Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 23386-23395
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Pengembangan Aplikasi Prediksi Kelulusan Akademik Siswa Berbasis Web Menggunakan Algoritma C4.5

Esa Andyarizky, Ahmad Nursodiq
Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang
esaandyarizky10@gmail.com*, dosen02526@unpam.ac.id

Abstrak

Pendidikan memiliki peran penting dalam pembinaan siswa di Yayasan Setia Hati Terate, dengan kelulusan sebagai indikator penguasaan kompetensi. Namun, proses evaluasi yang masih mengandalkan pencatatan dan pengamatan pelatih berpotensi menimbulkan subjektivitas, ketidakkonsistenan, serta keterlambatan pembinaan. Penelitian ini bertujuan mengembangkan aplikasi prediksi kelulusan siswa berbasis web menggunakan algoritma C4.5 sebagai alat bantu pengambilan keputusan yang objektif. Data penelitian berupa riwayat penilaian siswa berdasarkan kehadiran latihan, penguasaan materi, teknik senam dasar, kemampuan jurus, dan ketahanan fisik. Tahapan sistem mencakup pengunggahan data, pembersihan dan transformasi data, pembagian data latih dan data uji, pembentukan pohon keputusan, serta penyajian hasil prediksi. Aplikasi dibangun menggunakan PHP dan MySQL dengan hak akses Super Admin, Admin Pelatih, dan Siswa. Algoritma C4.5 memilih atribut berdasarkan nilai entropy dan information gain untuk menghasilkan aturan keputusan status lulus atau tidak lulus. Hasil pengolahan menunjukkan bahwa teknik senam memperoleh information gain tertinggi sebesar 0,3659 dan menjadi simpul akar. Pengujian confusion matrix pada 20 data uji menghasilkan akurasi 90%, sedangkan black box testing menunjukkan seluruh fitur berjalan sesuai kebutuhan. User Acceptance Test memperoleh tingkat kelayakan 94%. Dengan demikian, aplikasi mampu mendukung evaluasi kelulusan yang lebih cepat, konsisten, transparan, dan berbasis data serta membantu pelatih memberikan pembinaan yang lebih tepat kepada siswa yang berisiko tidak lulus.

Kata kunci: Algoritma C4.5, Aplikasi Web, Data Mining, Pohon Keputusan, Prediksi Kelulusan

1. Pendahuluan

Sebagai salah satu organisasi pencak silat terbesar di Indonesia, Yayasan Setia Hati Terate (PSHT) menerapkan sistem pelatihan berjenjang untuk melahirkan pesilat bermutu melalui ujian kenaikan tingkat. Status kelulusan ini menjadi indikator krusial penguasaan aspek teknis oleh Siswa, yang diukur berdasarkan lima parameter utama: kehadiran latihan, penguasaan materi, teknik senam dasar, kemampuan jurus, dan ketahanan fisik. Namun, mekanisme evaluasi saat ini masih didominasi oleh pencatatan manual dan observasi subjektif pelatih, sehingga rentan menimbulkan bias serta ketidakakuratan hasil. Kurangnya objektivitas tersebut berpotensi memicu ketidakpuasan dan memperbesar risiko kekeliruan fatal, seperti menggagalkan Siswa yang potensial atau meloloskan peserta yang belum kompeten. Oleh karena itu, implementasi pendekatan berbasis data yang sistematis sangat diperlukan untuk menjaga integritas institusi, meningkatkan kepercayaan Siswa, serta mengoptimalkan efisiensi waktu dan energi para pelatih selama proses penilaian.

Mekanisme evaluasi konvensional yang saat ini diterapkan oleh PSHT memiliki sejumlah kelemahan mendasar yang menjadikannya kurang efektif untuk skala organisasi yang besar. Tingginya ketergantungan pada ingatan serta pengamatan subjektif pelatih rentan memicu inkonsistensi penilaian, yang pada akhirnya dapat menimbulkan ketidakadilan bagi Siswa. Selain itu, volume data Siswa yang terus meningkat meliputi rekam kehadiran, nilai materi, teknik senam dasar, kemampuan jurus, serta ketahanan fisik sering kali hanya disimpan dalam dokumen fisik yang tidak terstruktur sehingga belum dimanfaatkan secara optimal. Akibatnya, pelatih mengalami kesulitan untuk mendeteksi Siswa yang berisiko tidak lulus maupun dalam memantau grafik perkembangan mereka secara efektif. Keterlambatan dalam memberikan intervensi dini ini berpotensi menurunkan motivasi Siswa atau bahkan menyebabkan mereka berhenti berlatih.

Pada era digital saat ini, metode evaluasi manual dinilai sangat tidak efisien, rentan terhadap kesalahan, dan tidak memadai untuk mengelola data yang kompleks. Kondisi ini secara langsung menghambat Yayasan Setia Hati Terate (PSHT) dalam merumuskan keputusan kelulusan yang proaktif, objektif, dan berbasis bukti. Fenomena inilah yang menjadi landasan utama penelitian ini, di mana prosedur penilaian pencak silat masih sangat didominasi oleh pendekatan yang subjektif. Dalam praktiknya, pelatih sering kali hanya mengandalkan intuisi serta catatan sederhana untuk menilai lima aspek fundamental, yaitu intensitas kehadiran latihan, pemahaman materi sejarah PSHT, kelancaran teknik senam gerak dasar, ketepatan jurus kombinasi, serta ketahanan fisik. Keterbatasan sistem tersebut memicu tingginya ketidakpastian dan memperbesar potensi bias, sehingga berisiko merugikan Siswa yang seharusnya memenuhi kriteria kelulusan maupun sebaliknya. Mekanisme evaluasi konvensional yang saat ini diterapkan oleh PSHT memiliki sejumlah kelemahan mendasar yang menjadikannya kurang efektif untuk skala organisasi yang besar Tingginya ketergantungan pada ingatan serta pengamatan subjektif pelatih rentan memicu inkonsistensi penilaian, yang pada akhirnya dapat menimbulkan ketidakadilan bagi Siswa. Pengimplementasian solusi tersebut ke dalam aplikasi berbasis Web diharapkan mampu mentransformasi pola evaluasi PSHT dari pendekatan reaktif menjadi lebih proaktif. Transformasi ini akan sangat mendukung perumusan keputusan yang objektif dan terukur, sekaligus mengoptimalkan alokasi sumber daya pelatihan. Untuk mewujudkan hal tersebut, penelitian ini difokuskan pada pengembangan sistem prediksi kelulusan dengan mengadopsi Algoritma C4.5. Metode klasifikasi ini dipilih secara khusus karena kapabilitasnya dalam memproduksi struktur *Decision Tree* yang sangat mudah diinterpretasikan. Keunggulan tersebut memfasilitasi pelatih untuk mengidentifikasi variabel penentu kelulusan secara langsung, melalui aturan logika yang spesifik (sebagai contoh: "Apabila skor jurus < 70% dan tingkat kehadiran < 80%, maka Siswa diprediksi berisiko tidak lulus"). Pembangunan perangkat lunak prediksi kelulusan ini tidak semata-mata ditujukan untuk menghasilkan arsitektur komputasi yang akurat, melainkan juga berorientasi pada penyediaan instrumen praktis yang mampu meningkatkan keadilan dan efisiensi prosedur kelulusan di lingkungan Yayasan Setia Hati Terate.

Tinjauan literatur singkat dari penelitian terkait menunjukkan bahwa algoritma klasifikasi dapat menjadi solusi atas permasalahan tersebut. Algoritma C4.5 terbukti sangat handal dalam menangani masalah klasifikasi akademik. Penelitian oleh Rahman et al. mengenai prediksi kelulusan mahasiswa menghasilkan akurasi sebesar 88,74%. Penelitian lain oleh Wicaksono et al. untuk memprediksi kelulusan siswa sekolah dasar menghasilkan tingkat akurasi mencapai 96,67%. Studi komparatif oleh Fakih et al. Juga membuktikan bahwa Algoritma Decision Tree C4.5 terbukti memiliki performa yang lebih unggul dibandingkan metode K-Nearest Neighbor (K-NN) yang hanya mencapai 71,79%. Ulfah & Jepara juga membuktikan Algoritma C4.5 memiliki performa yang lebih baik daripada Naïve Bayes Classifier. Keadaan canggih, analisis kesenjangan, dan kebaruan terlihat pada pemanfaatan algoritma C4.5 secara spesifik untuk penilaian pencak silat yang belum pernah terkomputerisasi.

Pertanyaan tujuan dari penelitian ini adalah: 1) Bagaimana merancang dan membangun sebuah aplikasi berbasis web yang mampu mengolah data kelulusan siswa? 2) Bagaimana mengimplementasikan algoritma C4.5 secara efektif untuk memprediksi kelulusan siswa? 3) Sejauh mana tingkat akurasi model prediksi yang dihasilkan oleh algoritma C4.5 dalam memprediksi kelulusan siswa? Pengimplementasian solusi ke dalam aplikasi berbasis Web diharapkan mampu mentransformasi pola evaluasi PSHT dari pendekatan reaktif menjadi lebih proaktif. Metode klasifikasi ini dipilih secara khusus karena kapabilitasnya dalam memproduksi struktur Decision Tree yang sangat mudah diinterpretasikan menjadi aturan logika (rule).

2. Metode Penelitian

Studi Penelitian ini mengadopsi pendekatan kuantitatif yang dipadukan dengan metode rekayasa perangkat lunak melalui studi kasus yang dilakukan langsung di Yayasan Setia Hati Terate. Data historis akademik siswa diekstraksi untuk dijadikan kumpulan data utama dalam pembangunan model klasifikasi. Prosedur sistem diimplementasikan melewati tahapan terstruktur yang dimulai dengan unggah data (upload data), dilanjutkan dengan tahapan pra-pemrosesan data (data cleansing dan data transform), dan diakhiri dengan pemrosesan menggunakan algoritma C4.5 untuk membentuk decision tree.

a. Jenis Data dan Sumber:

Observasi ini memanfaatkan data sekunder yang bersumber dari rekam jejak historis kelulusan Siswa pada Yayasan Setia Hati Terate (PSHT). Adapun variabel yang diekstraksi mencakup lima parameter penilaian, yaitu intensitas kehadiran, penguasaan materi, kelancaran teknik senam dasar, kemampuan jurus, serta ketahanan fisik.

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

b. Tahapan Implementasi:

1. Pra-pemrosesan Data: Himpunan data historis akan melalui tahap pra-pemrosesan yang secara spesifik meliputi proses *Data Cleansing* dan *Data Transform* agar struktur informasinya tervalidasi dan siap untuk dieksekusi.
2. Pemodelan: Komposisi Dataset akan didistribusikan ke dalam dua bagian utama, yakni 80% dialokasikan sebagai Data Latih dan 20% sisanya ditetapkan sebagai Data Uji. Pembentukan arsitektur prediktif ini dilakukan secara komputasional menggunakan Algoritma C4.5.
3. Pengembangan Sistem: Basis pengetahuan yang dihasilkan oleh model Algoritma C4.5 tersebut kemudian ditanamkan ke dalam perangkat lunak berbasis Web, yang ditujukan untuk beroperasi sebagai Sistem Pendukung Keputusan guna memfasilitasi pelatih.

c. Metode Analisis dan Pengujian:

1. Pengujian Model (Eksperimental): Validasi performa algoritma akan diukur secara presisi dengan mengimplementasikan instrumen *Confusion Matrix*. Dalam pengujian ini, kalkulasi metrik evaluasi difokuskan pada perolehan nilai Akurasi, Presisi, dan *Recall*.
2. Analisis Wawasan: Hierarki dari struktur *Decision Tree* yang dihasilkan akan diurai lebih lanjut untuk mengekstraksi kaidah aturan kondisional (*if-then*). Analisis ini berfungsi untuk mengidentifikasi variabel determinan (faktor paling dominan) yang mendasari penentuan status kelulusan.
3. Pengujian Sistem: Evaluasi teknis terhadap fungsionalitas aplikasi Web akan dieksekusi melalui pendekatan Black Box Testing untuk mengonfirmasi bahwa seluruh fitur berjalan sempurna dan selaras dengan spesifikasi perancangan.

2.1. Pra-pemrosesan Data

Tahap pra-pemrosesan berfokus pada pembersihan data dari atribut bernilai kosong serta transformasi data angka menjadi kategori. Variabel atau atribut masukan yang diekstraksi mencakup lima pilar utama, yakni intensitas kehadiran latihan, pemahaman materi, kelancaran senam dasar, kemampuan jurus, serta ketahanan fisik, beserta label status kelulusan akhir mereka. Sistem ini tidak didesain untuk memproyeksikan tingkatan sabuk maupun kualifikasi akademik lanjutan lainnya.

2.2. Algoritma C4.5

Algoritma C4.5 bekerja dengan memilih atribut berdasarkan perolehan nilai rasio perolehan informasi (Information Gain) tertinggi sebagai akar atau cabang pemisah. Langkah pertama sistem adalah mengukur tingkat heterogenitas kelas dari himpunan data secara keseluruhan melalui kalkulasi entropi menggunakan persamaan matematika seperti rumus 1.

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

S adalah himpunan kasus data, n adalah jumlah partisi kelas, dan pi adalah proporsi kasus pada kelas i terhadap total kasus [6]. Setelah entropi himpunan data diketahui, sistem mengukur tingkat signifikansi setiap atribut dalam memisahkan data dengan menghitung perolehan informasi seperti rumus 2.

$$Gain(S, A) = Entropy(S) - \sum_{v=1}^{n} \frac{|S_v|}{|S|} Entropy(S_v)$$

S adalah ruang sampel keseluruhan, A adalah atribut yang diuji, Sv adalah jumlah kasus pada partisi atribut, dan |S| adalah total keseluruhan kasus dalam ruang sampel. Setiap variabel dalam rumus dieksekusi secara berulang pada sistem.

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

2.3. Perancangan Sistem

Perancangan sistem dimodelkan melalui Unified Modelling Language (UML) yang mencakup Use Case, Activity, Sequence, dan Class Diagram. Terdapat tiga aktor utama yang berinteraksi: Super Admin yang mewarisi seluruh izin akses Admin Pelatih dengan tambahan akses absolut untuk manajemen pengguna dan pengelolaan keseluruhan konten aplikasi; Admin Pelatih yang memiliki wewenang mengelola data operasional pelatihan serta menjalankan algoritma prediksi; dan Siswa yang merupakan level otorisasi dasar dengan sifat view-only. Setelah berhasil masuk ke dalam sistem, siswa hanya diperkenankan melihat hasil prediksinya tanpa memiliki kemampuan untuk memicu proses prediksi atau memodifikasi data.

Perancangan basis data dilakukan melalui tahapan normalisasi hingga mencapai Bentuk Normal Ketiga (3NF) menghasilkan struktur basis data relasional. Struktur ini memastikan tidak ada duplikasi data dan inkonsistensi antara data pengguna, data latih, struktur pohon keputusan, dan riwayat hasil prediksi.

3. HASIL DAN DISKUSI

Hasil penelitian didasarkan pada urutan logis komputasi algoritma C4.5 untuk membentuk sebuah aturan keputusan. Integrasi algoritma ke dalam aplikasi berbasis web memungkinkan pemrosesan perhitungan yang rumit secara otomatis dan penyajian luaran prediksi secara seketika.

3.1. Kumpulan Data Latih

Kumpulan data yang diolah merupakan data historis siswa yang berjumlah 20 sampel kasus valid untuk proses komputasi iterasi. Data memuat nama siswa beserta evaluasi kelima atribut dan target keputusan Lulus atau Tidak Lulus. Selengkapnya disajikan pada Tabel 1.

Tabel 3.1 Data Latih Historis Siswa

No	Nama	Kehadiran	Materi	Senam	Jurus	Fisik	Kelulusan
1.	Andi	Rajin	Baik	Baik	Baik	Baik	Lulus
2.	Budi	Rajin	Cukup	Baik	Cukup	Baik	Lulus
3.	Citra	Kurang	Kurang	Baik	Baik	Baik	Tidak Lulus
4.	Deni	Rajin	Baik	Cukup	Baik	Baik	Lulus
5.	Eka	Rajin	Kurang	Cukup	Cukup	Kurang	Lulus
6.	Fajar	Kurang	Cukup	Cukup	Baik	Kurang	Tidak Lulus
7.	Gilang	Rajin	Baik	Kurang	Baik	Baik	Lulus
8.	Hadi	Kurang	Cukup	Kurang	Cukup	Kurang	Tidak Lulus
9.	Intan	Kurang	Kurang	Kurang	Kurang	Baik	Tidak Lulus
10.	Joko	Rajin	Cukup	Baik	Baik	Kurang	Lulus
11.	Kiki	Rajin	Cukup	Baik	Cukup	Kurang	Lulus
12.	Lutfi	Rajin	Kurang	Baik	Kurang	Cukup	Lulus
13.	Maya	Kurang	Baik	Baik	Baik	Baik	Lulus
14.	Nando	Kurang	Cukup	Baik	Baik	Baik	Lulus
15.	Oki	Kurang	Baik	Cukup	kurang	Cukup	Lulus
16.	Putri	Rajin	Kurang	Cukup	Cukup	Kurang	Lulus
17.	Qori	Rajin	Baik	Kurang	Cukup	Baik	Tidak Lulus
18.	Rendi	Rajin	Cukup	Kurang	Baik	Cukup	Tidak Lulus
19.	Santi	Rajin	Kurang	Kurang	Kurang	Kurang	Tidak Lulus
20.	Tono	Kurang	Baik	Kurang	Kurang	Kurang	Tidak Lulus

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

3.2 Perhitungan Entropy Keseluruhan (Iterasi 1)

Langkah pertama yang dilakukan oleh sistem adalah menghitung nilai ketidakpastian (*Entropy*) dari keseluruhan dataset (20 kasus). Dari data latih pada tabel 4.1, terdapat 12 siswa Lulus dan 8 siswa Tidak Lulus.

Rumus *Entropy:*

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

Perhitungan *Entropy* Total :

$$Entropy = \left(-\frac{12}{20}\log_2\frac{12}{20}\right) + \left(-\frac{8}{20}\log_2\frac{8}{20}\right)$$

$$Entropy = (-0.6 \times -0.7369) + (-0.4 \times -1.3219)$$

$$Entropy = 0.4421 + 0.5288$$

$$Entropy = 0.9709$$

3.3 Perhitungan Information Gain (Iterasi 1)

Setelah nilai *Entropy* Total didapatkan (0.9709), sistem akan menghitung nilai *Information Gain* untuk masing-masing kriteria (Kehadiran, Materi, Senam, Jurus, Fisik) untuk menentukan atribut mana yang menjadi akar (*Root Node*).

Rumus *Information Gain*:

$$Gain(S, A) = Entropy(S) - \sum_{v=1}^{n} \frac{|S_v|}{|S|} Entropy(S_v)$$

Entropy Senam = Baik (7 Lulus, 1 Tidak Lulus):

$$Entropy(\text{Senam_Baik}) = \left(-\frac{7}{8}\log_2\frac{7}{8}\right) + \left(-\frac{1}{8}\log_2\frac{1}{8}\right)$$

$$Entropy(\text{Senam_Baik}) = 0.1685 + 0.3750 = 0.5435$$

Entropy Senam = Cukup (4 Lulus, 1 Tidak Lulus):

$$Entropy(\text{Senam_Cukup}) = \left(-\frac{4}{5}\log_2\frac{4}{5}\right) + \left(-\frac{1}{5}\log_2\frac{1}{5}\right)$$

$$Entropy(\text{Senam_Cukup}) = 0.2575 + 0.4644 = 0.7219$$

Entropy Senam = Kurang (1 Lulus, 6 Tidak Lulus):

$$Entropy(\text{Senam_Kurang}) = \left(-\frac{1}{7}\log_2\frac{1}{7}\right) + \left(-\frac{6}{7}\log_2\frac{6}{7}\right)$$

$$Entropy(\text{Senam_Kurang}) = 0.4010 + 0.1907 = 0.5917$$

Perhitungan Gain Atribut SENAM:

$$Gain(\text{Total}, \text{Senam}) = 0.9709 - \left(\frac{8}{20} \times 0.5435 + \frac{5}{20} \times 0.7219 + \frac{7}{20} \times 0.5917\right)$$

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

$$Gain(\text{Total, Senam}) = 0.9709 - (0.2174 + 0.1805 + 0.2071)$$

$$Gain(\text{Total, Senam}) = 0.9709 - 0.6050$$

$$Gain(\text{Total, Senam}) = 0.3659$$

Sistem melakukan mekanisme kalkulasi matriks di atas terhadap seluruh atribut yang ada secara otomatis. Rekapitulasi dari seluruh komputasi pada Iterasi 1 terhadap 20 data sampel disajikan pada tabel berikut.

Tabel 3.2 Perhitungan Information Gain

Node	Atribut Penilaian	Kasus	Lulus	Tidak Lulus	Entropy	Information Gain
1	Total Data	20	12	8	0.9709	
	Kehadiran					0.1024
	- Rajin	12	9	3	0.8113	
	- Kurang	8	3	5	0.9544	
	Materi					0.0240
	- Baik	7	5	3	0.8631	
	- Cukup	7	4	3	0.9852	
	- Kurang	6	3	3	1.0000	
	Jurus					0.0395
	- Baik	9	6	3	0.9183	
	- Cukup	6	4	2	0.9183	
	- Kurang	5	2	3	0.9710	
	Senam					0.3659
	- Baik	8	7	1	0.5436	
	- Cukup	5	4	1	0.7219	
	- Kurang	7	1	6	0.5917	
	Fisik					0.0200
	- Baik	9	6	3	0.9183	
	- Cukup	3	2	1	0.9183	
	- Kurang	8	4	4	1.0000	

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tabel di atas merupakan simulasi cara kerja algoritma menggunakan 20 sampel data. Pada sampel data ini, atribut Senam secara matematis memperoleh *Information Gain* tertinggi yaitu sebesar 0.3659, sehingga ditetapkan sebagai Simpul Akar (*Root Node*).

3.4 Perhitungan Komputasi Lanjutan (Iterasi 2)

Karena ketiga cabang dari atribut Senam (Baik, Cukup, dan Kurang) masih memiliki nilai *Entropy* lebih dari 0 (masih terdapat percampuran data Lulus dan Tidak Lulus), maka algoritma akan melanjutkan komputasi ke Iterasi ke-2 untuk mencari atribut pemisah (*Node* Cabang) selanjutnya.

Sebagai contoh, berikut adalah perhitungan Iterasi 2 untuk cabang Senam = Cukup. Berdasarkan Tabel terdapat 5 data siswa dengan kategori Senam Cukup, yaitu Deni, Eka, Fajar, Oki, dan Putri. Rincian status kelulusannya adalah 4 Lulus dan 1 Tidak Lulus, sehingga nilai *Entropy* awalnya adalah 0.7219.

Sistem kemudian mengevaluasi atribut sisa (Kehadiran, Materi, Jurus, Fisik) pada subset 5 data tersebut. Apabila sistem menghitung *Information Gain* untuk atribut Materi, maka rincian distribusinya adalah:

Materi = Baik (2 kasus: Deni, Oki) → 2 Lulus, 0 Tidak Lulus. (*Entropy* = 0)

Materi = Cukup (1 kasus: Fajar) → 0 Lulus, 1 Tidak Lulus. (*Entropy* = 0)

Materi = Kurang (2 kasus: Eka, Putri) → 2 Lulus, 0 Tidak Lulus. (*Entropy* = 0)

Perhitungan *Gain* untuk atribut Materi pada cabang Senam Cukup:

$$Gain(Senam_Cukup, Materi) = 0.7219 - \left(\frac{2}{5}(0) + \frac{1}{5}(0) + \frac{2}{5}(0)\right)$$

$$Gain(Senam_Cukup, Materi) = 0.7219 - 0$$

$$Gain(Senam_Cukup, Materi) = 0.7219$$

Perolehan *Gain* yang sempurna ini (0.7219) menjadikan atribut Materi terpilih sebagai cabang turunan dari Senam = Cukup. Karena seluruh nilai *Entropy* pada cabang turunan Materi sudah mencapai kemurnian absolut (bernilai 0), maka iterasi dihentikan dan cabang tersebut berubah menjadi Simpul Daun (*Leaf Node* / Keputusan Akhir).

Algoritma di dalam sistem akan terus melakukan iterasi komputasi yang sama pada cabang Senam = Baik dan Senam = Kurang secara berulang hingga seluruh *Entropy* pada semua sub-cabang bernilai 0.

3.5 Pembentukan Aturan (*Rule*) Pohon Keputusan

Perhitungan matematis di atas mendemonstrasikan tahapan kerja sistem dalam memecah data. Ketika sistem aktual memproses himpunan *dataset* asli, algoritma C4.5 secara otomatis menyelesaikan seluruh iterasi dan merender visualisasi Pohon Keputusan (*Decision Tree*) yang utuh.

Struktur hierarki tersebut kemudian diekstraksi menjadi logika kondisional bersyarat (*If-Then*). Logika komputasional inilah yang ditanamkan pada sistem berbasis web sebagai mesin utama untuk mengeksekusi prediksi kelulusan siswa baru secara otomatis dan akurat. Ekstraksi kaidah aturan (*rules*) dari sistem adalah sebagai berikut:

IF Senam = Baik *AND* Materi = Baik / Cukup *THEN* Status Kelulusan = Lulus

IF Senam = Baik *AND* Materi = Kurang THEN Status Kelulusan = Tidak Lulus

IF Senam = Cukup *AND* Fisik = Kuat *THEN* Status Kelulusan = Lulus

IF Senam = Cukup *AND* Fisik = Lemah *AND* Kehadiran = Rajin *THEN* Status Kelulusan = Lulus

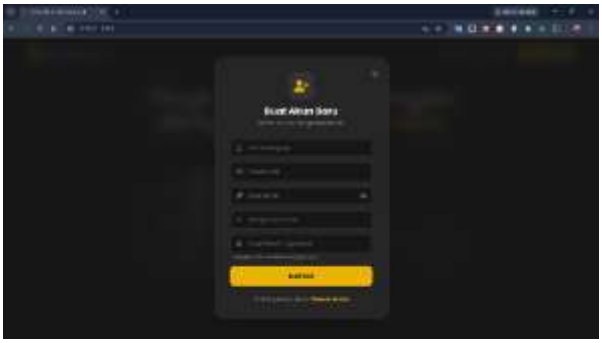
DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

IF Senam = Cukup *AND* Fisik = Lemah *AND* Kehadiran = Kurang *THEN* Status Kelulusan = Tidak Lulus

IF Senam = Kurang *AND* Jurus = Baik *THEN* Status Kelulusan = Lulus

IF Senam = Kurang *AND* Jurus = Cukup / Kurang *THEN* Status Kelulusan = Tidak Lulus

3.6. Implementasi Antarmuka



Gambar 3.1 Halaman Registrasi



Gambar 3.2 Homepage



Gambar 3.3 Olah Data

Antarmuka sistem dirancang secara dinamis. Halaman Login dan Registrasi Akun menerapkan autentikasi pop-up. Registrasi memiliki kolom "Kode Pelatih" opsional; jika diisi dengan kode valid, sistem mengenalinya sebagai Admin Pelatih. Jika dikosongkan, mendaftar sebagai Siswa dengan sifat view-only. Halaman Upload Dataset memfasilitasi admin mengimpor data spreadsheet. Halaman Pengolahan Data memvisualisasikan operasi otomatisasi sistem yang mengeliminasi anomali data (misal: "Total Data Awal 100", "Data Kosong Dihapus 21", "Data Siap Diproses 79"). Visualisasi Pohon Keputusan merender aturan C4.5 secara grafis interaktif. Formulir

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Prediksi memfasilitasi admin memasukkan data siswa baru, menekan "Hitung Prediksi", dan sistem mengeluarkan hasil Lulus/Tidak Lulus secara instan yang dapat diekspor ke PDF/Excel.

3.7. Evaluasi Akurasi dan Fungsionalitas

Pengujian Black Box terhadap seluruh modul fungsional sistem (meliputi Autentikasi, Manajemen Konten, Unggah Dataset, Pemrosesan C4.5, dan Validasi Akses View-Only Siswa) menunjukkan hasil 100% Valid. Kuesioner User Acceptance Test yang disebarkan kepada 20 responden menghasilkan skor 470 dari maksimal 500, mencapai persentase kepuasan sebesar 94% (Sangat Layak).

Pengujian performa sistem dilakukan melalui metode matriks kebingungan (Confusion Matrix) terhadap 20 sampel data uji untuk mengetahui keandalan sistem dalam memprediksi data baru. Hasil matriks mendeteksi True Positive (10 siswa aktual Lulus dan diprediksi Lulus), True Negative (8 siswa aktual Tidak Lulus dan diprediksi Tidak Lulus), False Positive (1 siswa aktual Tidak Lulus namun diprediksi Lulus), dan False Negative (1 siswa aktual Lulus namun diprediksi Tidak Lulus). Kalkulasi akurasi keseluruhan sistem dihitung menggunakan persamaan penjumlahan presisi:

$$Akurasi = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) \times 100\%$$

$$Akurasi = \left(\frac{10 + 8}{10 + 8 + 1 + 1}\right) \times 100\%$$

$$Akurasi = \left(\frac{18}{20}\right) \times 100\%$$

$$Akurasi = 0.90 \times 100\%$$

$$Akurasi = 90\%$$

Nilai akurasi 90% membuktikan model prediktif C4.5 yang ditanamkan pada aplikasi bekerja dengan tingkat presisi yang sangat tinggi, sehingga valid dijadikan landasan keputusan di Yayasan Setia Hati Terate.

4. Kesimpulan

Sistem prediksi kelulusan akademik berbasis web telah berhasil dirancang dan dibangun dengan mengintegrasikan algoritma klasifikasi data C4.5 secara utuh. Arsitektur sistem menerapkan alur pra-pemrosesan yang ketat berupa data cleansing dan data transform, serta membagi tingkat otorisasi secara proporsional. Proses inferensi matematika algoritma berhasil mengekstraksi atribut senam sebagai faktor paling determinan dalam kelulusan melalui perhitungan rasio perolehan informasi terbesar sebesar 0.3659. Model prediksi pohon keputusan yang dihasilkan di dalam sistem terbukti memiliki performa komputasi yang sangat andal, dengan mencatatkan tingkat akurasi mencapai 90% pada pengujian silang matriks kebingungan. Penerapan teknologi ini memberikan implikasi nyata bagi pelatih yayasan untuk merancang pembinaan terukur bagi siswa sejak dini, serta menciptakan transparansi evaluasi tanpa terikat kendala objektivitas maupun potensi kerusakan dokumen manual.

Referensi

1. Abdillah, M. F., Fitri, H., & Hardiansyah, M. R. (2025). Perkembangan Pencak Silat PSHT dan Dampaknya terhadap Masyarakat Medan, 1987–2023. *Polyscopia*, *2*(1), 29–33. https://doi.org/10.57251/polyscopia.v2i1.1552
2. Dina Amalia Putri, Naza Sefti Prianita, & Elkin Rilvani. (2025). Penerapan Metode C4.5 dan K-Nearest Neighbor0020untuk Klasifikasi Kelulusan Mahasiswa Berdasarkan Data Akademik. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, *3*(4), 256–267. https://doi.org/10.61132/jupiter.v3i4.1032
3. Fakih, A., Hamzami, M. A., Hadianto, M. R., & Alifah, N. I. S. (2025). Perbandingan Akurasi Algoritma C4.5 dan K-NN Untuk Prediksi Kelulusan Mahasiswa Penerima Beasiswa. *Jurnal Komputer Antartika*, *3*(1), 18–25. https://doi.org/10.70052/jka.v3i1.623
4. Fardiani, R., Responden, C., & Kunci Pendidikan, K. (2024). PENERAPAN ALGORITMA C4.5 UNTUK PREDIKSI PENERIMA BEASISWA SISWA BERPRESTASI DI MI AL-ISHLAH CIGANITRI. In *Jurnal Informatika) Universitas Muhammadiyah Tangerang P* (Vol. 8, Number 4).
5. Fatma Ayu Rahman, A., Wartulas, S., Raya Pagojengan, J. K., & Brebes, P. (2020). PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS DI UNIVERSITAS PERADABAN). In *Ade Fatma Ayu Rahman IJIR* (Vol. 1, Number 2).
6. Handoko, M. B., Fauzi, Z. W., Agung, H., & Sofia, D. (2025). Prediction of Science and Social Science Students Interests Using Decision Tree Algorithm with CRISP-DM. *Bit-Tech*, *8*(1), 437–447. https://doi.org/10.32877/bt.v8i1.2577
7. Hartono Renyut, D., Handayani Makassar, U., & Trinitas Ambon, S. (2022). *PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C.45 (Studi Kasus, Sekolah Tinggi Ilmu Administrasi Trinitas Ambon)*. *7*(2).

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

8. Reza Fauzy, Riki Winanjaya, & Susiani. (2022). Analisis Tingkat Kepuasan Pelanggan dengan Menerapkan Algoritma C4.5. *Bulletin of Computer Science Research*, *2*(2), 41–46. https://doi.org/10.47065/bulletincsr.v2i2.162
9. Siahaan, S. W., Sianipar, K. D. R., R.H Zer, P. P. P. A. N. W. F. I., & Hartama, D. (2020). Penerapan Algoritma C4.5 Dalam Meningkatkan Kemampuan Bahasa Inggris Pada Mahasiswa. *PETIR*, *13*(2), 229–239. https://doi.org/10.33322/petir.v13i2.1029
10. Syamsu, M., Terisia, V., Yusuf, D., Itb, A., & Dahlan, J. (2022). *PENERAPAN MODEL INFRASTRUKTUR ARTIFICIAL INTELLIGENCE SEBAGAI PENGGERAK INDUSTRI 4.0* (Vol. 3, Number 1).
11. Tintania Dwi Yuliani, D. S. R. Z. F. N. M. H. T. A. R. A. C. B. O. L. (2023). *PENERAPAN ALGORITMA C4.5 DALAM KLASIFIKASI DOMINASI JENIS KELAMIN PADA DATA PENDUDUK KOTA DEPOK BERDASARKAN KELOMPOK USIA DI TAHUN 2019-2023*.
12. Tukino. (2023). *Penerapan Metode Algoritma C4.5 dalam Penilaian Kelayakan Pemberian Kredit Kepada Mitra Usaha PT Arita Prima Sukses* (Vol. 15).
13. Ulfah, S., & Jepara, P. B. (2022). *KOMPARASI ALGORITMA C4.5 DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER UNTUK KLASIFIKASI STATUS GIZI BALITA COMPARISON OF C4.5 ALGORITHM WITH THE NAVE BAYES CLASSIFIER ALGORITHM FOR CLASSIFICATION OF NUTRITIONAL STATUS OF TOODLER*. *13*(1), 2548–4168. https://doi.org/10.34001/jdpt.v12i2
14. Wicaksono, A., Putro, G., & Setiadi, T. (2023). Penerapan Klasifikasi Decision Tree (C4.5) untuk Memprediksi Kelulusan Siswa Sekolah Dasar di Kecamatan Juai. *Jurnal Format*, *12*.
15. Yang, W., Long, H., Ma, L., & Sun, H. (2020). Research on clustering method based on weighted distance density and k-means. *Procedia Computer Science*, *166*, 507–511. https://doi.org/10.1016/j.procs.2020.02.056

DOI: https://doi.org/10.31004/riggs.v5i2.12534
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)