. The primary contribution of this study is providing a GIS-based tool that not only simplifies the visual mapping of harvest results but also enhances the effectiveness of decision-making by the Department of Agriculture. This includes identifying sub-districts with low harvest yields for prioritized aid and assisting farmers in monitoring harvest progress spatially. With more accurate predictions, this system has the potential to support efficient resource allocation and data-driven agricultural policies.

1. PENDAHULUAN

Teknologi informasi telah memberikan dampak signifikan dalam berbagai sektor, termasuk pertanian, dengan mendukung analisis data, perencanaan, dan evaluasi hasil produksi. Dalam konteks Kabupaten Simalungun, salah satu penghasil padi utama di Indonesia, tantangan utama adalah rendahnya kualitas hasil panen yang dapat menurunkan nilai jual dan keuntungan petani. Permasalahan ini menuntut adanya pendekatan berbasis data untuk memprediksi hasil panen secara lebih akurat dan berkelanjutan [1], [2].

Penelitian sebelumnya telah menggunakan Sistem Informasi Geografis (SIG) dan Algoritma C4.5 untuk berbagai aplikasi, termasuk pengelolaan data spasial dan prediksi. Namun, implementasi kombinasi SIG dengan Algoritma C4.5 dalam konteks prediksi hasil panen padi spesifik di Kabupaten Simalungun masih terbatas. Artikel ini bertujuan untuk mengisi gap tersebut dengan mengembangkan sistem berbasis SIG dan Algoritma C4.5 yang dapat membantu memprediksi hasil panen padi di setiap kecamatan, mempermudah pemetaan wilayah, dan mendukung pengambilan keputusan oleh dinas pertanian [3]–[6].

Adapun penelitian sebelumnya oleh Abdullah, M Zakki, Nawangnugraeni, Devi Astri, Hakim, Abdul Yuniarto, Prima (2021) dengan menerapkan Sistem Informasi Geografis berjudul “*Geographic Information System (GIS) For Mapping Greenpark Using Leaflet JS*”. Dengan menggunakan pemetaan *Leaflet js*, Sistem Informasi Geografis termasuk salah satu sistem yang berfungsi untuk mengolah data spasial sehingga menjadi sebuah informasi yang dapat di pahami oleh orang awam. *Leaflet js* merupakan salah satu teknologi dalam pengembangan Sistem Informasi Geografis berbasis online dan bersifat open *source*, sehingga dalam pengembangannya lebih dinamis dan fleksibel. Penelitian ini menerangkan mengenai pengolahan data spasial ke dalam Sistem Informasi Geografis yang berbasis *online*. Basis data yang digunakan dalam pengembangan ini meliputi Ruang Terbuka Hijau yang terbagi menjadi 2 klasifikasi, yaitu Publik dan Privat, dan terdapat 738 data total keseluruhan. Dengan banyaknya data tersebut memungkinkan untuk di olah secara informatif menggunakan Sistem Informasi Geografis yang berbasis *online* menggunakan teknologi *Leaflet js*. Hasil penelitian ini, menghasilkan persentase 19.61% data Ruang Terbuka Hijau Publik, hasil tersebut hampir memenuhi standar minimal Ruang Terbuka Hijau pada suatu wilayah. Sehingga dengan sistem informasi ini masyarakat umum dapat melihat data Ruang Terbuka Hijau dalam bentuk peta digital yang mudah dipahami karena disajikan dalam peta polygon.[7]

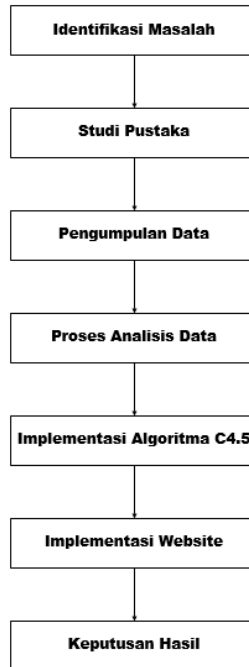
Penelitian sebelumnya oleh Nove Viktor Boyke Siahaan, Poningsih, Dedi Suhendro, Dedy Hartama & Suhada (2022) dengan menerapkan metode yang sama yaitu metode Algoritma C4.5 berjudul “Penerapan Data Mining Algoritma C4.5 Terhadap Prediksi Faktor Menurunnya Hasil Panen Padi”. Dengan menggunakan metode Algoritma C4.5, sumber data yang digunakan adalah data primer yang diperoleh dengan wawancara langsung kepada pemilik kilang padi dan petani untuk memprediksi faktor menurunnya hasil padi dengan melihat kondisi padi apakah terserang hama, apakah bulir pada padi banyak atau sedikit, cara tanam padi, musim tanam dan warna daun padi dengan begitu para petani mendapatkan solusi bagaimana menyelesaikan masalah menurunnya hasil panen padi dan mengurangi resiko kerugian [8].

Berdasarkan penelitian sebelumnya diatas, Hasil yang diperoleh dari pengujian Sistem informasi geografis menggunakan pemetaan *Leaflet js* dan metode algoritma C4.5 bahwa SIG merupakan sistem komputer yang digunakan untuk mengelola, menyimpan, dan menampilkan data geografis dan dengan mengimplentasi teknologi *Leaflet js* yang berbasis *online* dan bersifat *open source* maka dalam pengembangannya lebih dinamis dan fleksibel. dan untuk metode algoritma C4.5 dapat disimpulkan bahwa prediksi faktor menurunnya hasil panen padi didasari oleh keterhubungan Atribut Bulan Tanam, Hama, Bulir Padi, Warna Daun dan Cara Tanam. Namun dalam penelitian ini hanya menghitung nilai akurasi menggunakan aplikasi RapidMiner dari prediksi hasil panen padi tanpa melihat bagian mana atau daerah mana yang memiliki hasil panen terbanyak, sedang ataupun sedikit.

Dari permasalahan di atas, peneliti merancang sebuah Sistem Informasi Geografis (SIG) untuk prediksi hasil panen padi menggunakan algoritma C4.5. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk klasifikasi, segmentasi, atau pengelompokan data, dan bersifat prediktif. Hasil prediksi panen padi berdasarkan data yang ada akan divisualisasikan dalam bentuk pemetaan *WebGIS* untuk mempermudah pengguna membaca data. Variabel GIS, seperti *maps*, digunakan untuk menyimpan titik koordinat sehingga nama-nama kecamatan dapat dipetakan dengan akurat. Penelitian yang dilakukan diharapkan membantu Dinas Pertanian dalam memantau hasil panen padi di masa mendatang dan meningkatkan efisiensi pengambilan keputusan terkait hasil panen padi di setiap kecamatan.

2. METODE PENELITIAN

Adapun setiap proses dalam tahapan penelitian dilakukan seperti Gambar 1.



Gambar 1. Desain Penelitian

(1) Teknik Pengumpulan Data

Data merupakan sebuah kondisi dimana kondisi yang nyata terjadi dimasukkan kedalam sebuah elemen input yang selanjutnya akan diproses menjadi sebuah output dengan tujuan akhir sebagai informasi yang dibutuhkan[9]. Teknik pengumpulan data yang digunakan dalam penelitian ini merupakan tahapan yang ada dalam metode kualitatif. Dari pengumpulan data yang dilakukan ditemukan hasil data dengan 2 jenis data yaitu data primer dan data sekunder.

Data Primer merupakan data yang belum pernah diolah oleh pihak tertentu untuk kepentingan tertentu. Data Primer menunjukkan keaslian informasi yang terkandung didalam data tersebut[10]. Dalam penelitian ini sumber data primer dikumpulkan berdasarkan hasil wawancara dan observasi yang secara langsung dilakukan bersama Ibu Pengawas Benih Tanaman Ahli Muda Di Kantor Dinas Pertanian Kabupaten Simalungun.

Data Sekunder adalah data yang telah diolah, disimpan, disajikan dalam format atau bentuk tertentu oleh pihak tertentu untuk kepentingan tertentu. Data sekunder menunjukkan ketidakeaslian informasi yang terkandung di dalam data tersebut karena telah diolah untuk kepentingan tertentu, [11]. Sumber data sekunder dalam penelitian ini ialah artikel jurnal, buku, dan *website* serta penelitian sebelumnya.

Adapun teknik pengumpulan data yang dilakukan dalam penelitian ini antara lain:

1. Wawancara
Melakukan tanya jawab langsung dengan pihak – pihak yang berkaitan dengan penelitian. Peneliti mewawancarai langsung para mitra di Kantor Dinas Pertanian Kabupaten Simalungun.
2. Observasi
Melakukan pengamatan secara langsung di lingkungan kerja Kantor Dinas Pertanian Kabupaten Simalungun.
3. Studi Pustaka
Melakukan literasi dengan berbagai referensi untuk memperoleh informasi yang relevan dengan permasalahan yang diteliti. Seperti melalui jurnal, penelitian sebelumnya dan pengumpulan berbagai informasi tentang permasalahan yang diteliti.

(2) Analisis Data

Identifikasi masalah merupakan suatu metode untuk mengidentifikasi permasalahan yang timbul selama berjalannya sistem kerja untuk kemudian dicari solusi dari permasalahan tersebut [12].

(3) Proses Perhitungan Metode Algoritma C4.5

Pada proses perhitungan menggunakan metode algoritma C4.5, terdapat beberapa tahapan utama. Algoritma C4.5 merupakan metode pembelajaran mesin yang digunakan untuk membangun pohon keputusan (*decision tree*) dari data pelatihan. Pohon keputusan ini digunakan untuk klasifikasi dan prediksi berdasarkan atribut yang signifikan. Algoritma ini memiliki kemampuan menangani data kontinu, mengatasi nilai yang hilang (*missing values*), dan melakukan *pruning* untuk mengurangi kompleksitas pohon [13].

a. Input Data Set

Dataset awal terdiri dari data hasil panen padi, luas lahan, curah hujan, dan hasil panen tahun sebelumnya. Dataset ini diolah dengan memisahkan data menjadi atribut-atribut relevan yang akan digunakan dalam perhitungan.

b. Pembobotan Data

Setiap atribut diberi nilai berdasarkan kategorisasi tertentu:

- Luas Lahan: < 1 ha (kategori sempit) dan ≥ 1 ha (kategori luas).
- Curah Hujan: < 79 mm (rendah) dan ≥ 79 mm (tinggi).
- Hasil Panen Tahun Sebelumnya: Kategori naik atau turun berdasarkan ambang batas tertentu.

c. Perhitungan *Entropy*

Entropy digunakan untuk mengukur tingkat ketidakaturan atau ketidakpastian data. Rumus untuk menghitung *entropy* adalah:

$$Entropy(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

Ket:

s = himpunan kasus

n = jumlah partisi s

p_i = proporsi i terhadap s

d. Perhitungan *Gain*

Gain digunakan untuk menentukan atribut terbaik yang akan dijadikan sebagai *root* pada pohon keputusan. Rumus untuk menghitung *Gain* adalah:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (2)$$

Ket :

S = himpunan kasus

A = atribut

N = jumlah partisi atribut A

$|S_i|$ = jumlah kasus pada partisi ke- i

$|S|$ = jumlah kasus dalam S

e. Penetapan *Root*

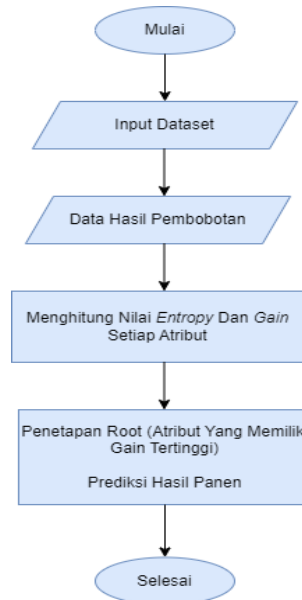
Setelah atribut dengan *Gain* tertinggi ditemukan, atribut tersebut dijadikan sebagai *root*, dan proses perhitungan diteruskan ke masing-masing cabang hingga semua data terklasifikasi.

f. Visualisasi dengan *Leaflet Js*

Setelah pohon keputusan terbentuk, hasil prediksi diolah menjadi peta spasial menggunakan *Leaflet.js*. Proses ini dimulai dengan pengolahan data spasial, di mana titik koordinat setiap kecamatan dimasukkan ke dalam sistem sebagai variabel GIS. Selanjutnya, data hasil prediksi dikaitkan dengan setiap kecamatan melalui atribut peta untuk memastikan hubungan antara data prediksi dan lokasi geografis. Terakhir, peta interaktif dibuat menggunakan *Leaflet.js*, di mana data tersebut divisualisasikan dengan label dan warna yang berbeda sesuai dengan kategori hasil prediksi, seperti hasil panen tinggi, sedang, atau rendah. Pendekatan ini memungkinkan pengguna

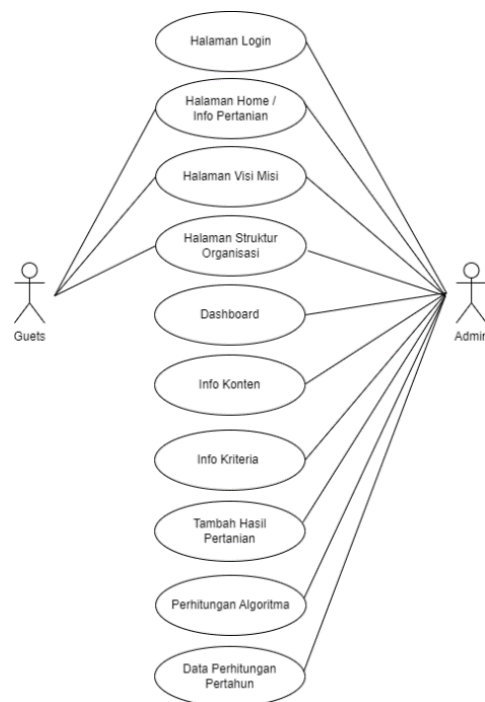
untuk memahami distribusi hasil panen secara lebih informatif dan mendukung pengambilan keputusan berbasis data.

Gambar 2 adalah langkah-langkah kerangka kerja metode Algoritma C4.5.

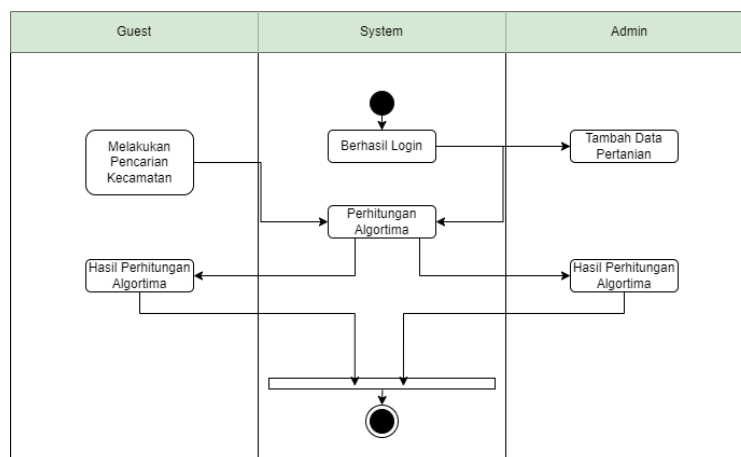


Gambar 2. Flowchart Algoritma C4.5

- (4) Usulan Sistem
Beberapa hal yang disarankan dalam sistem ini sebagai permasalahan yang akan ditawarkan jawaban atau penggantinya dalam upaya memperjelas persyaratan yang harus dipenuhi oleh sistem yang akan dibuat[14].
- (5) Implementasi
Fase ini menjelaskan tentang bagaimana membangun sebuah sistem SIG terhadap desain sistem yang dirancang dan menggunakan bahasa pemrograman berbasis web[15]. Dalam penerapannya, ada beberapa faktor yang perlu diperhatikan. Hal ini termasuk memastikan sistem beroperasi sebagaimana mestinya, mudah diakses oleh pengguna, dan mematuhi sistem yang dirancang.
- (6) Pengujian
Pengujian sistem (*system testing*) sebagai testing pada *software*. Pengujian sistem informasi dilakukan untuk memastikan sistem memenuhi persyaratan dan beroperasi sebagaimana mestinya [16]. *Black Box Testing* digunakan untuk melakukan pengujian. Pengujian Black Box berkonsentrasi pada persyaratan fungsional perangkat lunak. Penguji dapat melakukan pengujian terhadap spesifikasi fungsional dan menentukan serangkaian kondisi masukan.
- (7) Perancangan Desain Sistem
Desain sistem merupakan suatu proses yang digunakan untuk mendefinisikan arsitektur ataupun komponen antarmuka sehingga dapat menciptakan solusi yang efektif dan efisien. Desain sistem juga berisikan langkah-langkah operasi dalam pengolahan data yang mendukung operasi sistem. Diagram yang digunakan dalam desain sistem informasi ini yaitu *Use Case Diagram* (Gambar 3), *Activity Diagram* (Gambar 4)



Gambar 3. Use Case Diagram



Gambar 4. Activity Diagram

3. HASIL DAN PEMBAHASAN

Sistem informasi geografis adalah sistem komputer yang digunakan untuk mengelola, menganalisis, dan menyebarkan informasi geografis yang berkaitan dengan tata letak keruangan dan informasi-informasi yang berhubungan dengan permukaan bumi. Teknologi ini memungkinkan pengolahan data spasial menjadi informasi yang mudah dipahami, sehingga mendukung berbagai kebutuhan analisis, perencanaan, dan pengambilan keputusan [17]–[20]. SIG memiliki banyak manfaat salah satunya adalah memberikan kemudahan kepada para pengguna untuk menentukan kebijaksanaan yang akan diambil. Adapun kegunaan Sistem Informasi Geografis dalam penelitian ini yaitu untuk menentukan lokasi yang di ambil dari kecamatan yang ada di Kabupaten Simalungun dan dari data kecamatan tersebut dipetakan dan variabel yang digunakan dalam menyimpan data GIS yaitu menggunakan variabel dengan nama atribut maps yang mana tujuan variabel ini untuk menyimpan titik koordinat dari setiap kecamatan yang ada di kabupaten Simalungun.

Pada tahap ini merupakan hasil dari perhitungan dan semua proses yang dilakukan, dimana perhitungan menggunakan metode yang telah dibahas sebelumnya yaitu Algoritma C4.5 dengan pengumpulan data kemudian data hasil pembobotan lalu menghitung nilai *Entropy* dan *Gain* dan Penetapan *Root* (Atribut yang memiliki gain tertinggi).

a. Pengumpulan Data

Pada tabel 1 merupakan data asli yang diperoleh dari kantor Dinas Pertanian Kabupaten Simalungun setelah melakukan observasi dan pengumpulan data secara langsung ke kantor dinas.

Tabel 1. Data Produksi Padi Di Kabupaten Simalungun

No	Alamat Petani			Luas Lahan (Ha)	Hasil Panen (TON) 2023	Hasil Panen (TON) 2024	Curah Hujan (Mm)
	Nama Petani	Kecamatan	Desa				
1	Erikson Jaya Gultom	Huta Bayu Raja	Talang Bayu	1	5,85	6,28	86
2	Tiurma Simbolon	Dolok Panribuan	Dolok Tomuan	1	6,01	6,94	138
3	Walmar Haloho	Siantar	Karang Bangun	0,7	6,23	6,78	79
4	Ambrin	Pematang Bandar	Bandar Manis	1	6,01	6,72	138
5	Jenda Tatungan	Tanah Jawa	Tanjung Pasir	1	6,13	6,79	138
6	Cristian N Nahak	Sidamanik	Sarinatondang	0,7	6,02	6,78	138
7	Emron Manik	Bandar	Bandar Rakyat	0,5	6	6,78	138
8	Abdul Rahmat	Jawa Maraja Bahjambi	Bah Joga	1	5,8	6,99	117
9	Usman Simbolon	Panombean Pane	Saut Saroha	0,5	6,3	6,98	79
10	Heri Sucipto	Gunung Malela	Pematang Gajing	0,5	6,23	6,98	138
:							
1000	Yaser	Gunung Maligas	Karang Rejo	0,4	5,89	7,18	79

Pada tabel 1 diatas merupakan data asli yang diperoleh dari kantor Dinas Pertanian Kabupaten Simalungun setelah melakukan observasi dan pengumpulan data secara langsung ke kantor dinas.

b. Data Hasil Pembobotan

Algoritma C4.5 digunakan untuk mengetahui tingkat akurasi prediksi hasil panen pertanian tanaman padi di Kabupaten Simalungun. Langkah untuk proses algoritma C4.5 dengan membuat cabang setiap nilai akar tersaji pada Tabel 2.

Tabel 2. Variabel Atribut

Variabel Atribut	Nilai	Keterangan
Luas Lahan	<1 (Ha)	Sempit
	>=1 (Ha)	Luas
Hasil Panen 2023	<5,09 (Ton)	Turun
	>=5,09 (Ton)	Naik
Hasil Panen 2024	<6,09 (Ton)	Turun
	>=6,09 (Ton)	Naik

Curah Hujan	<79 (Mm) >=79 (Mm)	Rendah Tinggi
-------------	-----------------------	------------------

c. Menghitung Nilai *Entropy* dan *Gain*

• Nilai *Entropy*

Untuk menghitung nilai *Entropy* menggunakan rumus:

$$Entropy(S) = -\sum_{i=1}^n p_i \cdot \log^2 p_i \quad (1)$$

Keterangan:

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi i S terhadap S

$$\begin{aligned} - Entropy(Total) &= \left(\left(\frac{844}{1000} \right) * \log^2 \left(\frac{844}{1000} \right) + \left(\frac{156}{1000} \right) * \log^2 \left(\frac{156}{1000} \right) \right) \\ &= 0,62465382 \end{aligned}$$

- *Entropy* Luas Lahan

$$\begin{aligned} Luas &= \left(\left(\frac{379}{468} \right) * \log^2 \left(\frac{379}{468} \right) + \left(\frac{89}{468} \right) * \log^2 \left(\frac{89}{468} \right) \right) \\ &= 0,70182892 \end{aligned}$$

$$\begin{aligned} Sempit &= \left(\left(\frac{465}{532} \right) * \log^2 \left(\frac{465}{532} \right) + \left(\frac{67}{532} \right) * \log^2 \left(\frac{67}{532} \right) \right) \\ &= 0,54619712 \end{aligned}$$

- *Entropy* Hasil Panen 2023

$$\begin{aligned} Meningkatkan &= \left(\left(\frac{844}{844} \right) * \log^2 \left(\frac{844}{844} \right) + \left(\frac{0}{844} \right) * \log^2 \left(\frac{0}{844} \right) \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Menurun &= \left(\left(\frac{0}{156} \right) * \log^2 \left(\frac{0}{156} \right) + \left(\frac{156}{156} \right) * \log^2 \left(\frac{156}{156} \right) \right) \\ &= 0 \end{aligned}$$

- *Entropy* Hasil Panen 2024

$$\begin{aligned} Meningkatkan &= \left(\left(\frac{843}{957} \right) * \log^2 \left(\frac{843}{957} \right) + \left(\frac{114}{957} \right) * \log^2 \left(\frac{114}{957} \right) \right) \\ &= 0,52683255 \end{aligned}$$

$$\begin{aligned} Menurun &= \left(\left(\frac{1}{43} \right) * \log^2 \left(\frac{1}{43} \right) + \left(\frac{42}{43} \right) * \log^2 \left(\frac{42}{43} \right) \right) \\ &= 0,15935006 \end{aligned}$$

- *Entropy* Curah Hujan

$$\begin{aligned} Rendah &= \left(\left(\frac{278}{333} \right) * \log^2 \left(\frac{278}{333} \right) + \left(\frac{55}{333} \right) * \log^2 \left(\frac{55}{333} \right) \right) \\ &= 0,64652431 \end{aligned}$$

$$\begin{aligned} Tinggi &= \left(\left(\frac{566}{667} \right) * \log^2 \left(\frac{566}{667} \right) + \left(\frac{101}{667} \right) * \log^2 \left(\frac{101}{667} \right) \right) \\ &= 0,61339314 \end{aligned}$$

• Menghitung *Gain*

Untuk menghitung nilai *Gain* menggunakan rumus:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Dimana :

S = himpunan kasus

A = atribut

N = jumlah partisi atribut A

$|S_i|$ = jumlah kasus pada partisi ke- i

$|S|$ = jumlah kasus dalam S

- Nilai *Gain* Luas Lahan

$$= (0,62465382) - \left(\left(\frac{468}{1000} \right) * 0,701828917 \right) - \left(\left(\frac{532}{1000} \right) * 0,546197122 \right)$$

- = 0,005621022
- Nilai *Gain* Hasil Panen 2023
 - = $(0,62465382) - ((\frac{844}{1000}) * 0) - ((\frac{156}{1000}) * 0)$
 - = 0,624653823
- Nilai *Gain* Hasil Panen 2024
 - = $(0,62465382) - ((\frac{957}{1000}) * 0,526832546) - ((\frac{43}{1000}) * 0,159350063)$
 - = 0,113623024
- Nilai *Gain* Curah Hujan
 - = $(0,62465382) - ((\frac{333}{1000}) * 0,646524306) - ((\frac{667}{1000}) * 0,613393138)$
 - = 0,000228006

d. Penetapan *Root* (Atribut yang memiliki gain tertinggi)

Setelah mendapatkan data sampel, kemudian melakukan proses perhitungan jumlah data, *entropy* dan *gain*. Hasil tersebut terdapat pada Tabel 3.

Tabel 3. Penetapan *Root*

	Keterangan	Jumlah Kasus	Meningkat	Stabil	Entropy	Gain
Node						
1	total	1000	844	156	0,62465382	
	luas lahan					0,00562102
	Luas	468	379	89	0,70182892	
	Sempit	532	465	67	0,54619712	
	Hasil Panen 2023					0,62465382
	Meningkat	844	844	0	0	
	Menurun	156	0	156	0	
	Hasil Panen 2024					0,11362302
	Meningkat	957	843	114	0,52683255	
	Menurun	43	1	42	0,15935006	
	Curah Hujan					0,00022801
	Rendah	333	278	55	0,64652431	
	Tinggi	667	566	101	0,61339314	



Gambar 5. Rule Pohon Keputusan

Gambar 5 merupakan rule pohon keputusan untuk menentukan stabil dan meningkat, jika *Entropy* nya 0 maka dia stabil. Jika *Entropy* nya tidak =0 maka dia meningkat.

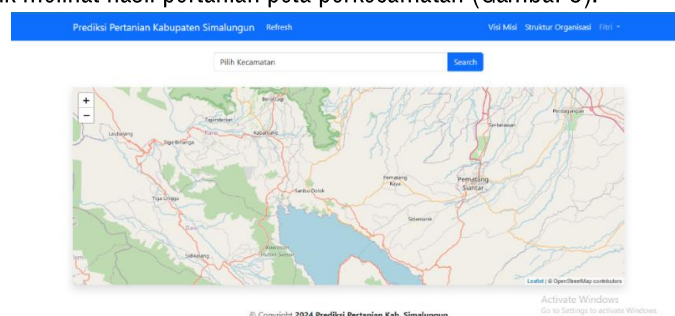
Pada hasil panen 2023, jika panen meningkat, hasil panen 2024 diprediksi stabil, sedangkan jika panen 2023 menurun, hasil panen 2024 tetap diperkirakan stabil. Pada hasil panen 2024, peningkatan hasil panen diprediksi akan terus meningkat, sedangkan penurunan hasil panen 2024 akan menghasilkan kondisi stabil. Berdasarkan atribut luas lahan, lahan yang luas cenderung menghasilkan panen meningkat, sedangkan lahan sempit cenderung stabil. Sementara itu, curah hujan tinggi mendukung peningkatan hasil panen, sedangkan curah hujan rendah masih dapat menghasilkan panen meningkat jika atribut lain mendukung. Prediksi ini dapat membantu Dinas Pertanian mengarahkan sumber daya, seperti memberikan bantuan teknologi atau pengelolaan air pada wilayah dengan keterbatasan tertentu. Penggunaan algoritma C4.5 memberikan kelebihan dalam menghasilkan pohon keputusan berbasis atribut penting yang mudah dipahami, dan integrasi dengan *Leaflet.js* memungkinkan visualisasi data dalam bentuk peta interaktif untuk identifikasi wilayah prioritas. Namun, keterbatasan algoritma ini, seperti ketergantungan pada kualitas data, dapat memengaruhi hasil prediksi, sementara *Leaflet.js* memerlukan optimasi untuk menangani dataset besar secara efisien. Secara keseluruhan, sistem ini mendukung pengambilan keputusan berbasis data dengan menyediakan prediksi yang informatif, meskipun validasi di lapangan tetap diperlukan untuk meningkatkan akurasi.

4. IMPLEMENTASI SISTEM

Berikut adalah contoh bagaimana sistem informasi geografis diintegrasikan kedalam web menggunakan algoritma C4.5. Adapun tampilan sistem yang sudah diimplementasikan adalah sebagai berikut, secara spesifik :

a. Tampilan User

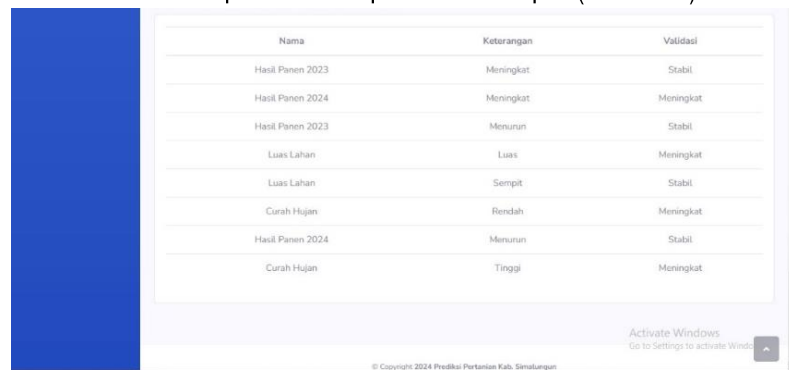
Tampilan *user* untuk melihat hasil pertanian peta perkecamatan (Gambar 5).



Gambar 5. Tampilan user untuk melihat hasil pertanian peta perkecamatan

b. Tampilan *Dashboard*

Tampilan *dashboard* untuk melihat prediksi hasil panen tahun depan (Gambar 6)



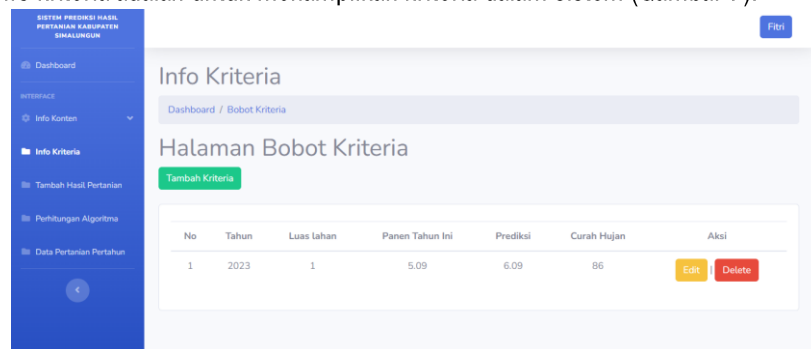
The screenshot shows a web application dashboard with a blue sidebar and a main content area. The main content area displays a table with three columns: Nama, Keterangan, and Validasi. The table contains eight rows of data related to harvest predictions for 2023 and 2024, including land area and rainfall. At the bottom right, there is a Windows watermark and a copyright notice for 2024.

Nama	Keterangan	Validasi
Hasil Panen 2023	Meningkat	Stabil
Hasil Panen 2024	Meningkat	Meningkat
Hasil Panen 2023	Menurun	Stabil
Luas Lahan	Luas	Meningkat
Luas Lahan	Sempit	Stabil
Curah Hujan	Rendah	Meningkat
Hasil Panen 2024	Menurun	Stabil
Curah Hujan	Tinggi	Meningkat

Gambar 6. Tampilan *dashboard* untuk melihat prediksi hasil panen

c. Tampilan Info Kriteria

Tujuan dari info kriteria adalah untuk menampilkan kriteria dalam sistem (Gambar 7).



The screenshot shows the 'Info Kriteria' page of the web application. It features a blue sidebar with navigation options. The main content area is titled 'Info Kriteria' and 'Halaman Bobot Kriteria'. It includes a 'Tambah Kriteria' button and a table with columns: No, Tahun, Luas Lahan, Panen Tahun Ini, Prediksi, Curah Hujan, and Aksi. The table contains one row of data for the year 2023. There are 'Edit' and 'Delete' buttons in the 'Aksi' column.

No	Tahun	Luas Lahan	Panen Tahun Ini	Prediksi	Curah Hujan	Aksi
1	2023	1	5.09	6.09	86	<button>Edit</button> <button>Delete</button>

Gambar 7. Tampilan Info Kriteria

d. Tampilan Hasil Perhitungan Algoritma C4.5

Form ini merupakan halaman pengolahan data prediksi hasil panen data (Gambar 8).



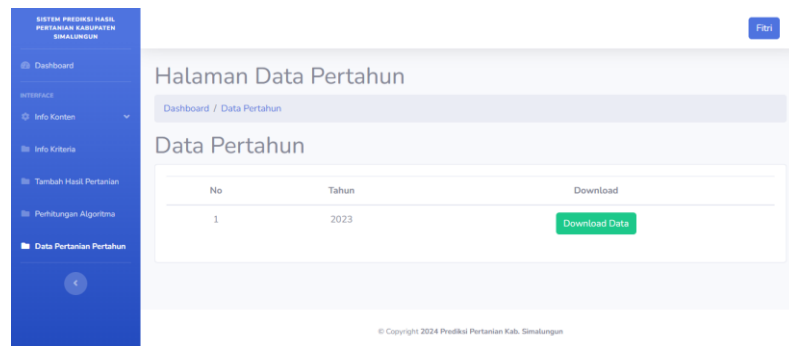
The screenshot shows the 'Hasil Perhitungan' page of the web application. It features a blue sidebar with navigation options. The main content area is titled 'Hasil Perhitungan' and displays a table with columns: Nama, Keterangan, Jumlah Kasus, Meningkat, Stabil, Entropy, and Gain. The table contains ten rows of data, including 'Total', 'Luas Lahan', and 'Hasil Panen' for 2023 and 2024, along with 'Curah Hujan' data.

Nama	Keterangan	Jumlah Kasus	Meningkat	Stabil	Entropy	Gain
Total		1000	844	156	0.62465382325071	0
Luas Lahan	Luas	468	379	89	0.70182891654312	0.0056210216704234
Luas Lahan	Sempit	532	465	67	0.5461971215002	0
Hasil Panen 2023	Meningkat	844	844	0	0	0.62465382325071
Hasil Panen 2023	Menurun	156	0	156	0	0
Hasil Panen 2024	Meningkat	957	843	114	0.52683254638889	0.11362302366106
Hasil Panen 2024	Menurun	43	1	42	0.15935006268563	0
Curah Hujan	Rendah	333	278	55	0.64652430562684	0.00022800623762886
Curah Hujan	Tinggi	667	566	101	0.61339313828987	0

Gambar 8. Tampilan Hasil Perhitungan Algoritma C4.5

e. Tampilan Data Pertahun

Form ini menampilkan data pertahun hasil panen padi di Kabupaten Simalungun (Gambar 9).



Gambar 9. Tampilan Data Pertahun

Pengujian (*Testing*)

Setelah menyelesaikan tahap implementasi aplikasi, langkah selanjutnya adalah melakukan tahap pengujian atau testing untuk memastikan bahwa fungsi aplikasi berjalan dengan baik. Pengujian ini dilakukan menggunakan metode *Blackbox Testing*, yang fokus pada pemeriksaan fungsionalitas aplikasi tanpa mempertimbangkan struktur internalnya. Dalam proses ini, dibuatlah kolom-kolom seperti indikator Variabel, Kegiatan *Testing*, dan Hasil Uji untuk mendokumentasikan berbagai aspek pengujian, mencatat indikator yang di uji, kegiatan yang dilakukan selama testing, dan hasil yang diperoleh dari setiap uji coba tersebut disajikan Tabel 4.

Tabel 4. *Blackbox Testing*

No	Indikator Variabel	Kegiatan <i>Testing</i>	Hasil Uji
1	Menjalankan <i>Website</i>	Tampilan halaman menu utama	Sesuai
2	Klik tombol <i>login</i>	Tampil halaman <i>dashboard</i>	Sesuai
3	Klik menu konten aktivitas	Tampil halaman konten aktivitas	Sesuai
4	Klik menu visi misi	Tampil halaman visi misi	Sesuai
5	Klik menu struktur organisasi	Tampil halaman struktur organisai	Sesuai
6	Klik menu info kriteria	Tampil halaman info kriteria	Sesuai
7	Klik menu tambah hasil pertanian	Tampil halaman tambah hasil pertanian	Sesuai
8	Klik menu perhitungan algoritma	Tampil halaman perhitungan algoritma	Sesuai
9	Klik data pertanian pertahun	Tampil halaman data pertahun lalu <i>download data</i>	Sesuai

5. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa pengembangan sistem informasi geografis (SIG) dengan implementasi algoritma C4.5 memberikan kontribusi signifikan dalam mendukung prediksi hasil panen padi di Kabupaten Simalungun. Sistem ini membantu Dinas Pertanian memantau hasil panen secara terstruktur, mengidentifikasi wilayah yang membutuhkan intervensi, dan menyediakan alat visual berupa peta interaktif yang mudah dipahami. Dengan 1.000 sampel data dan lima atribut utama, sistem ini mampu memberikan prediksi hasil panen untuk tahun mendatang, memprioritaskan petani dengan hasil panen rendah untuk menerima bantuan, serta memungkinkan petani memantau perkembangan hasil panen di setiap kecamatan. Kontribusi utama penelitian ini adalah penerapan SIG dan algoritma C4.5 yang terintegrasi untuk mendukung pengambilan keputusan berbasis data di sektor pertanian. Untuk meningkatkan akurasi dan efisiensi sistem, disarankan eksplorasi metode *machine learning* lain, seperti *Random Forest* atau *Gradient Boosting*, serta penambahan variabel masukan yang lebih kaya, seperti data cuaca terperinci atau informasi praktik pertanian, sehingga sistem dapat terus dikembangkan untuk mendukung kebijakan pertanian yang lebih efektif.

DAFTAR PUSTAKA

- [1] P. A. G. Taek, D. Supriadi, and S. M. Taek, "Upaya Pemberdayaan Petani Lahan Kering Untuk Mewujudkan Pertanian Berkelanjutan Dan Ketahanan Pangan," *JISIP (Jurnal Ilmu Sos. dan Pendidikan)*, vol. 6, no. 1, pp. 2345–2359, Jan. 2022, doi: 10.58258/jisip.v6i1.2829.
- [2] R. Rizal and R. Togimin, "Pendampingan Penyusunan E-RDKK Menggunakan Teknologi Google Form pada Gapoktan 'Rejo Mulyo' Desa Rejoagung, Kecamatan Sumberwringin Kabupaten Bondowoso," *Perigel J. Penyul. Masy. Indones.*, vol. 1, no. 3, pp. 443–448, Sep. 2023, doi: 10.56444/perigel.v2i3.878.
- [3] D. P. Lamondjong and M. Hardjianto, "Data Mining untuk Memprediksi Jumlah Penjualan Hasil Pertanian Menggunakan Algoritma Forecasting (Studi Kasus : Dinas Pertanian Kabupaten Banggai)," *J. Teknol. Inf.*, vol. 14, no. 2, pp. 94–102, 2021, doi: <https://doi.org/10.35842/jtir.v16i2.405.g351>.
- [4] W. Fuadi, R. Wandu, and M. W. Pohan, "Aplikasi Geografis Prediksi Hasil Padi Menggunakan Metode Double Moving Average di Kabupaten Aceh Utara," *J. Tek. Inform.*, vol. 13, no. 1, pp. 50–60, Apr. 2021, doi: 10.29103/techsi.v13i1.2831.
- [5] L. N. Rani, "Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit," *J. Invotek Polbeng*, vol. 1, no. 2, pp. 126–132, Nov. 2016, doi: 10.35314/isi.v1i2.131.
- [6] T. W. Utami, "Sistem Informasi Geografis Pemetaan Penyebaran Penyakit Berbasis Web Di Pekalongan," *J. Ilm. Sains Teknol. Dan Inf.*, vol. 1, no. 4, pp. 35–40, Jan. 2023, doi: 10.59024/jiti.v1i4.597.
- [7] M. Z. Abdillah, D. A. Nawangnugraeni, A. Hakim, and P. Yuniarto, "Geographic Information System (Gis) for Mapping Greenpark Using Leaflet Js," *J. Tek. Inform. Kaputama*, vol. 5, no. 2, pp. 259–266, 2021.
- [8] N. V. B. Siahaan, P. Poningsih, D. Suhendro, D. Hartama, and S. Suhada, "Penerapan Data Mining Algoritma C4.5 Terhadap Prediksi Faktor Menurunnya Hasil Panen Padi," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 7, no. 1, p. 27, 2022, doi: 10.30645/jurasik.v7i1.412.
- [9] S. Mulyanda, S. Defit, and Sumijan, "Analisis Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Harga Pasar Mobil Bekas," *J. KomtekInfo*, vol. 10, pp. 116–121, 2023, doi: 10.35134/komtekinfo.v10i3.427.
- [10] W. Kurniawati, A. S. Fitriani, S. Aji, and S. Suprianto, "Analisis Kehadiran Pemilih di Tempat Pemungutan Suara Pada Pemilihan Presiden dengan Menggunakan Algoritma C4.5 (Studi Kasus di Desa Wonokasian Sidoarjo)," *J. Electr. Eng.*, vol. 1, no. 1, p. 18, 2024, doi: 10.47134/jte.v1i1.2475.
- [11] M. Mazwar and Y. Kuleh, "Pengaruh E-WOM dan Kualitas Produk Serta Lokasi Terhadap Keputusan Pembelian," *Ekon. Manajemen, dan Akunt.*, vol. 25, no. 2, pp. 408–417, 2023.
- [12] T. Menggunakan and C. Algoritma, "Data Mining Prediksi Produksi Garam Pada Pertanian," *Oktal*, vol. 2, no. 2, pp. 460–471, 2023.
- [13] D. M. A. M. Sanjaya, "Penerapan Data Mining untuk Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Algoritma C4.5: Studi Kasus STMIK Primakara," *J. Ilm. Ilmu Terap. Univ. Jambi*, vol. 6, no. 1, pp. 84–97, 2022, doi: 10.22437/jiituj.v6i1.19600.
- [14] T. Walyadi, "Implementasi Algoritma C4.5 Untuk Menentukan Musim Tanam Cabai Merah Di Daerah Magelang," pp. i–59, 2019.
- [15] A. Faidlon *et al.*, "URANUS : Jurnal Ilmiah Teknik Elektro, Sains dan Informatika Filtering Konten Negatif Menggunakan Metode Algoritma C4.5 dan API Bing Penerapan di Desa Kelet," vol. 2, no. 1, 2024.
- [16] P. Potensi, M. Dan, and B. Siswa, "Implementasi algoritma c4.5 pada sistem pakar penentuan potensi minat dan bakat siswa smp muhammadiyah kota pangkalpinang," 2023.
- [17] Y. Noorollahi, A. Ghenaatpisheh Senani, A. Fadaei, M. Simaee, and R. Moltames, "A framework for GIS-based site selection and technical potential evaluation of PV solar farm using Fuzzy-Boolean logic and AHP multi-criteria decision-making approach," *Renew. Energy*, vol. 186, no. 1, pp. 89–104, Mar. 2022, doi: 10.1016/j.renene.2021.12.124.
- [18] Y. Gao, J. Zhao, and L. Han, "Exploring the spatial heterogeneity of urban heat island effect and its relationship to block morphology with the geographically weighted regression model," *Sustain. Cities Soc.*, vol. 76, no. 2, pp. 103431–103440, Jan. 2022, doi: 10.1016/j.scs.2021.103431.
- [19] A. M. Khan *et al.*, "MaxEnt Modelling and Impact of Climate Change on Habitat Suitability Variations of Economically Important Chilgoza Pine (*Pinus gerardiana* Wall.) in South Asia," *Forests*, vol. 13, no. 5, pp. 715–721, May 2022, doi: 10.3390/f13050715.
- [20] W. L. Hakim *et al.*, "Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea," *J. Environ. Manage.*, vol. 305, no. 4, pp. 114367–114374, Mar. 2022, doi: 10.1016/j.jenvman.2021.114367.