

Implementasi Data Mining untuk Prediksi Daerah Rawan Penyakit DBD Menggunakan Algoritma C4.5 (Studi Kasus: Kabupaten Sanggau)

Dayang Nur Kalsum^{*1}, Syarifah Putri Agustini Alkadri², Istikoma³

^{1,2,3}Teknik Informatika, Teknik Dan Ilmu Komputer, Universitas Muhammadiyah Pontianak

Email: ^{*}dnurkalsum@gmail.com, ²agustini.putri1108@gmail.com, ³istikoma@gmail.com

(Naskah masuk: 25 Agustus 2024, diterima untuk diterbitkan: 10 Januari 2025)

Abstrak: Penerapan data mining untuk memprediksi daerah rawan demam berdarah adalah hal yang tepat. Algoritma C4.5 atau dikenal juga dengan algoritma pohon keputusan merupakan salah satu teknik data mining yang dapat digunakan untuk membuat model prediktif berdasarkan data historis. Penelitian ini bertujuan untuk mengembangkan model prediksi daerah rawan penyakit DBD (Demam Berdarah Dengue) di Kabupaten Sanggau menggunakan algoritma C4.5 dengan metodologi penelitian seperti identifikasi masalah, pengumpulan data, analisis kebutuhan data, perancangan sistem, pengembangan sistem, pengujian sistem, analisis hasil pengujian sistem, penarikan kesimpulan. Penulis dapat membangun aplikasi website untuk membantu memprediksi daerah rawan penyakit DBD. aplikasi yang dibangun dapat membantu Dinas kesehatan dalam memprediksi penyakit DBD walaupun ada kekurangan dalam tingkat akurasi. Dalam konteks ini, data historis mengenai kasus DBD, faktor risiko dan karakteristik daerah di Kabupaten Sanggau dapat dimanfaatkan untuk membuat prediksi yang akurat terkait dengan daerah rawan penyakit DBD. Pembuatan aplikasi prediksi daerah rawan DBD dengan mengambil data di dinas Kabupaten sanggau dari tahun 2018-2023 diharapkan akan lebih banyak membantu dalam menyampaikan informasi khususnya untuk daerah-daerah yang mengalami penyakit demam berdarah di masa akan datang.

Kata Kunci – Data Mining; Algoritma C4.5; Aplikasi Website; DBD dan Confusion Matrix;

Implementation of Data Mining to Predict Dengue Prone Areas Using C4.5 Algorithm (Case Study: Sanggau Regency)

Abstract: Applying data mining to predict areas prone to dengue fever is the right thing. The C4.5 algorithm or also known as the decision tree algorithm is a data mining technique that can be used to create predictive models based on historical data. This research aims to develop a prediction model for dengue-prone areas in Sanggau Regency using the C4.5 algorithm with research methodology such as problem identification, data collection, data needs analysis, system design, system development, system testing, analysis of test results system, drawing conclusions. The author can build a website application to help predict areas prone to dengue fever. The application that was built can help the Health Service in predicting dengue fever even though there is a lack of accuracy. In this context, historical data regarding dengue cases, risk factors and regional characteristics in Sanggau Regency can be used to make accurate predictions regarding dengue-prone areas. It is hoped that the creation of a prediction application for dengue-prone areas by taking data from the Sanggau Regency office from 2018-2023 will be more helpful in providing information, especially for areas experiencing dengue fever in the future.

Keywords – Data Mining; C4.5 Algorithm; Website Application; DBD; Confusion Matrix;

1. PENDAHULUAN

Penyakit Demam Berdarah Dengue (DBD) merupakan salah satu masalah kesehatan masyarakat yang signifikan di Indonesia. DBD disebabkan oleh virus dengue yang ditularkan oleh nyamuk *Aedes aegypti* dan *Aedes albopictus*. Penyakit ini dapat menyebabkan gejala ringan hingga parah, bahkan bisa berujung pada kematian jika tidak ditangani dengan cepat dan tepat. Wilayah dengan kondisi lingkungan tertentu, seperti genangan air yang memungkinkan nyamuk

berkembang biak, menjadi daerah rawan penyakit DBD [1]. Dalam upaya pencegahan dan pengendalian penyakit DBD, langkah-langkah yang efektif diperlukan untuk mengidentifikasi daerah-daerah yang berpotensi menjadi tempat penyebaran penyakit ini. Data mining, cabang dari ilmu komputer yang fokus pada ekstraksi informasi yang bermanfaat dari data besar, telah terbukti berguna dalam mendukung pengambilan keputusan di berbagai bidang, termasuk kesehatan [2].

Dalam konteks ini, penerapan data mining untuk memprediksi area risiko demam berdarah adalah hal yang tepat. Algoritma C4.5 atau dikenal juga dengan algoritma pohon keputusan merupakan salah satu teknik data mining yang dapat digunakan untuk membuat model prediktif berdasarkan data historis [3]. Algoritma ini berfokus pada membagi kumpulan data menjadi subkumpulan variabel target yang semakin homogen, menciptakan pohon keputusan yang dapat diinterpretasikan untuk pengambilan keputusan. Melakukan prediksi menggunakan algoritma data mining seperti C4.5 juga berkontribusi pada pengembangan metodologi dalam bidang kesehatan masyarakat [4]. Ini membuka peluang untuk penelitian lebih lanjut mengenai penerapan data mining dalam konteks kesehatan dan potensi penggunaan algoritma lain untuk analisis yang lebih mendalam. Dengan mengintegrasikan prediksi dalam studi ini, tidak hanya membantu Dinas Kesehatan Kabupaten Sanggau dalam menghadapi tantangan penyakit DBD, tetapi juga memberikan kontribusi signifikan terhadap pengembangan metode analisis dan pengelolaan kesehatan berbasis data. Kabupaten Sanggau dipilih sebagai studi kasus karena memiliki karakteristik geografis dan lingkungan yang berpotensi menjadi tempat berkembang biaknya nyamuk *Aedes aegypti* dan *Aedes albopictus* [5].

Dinas Kesehatan Kabupaten Sanggau memiliki data historis terkait penyakit DBD, seperti jumlah kasus, lokasi, dan faktor-faktor lingkungan. Dengan memanfaatkan data ini, penelitian ini bertujuan untuk mengembangkan model prediksi yang dapat membantu identifikasi daerah-daerah rawan penyakit DBD di kabupaten tersebut. Salah satu metode yang digunakan dalam analisis prediktif adalah algoritma C4.5. Algoritma ini merupakan salah satu algoritma pohon keputusan yang populer dan efisien dalam mengekstraksi aturan-aturan keputusan dari data yang ada [6]. Dengan menggunakan algoritma C4.5, Dinas Kesehatan Kabupaten Sanggau dapat memanfaatkan data historis kasus DBD, seperti faktor risiko lingkungan, kondisi iklim, dan distribusi populasi nyamuk vektor, untuk membangun model prediksi. Penelitian ini bertujuan untuk mengembangkan model prediksi daerah rawan penyakit DBD di Kabupaten Sanggau menggunakan algoritma C4.5. Model ini akan membantu Dinas Kesehatan dalam mengidentifikasi lokasi-lokasi yang berisiko tinggi terhadap DBD, sehingga upaya pencegahan dan pengendalian dapat difokuskan dengan lebih tepat dan efisien [7].

2. METODE PENELITIAN

Metodologi penelitian ini akan mencakup beberapa tahapan yang perlu dilakukan untuk mencapai tujuan penelitian. Berikut adalah metodologi yang akan digunakan dalam penelitian ini:

Penjelasan dari *flowchart* metode penelitian diatas adalah sebagai berikut:

1. Identifikasi Masalah

Mengidentifikasi masalah adalah tahap awal yang dilakukan untuk menentukan apa saja latar belakang, rumusan masalah, batasan masalah dan tujuan yang akan di akan diterapkan pada sistem.

2. Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi mengenai penelitian yang akan dilakukan serta gambaran sistem yang akan dibangun. Tahap pengumpulan data dilakukan dengan cara studi literatur yang mencakup artikel, jurnal, skripsi, buku dan bahan bacaan terkait dengan judul penelitian, observasi.

3. Analisis Kebutuhan Sistem

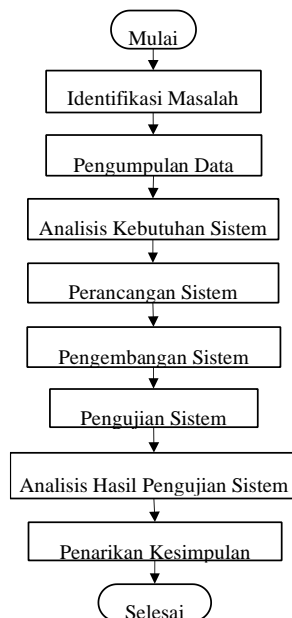
Analisis kebutuhan sistem untuk menentukan batasan yang akan diberikan pada sistem serta menentukan kebutuhan fungsional yang akan diberikan pada sistem.

4. Perancangan Sistem

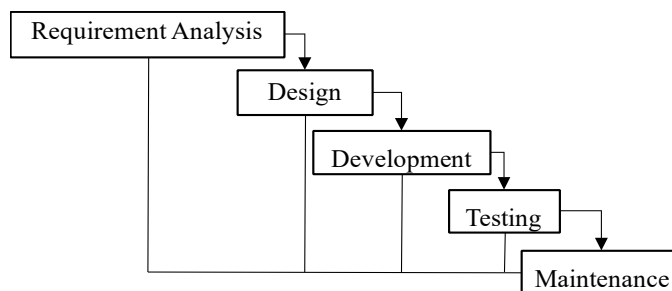
Perancangan sistem dilakukan untuk menggambarkan sistem yang akan dibangun. Perancangan sistem dimulai dari perancangan diagram arus data dengan flowchart dan DFD (Data Flow Diagram), perancangan basis data menggunakan ERD (Entity Relationship Diagram), antarmuka sistem, serta perancangan sistematis metode yang akan diterapkan.

5. Pengembangan Sistem

Pengembangan sistem merupakan proses pembuatan aplikasi menggunakan bahasa pemrograman berdasarkan perancangan yang sudah dilakukan. Adapun metode pengembangan sistem yang digunakan adalah metode waterfall. Metode ini melakukan pendekatan secara sistematis dan berurutan. Disebut waterfall karena tahap demi tahap yang harus dilalui menunggu selesainya tahap sebelumnya dan berjalan berurutan. Kelebihan metode waterfall adalah dokumen pengembangan sistem sangat terorganisir karena setiap fase harus terselesaikan dengan lengkap sebelum melangkah ke fase berikutnya sehingga dapat menghasilkan kualitas yang lebih baik. Hal ini berkaitan dengan pengembangan sistem yang diteliti, dimana tahap awal dilakukan analisa kebutuhan sistem, kemudian melakukan perancangan sistem, membangun sistem, melakukan ujicoba sistem dan melakukan pemeliharaan sistem. Untuk tahapan dari metode waterfall dapat dilihat pada gambar berikut:



Gambar 1. Metode Penelitian



Gambar 2. Metode Waterfall

2.1. Metode Algoritma C4.5

Algoritma C4.5 adalah algoritma pembelajaran mesin yang digunakan untuk membuat pohon keputusan dalam proses pengklasifikasian data. Algoritma ini pertama kali diperkenalkan oleh Ross Quinlan pada tahun 1993 sebagai perbaikan dari algoritma sebelumnya yang dikenal sebagai ID3 (Iterative Dichotomiser 3). C4.5 adalah salah satu algoritma pohon keputusan yang paling populer

dan banyak digunakan dalam bidang data mining dan pembelajaran mesin. Berikut adalah langkah-langkah utama yang diikuti oleh algoritma C4.5 untuk membangun pohon keputusan:

1. Pemilihan Atribut

Algoritma C4.5 memilih atribut yang paling baik untuk digunakan sebagai atribut pemisah pada setiap simpul dalam pohon keputusan. Kriteria yang digunakan untuk pemilihan ini adalah informasi gain atau rasio gain. Informasi gain mengukur seberapa baik atribut tersebut membagi data menjadi kelas-kelas yang berbeda, sedangkan rasio gain mengompensasi atribut yang memiliki banyak nilai atau kelas yang berbeda. Gain dan rasio gain adalah metrik yang sering digunakan dalam analisis data, terutama dalam konteks data mining dan machine learning. Berikut adalah cara umum untuk mendapatkan informasi gain dan rasio gain dari suatu data: Pilih fitur, Hitung entropy awal, Hitung entropy setelah splitting, Hitung informasi gain, Pilih fitur terbaik, dan iterasi.

2. Pembagian Data

Data pelatihan dibagi berdasarkan atribut yang dipilih pada langkah pertama. Masing-masing cabang dalam pohon keputusan menggambarkan nilai yang berbeda dari atribut tersebut seperti riwayat status DBD sebelumnya pada daerah yang ada di dataset. Pembagian data meliputi atribut pendukung dan atribut label yang dimana atribut pendukung menampilkan data seperti nama daerah, jumlah penduduk, dan lainnya. Sedangkan atribut label adalah atribut hasil akhir dari data atribut pendukung yaitu keterangan daerah tersebut rawan atau tidak rawannya DBD.

3. Rekursi

Proses ini diulangi untuk setiap cabang dari simpul saat ini. Algoritma C4.5 membangun pohon keputusan secara rekursif hingga salah satu dari tiga kondisi terpenuhi:

- Semua sampel dalam simpul memiliki label yang sama.
- Tidak ada atribut lagi yang dapat digunakan untuk memisahkan data.
- Mencapai batasan kedalaman maksimum pohon.

4. Pruning (Pemangkasan)

Setelah pohon keputusan dibangun, algoritma C4.5 dapat melakukan pemangkasan untuk menghindari overfitting. Pemangkasan menghapus cabang-cabang yang tidak memberikan kontribusi signifikan pada klasifikasi

5. Labeling

Setiap simpul daun dalam pohon keputusan diberi label kelas yang paling umum dari data pelatihan yang ada di simpul tersebut.

6. Pohon Keputusan Selesai

Proses berakhir ketika pohon keputusan lengkap dan siap digunakan untuk mengklasifikasikan data yang belum pernah dilihat sebelumnya.

2.2. Confusion Matrix

Confusion matrix adalah alat evaluasi yang digunakan dalam pemodelan klasifikasi untuk mengevaluasi kinerja model dengan membandingkan prediksi model dengan nilai sebenarnya dari data uji. Confusion matrix umumnya digunakan untuk evaluasi model klasifikasi yang memiliki dua atau lebih kelas.

Untuk studi kasus implementasi data mining untuk prediksi daerah rawan penyakit DBD menggunakan algoritma C4.5 di Kabupaten Sanggau, Anda dapat menghitung confusion matrix sebagai berikut:

Misalkan anda memiliki empat kelas dalam prediksi DBD:

1. True Positive (TP): Jumlah observasi yang diprediksi sebagai kelas rawan DBD dan benar-benar merupakan daerah rawan DBD.
2. False Positive (FP): Jumlah observasi yang diprediksi sebagai kelas rawan DBD tetapi sebenarnya bukan daerah rawan DBD.
3. True Negative (TN): Jumlah observasi yang diprediksi sebagai kelas tidak rawan DBD dan benar-benar bukan daerah rawan DBD.

4. False Negative (FN): Jumlah observasi yang diprediksi sebagai kelas tidak rawan DBD tetapi sebenarnya merupakan daerah rawan DBD.

Dalam notasi confusion matrix untuk kasus ini akan menjadi:

Tabel 1. Confusion Matrix

Keterangan	Kelas Rawan
Prediksi Rawan DBD	TP
Prediksi Tidak Rawan DBD	FP

2.3. Perancangan Pengujian

Pengujian ini dilakukan pada bagian pengguna. Pada tahap ini dilakukan pengujian sistem untuk memastikan apakah sistem yang dibangun dapat berfungsi dengan baik sesuai yang diharapkan. Pengujian blackbox dilakukan dengan menjalankan atau mengeksekusi data dengan proses yang dijalankan. Untuk rancangan pengujian blackbox bagian admin dapat dilihat pada Tabel dibawah:

Tabel 2. Perancangan Pengujian

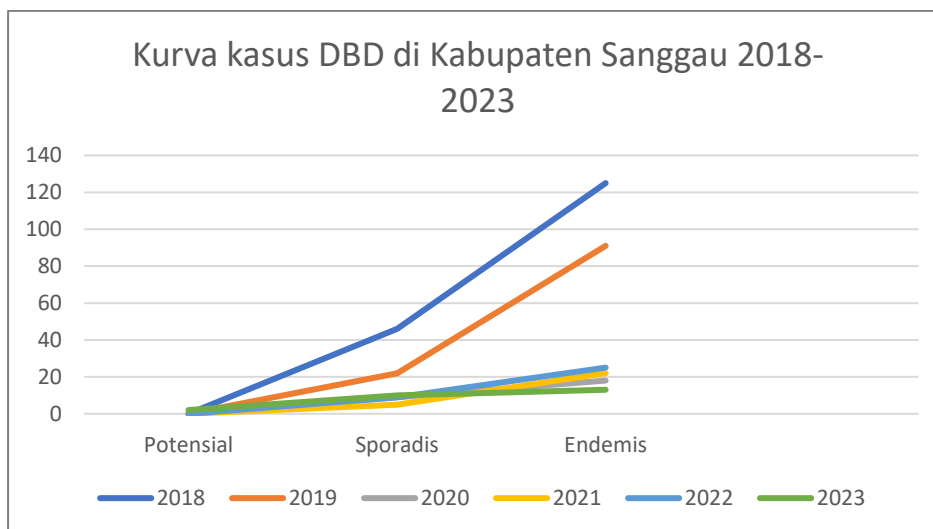
No.	Proses yang diuji	Hasil yang diharapkan	Berhasil/Tidak
1	Login	Menampilkan halaman login admin	
2	Beranda	Menampilkan halaman beranda	
3	Data Kasus	Menampilkan halaman data kasus dan dapat mengimport data excel	
4	Prediksi C4.5	Menampilkan halaman prediksi C4.5 dan pada halaman ini dapat menerima inputan dari data kasus atau daerah rawan yang kita ingin prediksi	
5	Hasil Prediksi	Menampilkan halaman hasil prediksi dan dapat menampilkan hasil seperti Potensial, Sporadis, atau pun Endemis pada data yang telah kita input	

2.4. Proses Perhitungan Manual Metode C4.5

Pada rancangan proses akan dilakukan bagaimana menghitung data penderita demam berdarah sesuai dengan algoritma yang digunakan oleh peneliti yaitu dengan algoritma C4.5 menggunakan dataset training penduduk seperti gambar dibawah ini:

STATUS ENDEMISITAS KECAMATAN BERDASARKAN KASUS DEMAM BERDARAH DENGUE														
Data Sampai dengan 30 Agustus 2023														
No	Kecamatan	Jumlah Kasus												Stratifikasi Daerah
		2018		2019		2020		2021		2022		2023		
		P	M	P	M	P	M	P	M	P	M	P	M	
1	TOBA	0	0	0	0	0	0	0	0	0	0	2	0	Potensial
2	MELIAU	9	1	3	0	3	0	0	0	0	0	0	0	Sporadis
3	KAPUAS	99	1	67	1	15	0	18	0	22	0	8	0	Endemis
4	MUKOK	11	0	7	0	0	0	2	0	1	0	1	0	Sporadis
5	JANGKANG	4	0	2	0	0	0	0	0	2	0	0	0	Sporadis
6	BONTI	2	0	2	0	0	0	1	0	0	0	0	0	Sporadis
7	PARINDU	25	0	23	0	3	0	4	0	3	0	5	0	Endemis
8	TAYAN HILIR	2	0	0	0	3	0	0	0	3	0	2	0	Sporadis
9	BALAI	1	0	1	0	0	0	0	0	0	0	3	0	Potensial
10	TAYAN HULU	10	1	5	0	0	0	0	0	0	0	1	0	Potensial
11	KEMBAYAN	1	0	0	0	2	0	0	0	2	0	0	0	Sporadis
12	BEDUAI	4	0	0	0	0	0	1	0	0	0	1	0	Sporadis
13	NOYAN	0	0	0	0	0	0	0	0	0	0	0	0	Potensial
14	SEKAYAM	0	1	1	0	1	0	1	0	0	0	0	0	Sporadis
15	ENTIKONG	3	0	2	0	0	0	0	0	1	0	0	0	Sporadis
	KABUPATEN	171	4	113	1	27	-	27	-	34	-	23	-	Endemis

Gambar 3. Data mentah dari Dinas Kesehatan Kabupaten Sanggau



Gambar 4. Kurva Kasus DBD di Kabupaten Sanggau 2018-2023

Catatan:

- Potensial (Daerah aman)
- Sporadis (Daerah kurang aman)
- Endemsi (Daerah tidak aman)

Dari data mentah diatas dengan total kasus 399 kasus DBD dapat di preprocessing menjadi data sampling pada tabel dibawah ini:

Tabel 3. Data Sampling Daerah Penderita DBD

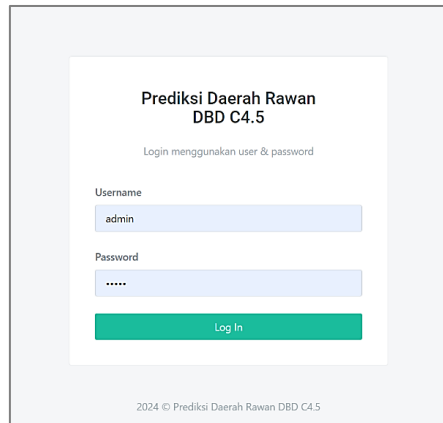
No.	Kecamatan	Jumlah Penderita	Jumlah Penduduk	Rata-rata Penduduk Per/Km	Status	Keterangan
1	Toba	2	16571	24	Potensial	Tidak rawan
2	Meliau	16	50773	134	Sporadis	Tidak rawan
3	Kapuas	231	86335	361	Endemis	Rawan
4	Mukok	22	21746	108	Sporadis	Tidak rawan
5	Jangkang	8	29572	39	Sporadis	Tidak rawan
6	Bonti	5	23800	37	Sporadis	Tidak rawan
7	Parindu	63	39154	248	Endemis	Rawan
8	Tayan Hilir	10	37826	1830	Sporadis	Tidak rawan
9	Balai	5	28937	253	Potensial	Tidak rawan
10	Tayan Hulu	17	39277	274	Potensial	Tidak rawan
11	Kembayan	5	32502	166	Sporadis	Tidak rawan
12	Beduai	6	13170	37	Sporadis	Tidak rawan
13	Noyan	4	11622	27	Potensial	Tidak rawan
15	Sekayam	4	40910	177	Sporadis	Tidak rawan
16	Entikong	6	18630	86	Sporadis	Tidak rawan

3. HASIL DAN PEMBAHASAN

Pada aplikasi yang dibangun dengan menggunakan HTML, CSS dan bahasa pemrograman PHP, Javascript dan MySql sebagai basis data. Implementasi sistem memiliki terdapat beberapa bagian diataranya sebagai berikut:

3.1. Halaman Tampilan Login Admin

Sebelum admin masuk pada aplikasi, admin harus mengisi form login pada halaman login admin dengan menggunakan Username dan Password yang telah disediakan.



Gambar 4. Halaman Tampilan Login Admin

3.2. Halaman Dashboard

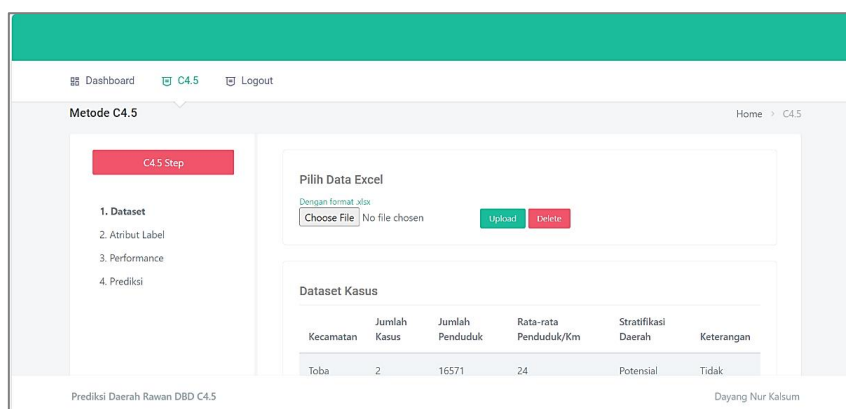
Perancangan tampilan halaman dashboard dapat dilihat pada gambar 5.4 Halaman ini merupakan tampilan utama yang berfungsi untuk melihat informasi menu seperti C4.5 dan logout.



Gambar 5. Halaman Dashboard

3.3. Halaman C4.5

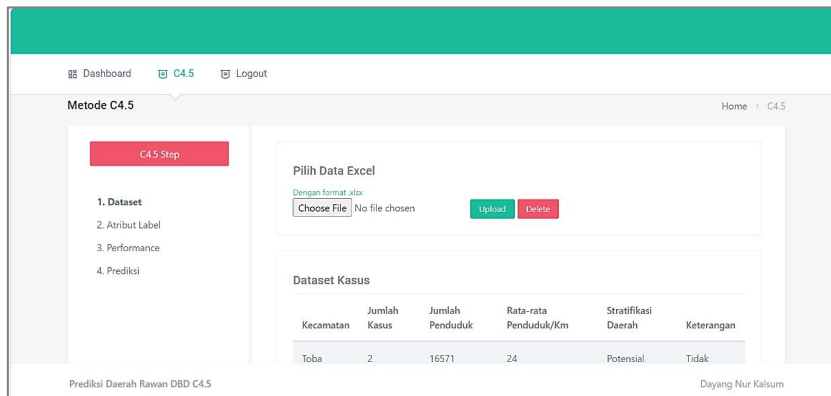
Pada halaman ini akan masuk padah tahapan pemrosesan C4.5 yang mencakup menu seperti input dataset, atribut label, performance, dan prediksi.



Gambar 6. Halaman C4.5

3.4. Dataset

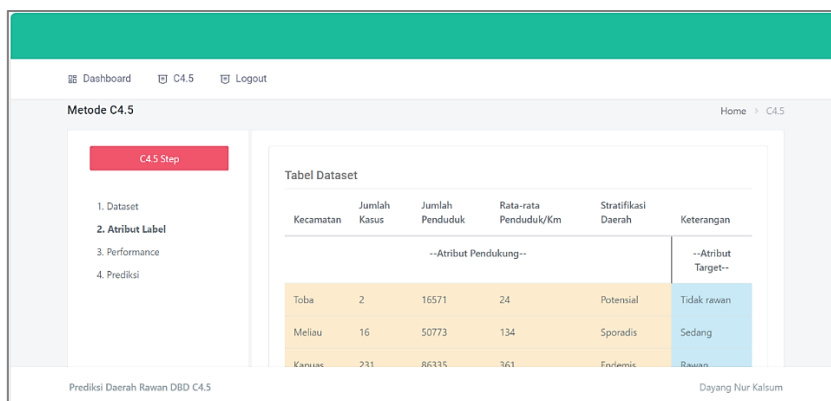
Pada menu Dataset ini menampilkan informasi untuk menginput dataset yang telah dibuat pada excel.



Gambar 7. Dataset

3.5. Menu Atribut Label

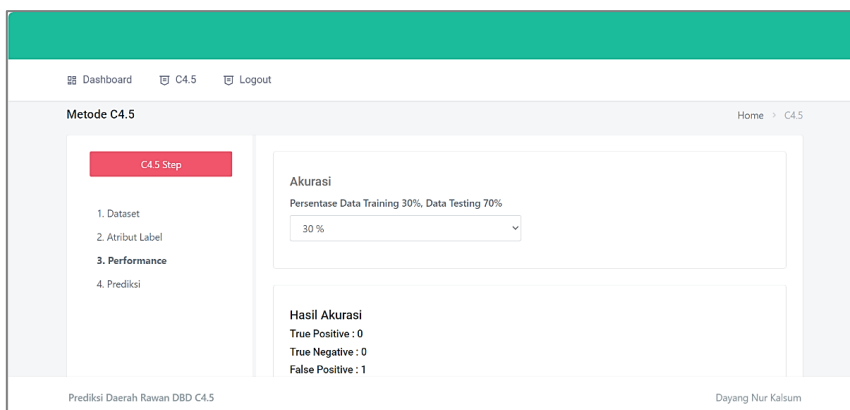
Pada menu Atribut Label terdapat informasi yang menampilkan 2 kateori seperti Atribut Pendukung dan Atribut Label pada data yang telah di input.



Gambar 8. Menu Atribut Label

3.6. Menu Performance

Pada menu Performance menampilkan informasi penginputan dan hasil dari akurasi yang diinginkan seperti True Positive, True Negative, FalsePositive, dan False Negative dengan cara menginput presentase data yang diinginkan.



Gambar 9. Menu Performance

3.7. Menu Prediksi

Pada menu Prediksi terdapat form Input Data Prediksi seperti Kecamatan, Jumlah Kasus, Jumlah Penduduk, Rata-rata Penduduk/Km, dan Stratifikasi Daerah.

Gambar 10. Menu Prediksi

3.8. Menu Prediksi Hasil

Setelah penginputan data berhasil akan ditampilkan informasi mengenai pohon keputusan dan hasil prediksi dari data yang kita input.

Gambar 11. Menu Prediksi Hasil

4. KESIMPULAN

Dari hasil penelitian diatas dapat diambil kesimpulan berdasarkan rumusan masalah sebagai berikut:

1. Kesimpulan dari penelitian ini adalah penulis dapat membangun aplikasi website untuk memprediksi daerah rawan penyakit DBD dan aplikasi yang dibangun dapat membantu dinas kesehatan dalam memprediksi penyakit DBD dengan baik walaupun ada kekurangan dalam tingkat akurasi.
2. Penerapan data mining dengan menggunakan algoritma C4.5 untuk memprediksi daerah rawan penyakit DBD di Kabupaten Sanggau merupakan langkah yang potensial dan relevan. Algoritma C4.5 adalah salah satu algoritma pohon keputusan yang dapat digunakan untuk menghasilkan model prediksi berdasarkan data historis. Dalam konteks ini, data historis mengenai kasus DBD, faktor risiko dan karakteristik daerah di Kabupaten Sanggau dapat dimanfaatkan untuk membuat prediksi yang akurat terkait dengan daerah rawan penyakit DBD.

3. Hasil dari 16 data diatas dapat di simpulkan ada 2 daerah rawan DBD yang berlokasi di Kecamatan Kapuas dan Kecamatan Entikong dan 14 tidak rawan dan dapat disimpulkan bahwa daerah yang penduduk per/km nya padat rawan terkena DBD.

UCAPAN TERIMA KASIH

Terimakasih kepada Allah swt, kedua orangtua, keluarga, para dosen pembimbing, dan teman-teman yang sudah memberikan support kepada saya untuk menyelesaikan skripsi dan jurnal.

DAFTAR PUSTAKA

- [1] T. M. Astuti, "Keragaman Karakter Fenotipik Nyamuk Aedes di Daerah Endemis Demam Berdarah di Kalimantan Barat," *Biocelbes*, vol. 5, no. 1, pp. 43-58, 2011. Alkrimi, Jameela Ali, et al. "Comparative Study Using Weka for Red Blood Cells Classification." *International Journal of Medical, Health, Pharmaceutical and Biomedical Engineering*, vol. 9, no. 1, 2015, pp. 19-22.
- [2] M. M. Sintorini, "Pengaruh Iklim terhadap Kasus Demam Berdarah Dengue," *Kesmas: National Public Health Journal*, vol. 2, no. 1, p. 11, 2007. Azahari, Azahari, et al. "Komparasi Data Mining Naive Bayes Dan Neural Network Memprediksi Masa Studi Mahasiswa S1." *Jurnal Teknologi Informasi Dan Ilmu Komputer*, vol. 7, no. 3, 2020, p. 443, doi:[10.25126/jtiik.2020732093](https://doi.org/10.25126/jtiik.2020732093).
- [3] Y. D. S. N. G. W. Fitriani, "Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4.5," *Jurnal Sistim Informasi dan Teknologi*, vol. 3, pp. 120-127, 2021.
- [4] W. A. Fatmawati K, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," vol. 2502, p. 714, 2018.
- [5] T. A. L. A. I. A. Kurniawan, "Implementasi Data Mining Untuk Prediksi Daerah Rawan Penyakit Demam Berdarah Menggunakan Algoritma C45," *Journal Satya*, pp. 50-58, 2018.
- [6] M. A. A. T. A. F. M. S. L. Sembiring, "Penerapan Metode Algoritma K-Means Clustering Untuk Pemetaan Penyebaran Penyakit Demam Berdarah Dengue (DBD)," *Jurnal Goresan Pena*, pp. 336-341, 2021.
- [7] M. Jailani, "Penerapan Algoritma C4.5 Terhadap Diagnosa Penyakit Demam Tifoid Berbasis Mobile," *UM Solo Jurnal TI*, vol. 11, no. 1, p. 27, 2015.
- [8] D. T. B. D. P. R. M. T. P. V. Meylitasari, "Perancangan Data Mining untuk Klasifikasi Prediksi Penyakit ISPA dengan Algoritma C4.5," 2017.
- [9] P. E. Amilia R, "Klasifikasi Diagnosa Penyakit Demam Berdarah Dengue Pada Anak Menggunakan Metode K-Nearest Neighbor Studi Kasus Rumah Sakit Pku Muhammadiyah Ujung," *INDEXIA*, vol. 2, no. 2, pp. 1-10, 2020.
- [10] R. S. J. Baker, "Data Mining," *Encyclopedia of Data Warehousing and Mining*, vol. 9, no. 1, p. 11, 2018.