JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 5, No. 3, November 2026, pp. 376-390
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/175

Application of the C4.5 Algorithm for Classification of Brick Production Quality Results

Ari Syahputra*[1], Eva Darnila[2], Cut Agusniar[3]
[1,2,3] Teknik Informatika, Universitas Malikussaleh
Jl. Kampus Unimal Bukit Indah, Blang Pulo, Kec. Muara Satu, Kota Lhokseumawe, Aceh, indonesia
e-mail: ari.210170114@mhs.unimal.ac.id*[1], eva.darnila@unimal.ac.id[2], cut.agusniar@unimal.ac.id[3]

ARTICLE INFO

Article History:

Received:
Revised:
Accepted:
Published:

ABSTRACT

Red clay bricks are one of the most widely used construction materials in the building industry. However, the determination of brick quality is often carried out manually, which may result in inconsistent product quality due to subjective assessments. This study aims to apply the C4.5 algorithm to classify the quality of red clay bricks objectively based on physical testing results. The research data were obtained from measurements of burning duration, water absorption, compressive strength, and brick color. The research methodology consisted of data collection, data preprocessing, entropy and information gain calculations, and decision tree construction using the C4.5 algorithm. A total of 300 brick samples were analyzed and categorized into three quality classes: Good, Moderate, and Poor. The classification model was evaluated using a confusion matrix to measure its performance. The results showed that the C4.5 algorithm successfully classified brick quality with an accuracy rate of 81.11%, indicating a good level of classification performance. Furthermore, the decision tree analysis revealed that water absorption was the most influential attribute in determining brick quality. The developed system provides a practical and objective approach for quality control by reducing subjectivity and improving consistency in the evaluation process. The findings of this study contribute to the implementation of data mining techniques in the construction materials industry and support more effective, measurable, and reliable brick quality assessment processes.

Keywords: *C4.5 Algorithm, Brick Quality Classification, Data Mining, Decision Tree, Quality Control*

Abstrak

Batu bata merah merupakan salah satu material konstruksi yang banyak digunakan dalam industri pembangunan. Namun, penentuan kualitas batu bata masih sering dilakukan secara manual sehingga dapat menyebabkan ketidakkonsistenan mutu produk akibat adanya penilaian yang bersifat subjektif. Penelitian ini bertujuan untuk menerapkan algoritma C4.5 dalam mengklasifikasikan kualitas batu bata merah secara objektif berdasarkan hasil pengujian fisik. Data penelitian diperoleh dari pengukuran durasi pembakaran, daya serap air, kuat tekan, dan warna batu bata. Metodologi penelitian meliputi pengumpulan data, prapemrosesan data, perhitungan nilai *entropy*, *information gain*, *split info dan gain ratio,* serta pembentukan pohon keputusan menggunakan algoritma C4.5. Sebanyak 300 sampel batu bata dianalisis dan dikelompokkan ke dalam tiga kelas kualitas, yaitu Baik, Sedang, dan Buruk. Kinerja model klasifikasi dievaluasi menggunakan *confusion matrix* untuk mengukur tingkat performanya. Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan kualitas batu bata dengan tingkat akurasi sebesar 81,11%, yang menunjukkan performa klasifikasi yang baik. Selain itu, hasil analisis pohon keputusan mengungkapkan bahwa atribut daya serap air merupakan faktor yang paling dominan dalam menentukan kualitas batu bata. Sistem yang dikembangkan memberikan pendekatan yang praktis dan objektif dalam proses pengendalian kualitas dengan mengurangi subjektivitas serta meningkatkan konsistensi penilaian.

DOI: https://doi.org/10.51903/bc8vh970

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Temuan penelitian ini berkontribusi pada penerapan teknik data mining di bidang industri material konstruksi dan mendukung proses penilaian kualitas batu bata yang lebih efektif, terukur, dan andal.
Kata Kunci: Algoritma C4.5, Klasifikasi Kualitas Batu Bata, Data Mining, Pohon Keputusan, Pengendalian Kualitas

1. PENDAHULUAN

Batu bata merah merupakan salah satu material konstruksi yang masih banyak digunakan dalam pembangunan rumah, gedung, maupun infrastruktur karena memiliki harga yang relatif terjangkau, mudah diperoleh, dan memiliki daya tahan yang baik [1]. Kualitas batu bata menjadi faktor penting yang memengaruhi kekuatan, keamanan, dan umur bangunan yang dibangun. Batu bata yang tidak memenuhi standar kualitas dapat menyebabkan berbagai permasalahan konstruksi, seperti retak pada dinding, berkurangnya kekuatan struktur, dan meningkatnya biaya perawatan bangunan [2]. Oleh karena itu, diperlukan proses pengendalian mutu yang mampu memastikan kualitas batu bata sesuai dengan standar yang telah ditetapkan. Pada industri batu bata, proses penentuan kualitas masih banyak dilakukan secara manual berdasarkan pengamatan visual dan pengalaman pengrajin [3]. Metode tersebut cenderung menghasilkan penilaian yang berbeda karena dipengaruhi oleh faktor subjektivitas masing-masing individu. Dalam penelitian ini digunakan sebanyak 300 data batu bata yang diperoleh melalui pengujian fisik terhadap beberapa parameter, yaitu durasi pembakaran, daya serap air, kuat tekan, dan warna batu bata. Hasil pengamatan menunjukkan adanya variasi karakteristik pada setiap sampel sehingga kualitas batu bata yang dihasilkan tidak selalu seragam [2], [3]. Kondisi tersebut menyebabkan proses evaluasi kualitas menjadi lebih kompleks apabila dilakukan secara konvensional. Oleh karena itu, diperlukan suatu metode yang mampu membantu proses klasifikasi kualitas batu bata secara objektif dan berbasis data.

Pemanfaatan *Data Mining* kini semakin krusial di era teknologi informasi guna mengekstraksi wawasan berharga dari himpunan data yang massif [4]. Teknik klasifikasi menjadi instrumen utama dalam *Data Mining* yang berfungsi untuk mengkategorikan data ke dalam kelas-kelas spesifik berdasarkan atribut yang melekat pada data tersebut. Di antara berbagai metode yang ada, algoritma C4.5 sering menjadi pilihan utama dalam riset karena keunggulannya dalam menghasilkan model *Decision Tree* yang intuitif dan mudah diinterpretasikan [5]. Mekanisme kerja algoritma ini bertumpu pada kalkulasi *Entropy* dan *Information Gain* untuk mengidentifikasi atribut paling dominan dalam pembangunan struktur pohon Keputusan [6]. Fleksibilitas algoritma C4.5 dalam menangani data bertipe numerik maupun kategorikal menjadikannya sangat relevan untuk diimplementasikan pada beragam skenario klasifikasi. Eksistensi algoritma C4.5 dalam menyelesaikan berbagai persoalan klasifikasi telah divalidasi oleh banyak studi terdahulu. Misalnya, penelitian oleh Diah et al. [6] membuktikan bahwa model ini mampu memberikan tingkat akurasi yang signifikan dalam klasifikasi mutu produk.

Selain itu, temuan oleh Nursyahfitri et al. [7] mengonfirmasi bahwa metode ini mampu merumuskan aturan keputusan yang sistematis, sehingga mempermudah pengambilan keputusan. Efektivitas C4.5 dalam menyeleksi atribut paling berpengaruh terhadap akurasi klasifikasi juga telah dibuktikan dalam penelitian Ferdinan et al. [8] menunjukkan bahwa algoritma C4.5 efektif dalam mengidentifikasi atribut yang paling berpengaruh terhadap hasil klasifikasi. Lebih lanjut, literatur Damanik et al. [9] dan Widia et al. [10] memberikan kesimpulan bahwa algoritma C4.5 memiliki performa yang tangguh dalam menangani data yang heterogen serta mampu menyajikan model klasifikasi yang ramah bagi pengguna. Meskipun algoritma C4.5 telah banyak diterapkan pada berbagai bidang, penelitian mengenai klasifikasi kualitas batu bata berdasarkan kombinasi parameter durasi pembakaran, daya serap air, kuat tekan, dan warna masih relatif terbatas [11]. Sebagian besar penelitian sebelumnya lebih banyak berfokus pada objek klasifikasi lain, seperti kualitas produk pangan, prediksi penyakit, dan evaluasi kinerja. Selain itu, belum banyak penelitian yang mengkaji atribut fisik yang paling dominan dalam menentukan kualitas batu bata melalui pembentukan *Decision Tree*. Keterbatasan tersebut menunjukkan adanya *research gap* yang perlu diteliti lebih lanjut untuk mendukung proses pengendalian mutu batu bata secara lebih efektif dan terukur.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menerapkan algoritma C4.5 dalam mengklasifikasikan kualitas batu bata berdasarkan parameter durasi pembakaran, daya serap air, kuat tekan, dan warna. Penelitian ini menggunakan 300 data batu bata yang telah melalui proses pengujian fisik sebagai dasar pembentukan model klasifikasi. Kontribusi penelitian ini terletak pada penerapan algoritma C4.5 untuk menghasilkan model klasifikasi kualitas batu bata yang objektif serta mengidentifikasi atribut yang paling dominan dalam menentukan kualitas batu bata. Hasil penelitian diharapkan dapat membantu pengrajin dan pelaku industri dalam melakukan pengendalian mutu secara lebih efektif, konsisten, dan terukur sehingga dapat mendukung peningkatan kualitas produk batu bata yang dihasilkan.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

2. TINJAUAN PUSTAKA

2.1 Batu Bata

Sebagai material konstruksi fundamental, batu bata merah menempati posisi krusial dalam pembangunan dinding baik untuk hunian, bangunan komersial, maupun berbagai jenis infrastruktur. Secara teknis, proses produksinya melibatkan pencetakan tanah liat yang kemudian melalui tahapan pengeringan dan pembakaran pada suhu terukur demi memperoleh karakteristik kekerasan yang optimal [12]. Variabel penentu mutu batu bata sangat beragam, mencakup proporsi bahan penyusun, teknik cetak, lamanya durasi pembakaran, tingkat absorpsi air, ketahanan tekan, hingga warna yang dihasilkan setelah proses pembakaran [11]. Standar kualitas batu bata yang unggul ditandai dengan daya serap air yang minim, kapasitas kuat tekan yang memadai, serta keseragaman warna hasil pembakaran, yang secara kolektif meningkatkan durabilitas konstruksi terhadap beban struktural. Dalam industri konstruksi, kualitas batu bata menjadi aspek yang sangat penting karena berpengaruh langsung terhadap kekuatan dan keamanan bangunan. Batu bata yang tidak memenuhi standar kualitas dapat menyebabkan kerusakan pada struktur bangunan serta menurunkan umur pakai konstruksi [1], [2]. Oleh karena itu, diperlukan metode yang mampu membantu proses penilaian kualitas batu bata secara objektif dan konsisten. Penilaian kualitas yang dilakukan berdasarkan data fisik diharapkan mampu menghasilkan keputusan yang lebih akurat dibandingkan penilaian secara visual.

2.2 *Data Mining*

Secara fundamental, Data Mining merupakan prosedur sistematis untuk menelusuri pola, kecenderungan, maupun wawasan strategis yang tersimpan di dalam repositori data berskala besar, atau yang lebih dikenal dengan terminologi *Knowledge Discovery in Database (KDD)* [13]. Dalam praktiknya, proses ini mengintegrasikan berbagai teknik komputasi kompleks yang mencakup disiplin statistika, pemodelan matematis, serta kecerdasan buatan (*Artificial Intelligence*), termasuk di dalamnya penerapan *Machine Learning*. Melalui pendekatan ini, data mentah dianalisis secara mendalam untuk mengungkap korelasi tersembunyi, yang kemudian ditransformasikan menjadi basis pengetahuan krusial guna meningkatkan presisi prediksi serta pengambilan keputusan strategis dalam ranah bisnis [14].

2.3 Klasifikasi

Dalam ranah *Data Mining*, teknik klasifikasi difungsikan sebagai metode pengategorian data ke dalam kelas-kelas spesifik yang ditentukan berdasarkan karakteristik atribut yang melekat pada setiap entitas data [5]. Mekanisme klasifikasi beroperasi melalui pembentukan model berbasis *training data* yakni data yang kelasnya telah terdefinisi sebelumnya yang kemudian diaplikasikan untuk memprediksi label kelas pada data yang belum terklasifikasi [7]. Output dari proses ini berfungsi krusial sebagai landasan pengambilan keputusan strategis karena kemampuannya dalam menyingkap pola tersembunyi di balik kumpulan data. Terdapat berbagai metode klasifikasi yang mapan dalam literatur, di antaranya *Naïve Bayes*, *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM), serta *Decision Tree*. Setiap algoritma tersebut membawa karakteristik unik dengan tingkat kinerja yang bervariasi, sangat bergantung pada sifat serta struktur data yang diolah. Mengingat signifikansi pemilihan metode terhadap akurasi model yang dihasilkan, diperlukan seleksi algoritma yang tidak hanya unggul dari sisi performa, tetapi juga memiliki aspek transparansi yang memudahkan pengguna dalam menginterpretasikan aturan keputusan yang terbentuk.

2.4 *Decision Tree*

Decision Tree merupakan metode klasifikasi yang merepresentasikan proses pengambilan keputusan dalam bentuk struktur pohon. Struktur tersebut terdiri atas akar (*root node*), cabang (*branch*), dan daun (*leaf*) yang menunjukkan hasil klasifikasi [9]. Setiap simpul pada pohon keputusan merepresentasikan atribut yang digunakan sebagai dasar pemisahan data, sedangkan cabang menunjukkan nilai atribut yang digunakan dalam proses klasifikasi. Keunggulan utama *decision tree* adalah kemampuannya menghasilkan aturan keputusan (*rules*) yang mudah dipahami oleh pengguna [6]. Selain itu, metode ini dapat digunakan untuk mengolah data numerik maupun kategorikal sehingga banyak diterapkan dalam berbagai penelitian klasifikasi. Hasil klasifikasi yang diperoleh dapat dijadikan dasar dalam proses pengambilan keputusan karena mampu menunjukkan hubungan antar atribut secara jelas [10]. Oleh karena itu, *decision tree* menjadi salah satu metode yang banyak digunakan dalam bidang *data mining*.

2.5 *Algoritma C4.5*

Algoritma C4.5 merupakan metode *machine learning* yang digagas oleh Ross Quinlan sebagai evolusi dari *algoritma ID3* [15]. Inovasi utama dalam C4.5 terletak pada kemampuannya untuk menutupi keterbatasan pada versi pendahulunya, terutama terkait fleksibilitas dalam memproses atribut bernilai kontinu serta

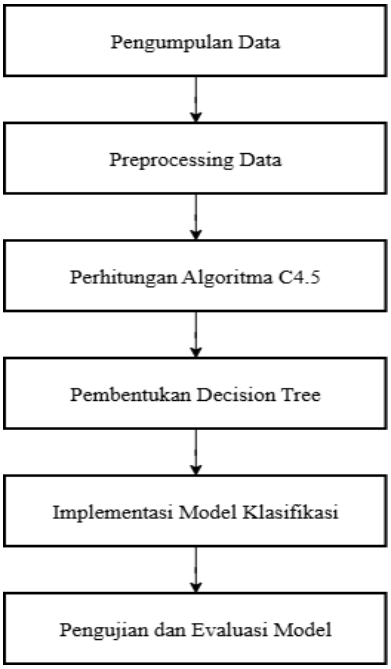
p-ISSN : 2803-1507 e-ISSN : 2803-1531

penanganan data yang tidak lengkap (*missing values*) [16]. Melalui mekanisme pembentukan pohon keputusan yang sistematis, algoritma ini mampu memetakan klasifikasi data berdasarkan atribut-atribut yang tersedia, sehingga menghasilkan model yang tidak hanya akurat, tetapi juga intuitif bagi pengguna untuk memahami proses logika di balik pengambilan keputusan.

3. METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Prosedur riset ini dirancang secara sistematis guna mengonstruksi model klasifikasi mutu batu bata berbasis algoritma C4.5. Rangkaian investigasi dijalankan secara terstruktur, mencakup seluruh fase mulai dari akuisisi data hingga tahap validasi performa model, guna menjamin ketepatan serta reliabilitas hasil klasifikasi. Metodologi yang diterapkan meliputi beberapa tahapan esensial: penyusunan rencana penelitian, penghimpunan data, *preprocessing* data, pembangunan model menggunakan algoritma C4.5, serta fase pengujian dan evaluasi untuk menarik simpulan akhir. Representasi visual dari alur kerja penelitian ini disajikan dalam Gambar 1.



Gambar 1. Tahapan Penelitian

3.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh melalui pengujian fisik terhadap 300 sampel batu bata yang diproduksi oleh UKM di Desa Ule Pulo, Kecamatan Dewantara. Pengumpulan data dilakukan melalui observasi lapangan dan pengujian laboratorium di Universitas Malikussaleh. Parameter yang diukur meliputi durasi pembakaran, kuat tekan, daya serap air, dan warna fisik batu bata.

Table 1. Data Sampel Hasil Pengujian Kualitas Batu Bata

Durasi Pembakaran (Jam)	Kekuatan Tekan (MPa)	Daya Serap Air (%)	Warna Fisik Batu Bata	Kualitas
48	16	8	Merah Tua	Baik
48	18.8	11.6	Hitam	Baik
60	9.7	8.9	Merah Tua	Buruk
60	15.3	16.1	Merah Muda	Sedang
72	13.2	13.4	Merah Muda	Sedang
72	16	17	Merah Tua	Sedang
84	16	21.5	Merah Tua	Buruk
84	18.8	25.1	Hitam	Buruk

Pengujian daya serap air dilakukan dengan membandingkan berat batu bata sebelum dan sesudah perendaman, sedangkan pengujian kuat tekan dilakukan menggunakan alat uji tekan untuk mengukur

p-ISSN : 2803-1507 e-ISSN : 2803-1531

kemampuan batu bata menahan beban maksimum. Warna fisik diamati secara visual sebagai indikator tingkat kematangan pembakaran. Penentuan label kualitas (Baik, Sedang, dan Buruk) mengacu pada standar SNI 15-2094-2000 berdasarkan hasil pengujian kuat tekan dan daya serap air. Seluruh data kemudian disusun menjadi dataset terstruktur yang digunakan dalam pembentukan model klasifikasi. Contoh data hasil pengujian ditunjukkan pada Tabel 1.

3.3 *Preprocessing Data*

Langkah *preprocessing* merupakan fase fundamental dalam menyiapkan dataset agar siap diolah oleh algoritma C4.5. Objektif utama dari tahapan ini adalah memperbaiki kualitas serta integritas data, sehingga proses klasifikasi menjadi lebih efisien dan akurat [17]. Kegiatan *preprocessing* mencakup deteksi dan penanganan *missing values*, seleksi atribut yang relevan, serta verifikasi konsistensi data terhadap parameter penelitian, yaitu durasi pembakaran (Jam), keuatan tekan (Mpa), daya serap air (%), serta warna batu bata [2],[18]. Pada tahap ini dilakukan penentuan interval nilai untuk atribut kontinu, yaitu kuat tekan dan daya serap air, berdasarkan standar SNI 15-2094-2000. Penetapan interval bertujuan untuk mempermudah proses klasifikasi serta memastikan bahwa setiap kategori mencerminkan kondisi mutu fisik batu bata secara objektif. Interval yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2.

Table 2. Interval Nilai Kuat Tekan dan Daya Serap Air Berdasarkan SNI 15-2094-2000

Kriteria	Kuat Tekan (MPa)	Daya Serap Air (%)
Sangat Rendah	≤ 10	≥ 20
Rendah	10 – 13	15 – 20
Sedang	13 – 15	10 – 15
Tinggi	≥ 15	≤ 10

Berdasarkan interval pada Tabel 2, sampel dengan kategori kuat tekan tinggi dan daya serap air rendah diklasifikasikan sebagai kualitas Baik. Sampel dengan nilai atribut pada rentang menengah diklasifikasikan sebagai kualitas Sedang, sedangkan sampel dengan kuat tekan sangat rendah dan daya serap air tinggi diklasifikasikan sebagai kualitas Buruk sesuai dengan ketentuan SNI 15-2094-2000. Setelah data dinyatakan valid dan terklasifikasi berdasarkan standar tersebut, dilakukan pembagian dataset sebanyak 300 entri ke dalam dua kelompok: *training set* dan *testing set* dengan rasio 70:30. Secara teknis, sebanyak 210 data dialokasikan untuk pembangunan model *decision tree*, sementara 90 data sisanya digunakan sebagai basis pengujian serta validasi performa model C4.5 yang telah terbentuk.

3.4 Penerapan Algoritma C4.5

Sebagai salah satu teknik klasifikasi terkemuka dalam *data mining*, algoritma C4.5 berfungsi untuk mengonstruksi model *decision tree* (pohon keputusan). Mekanisme kerjanya berfokus pada seleksi atribut paling signifikan yang ditentukan melalui kalkulasi *gain ratio* tertinggi untuk ditempatkan sebagai *root node* [19]. Proses pengelompokan dilakukan secara rekursif hingga seluruh data berhasil terklasifikasi ke dalam kelas-kelas target yang relevan. Dalam konteks penelitian ini, C4.5 diimplementasikan untuk mengevaluasi mutu hasil produksi batu bata berdasarkan parameter pengujian yang telah ditetapkan [12]. Implementasi algoritma ini melibatkan serangkaian perhitungan matematis yang mencakup *entropy*, *information gain*, *split information*, serta *gain ratio* pada setiap atribut variabel. Atribut yang memiliki nilai *gain ratio* paling optimal kemudian akan ditetapkan sebagai titik pusat (*root node*) dalam pembentukan struktur pohon keputusan [20].

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

Keterangan:
S: himpunan data (*dataset*)
Pi: probabilitas data pada kelas ke-i
n: jumlah kelas

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

Keterangan:
S: himpunan data
A: atribut yang diuji
S_i: subset data hasil pembagian berdasarkan atribut A
$| S_i |$: jumlah data pada subset ke-i
$| S |$: jumlah seluruh data

$$SplitInfo(S, A) = -\sum_{i=1}^{n} \frac{|S_i|}{|S|} \log_2 \left(\frac{|S_i|}{|S|}\right) \quad (3)$$

p-ISSN : 2803-1507 e-ISSN : 2803-1531

$$GainRatio(S,A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (4)$$

Keterangan:
$Gain(S,A)$: nilai information gain
$SplitInfo(S,A)$: nilai pembagian informasi
S_i: subset data
$|\ S_i\ |$: jumlah data pada subset ke-i
$|\ S\ |$: total data

3.5 Pembentukan Decision Tree

Output komputasi dari algoritma C4.5 berupa model struktur *decision tree* (pohon keputusan) yang berfungsi sebagai kerangka klasifikasi untuk menentukan mutu batu bata [21]. Struktur pohon yang telah terbentuk selanjutnya dikonversi ke dalam serangkaian aturan klasifikasi (*classification rules*). Aturan-aturan ini menjadi pedoman logis dalam menentukan kategori kualitas batu bata, yang didasarkan pada kombinasi nilai dari berbagai atribut input yang teridentifikasi dalam dataset [22].

3.6 Implementasi Sistem

Model klasifikasi yang dihasilkan dari algoritma C4.5 diimplementasikan ke dalam sistem berbasis web menggunakan bahasa pemrograman PHP dan database MySQL [23]. Sistem yang dikembangkan menyediakan fitur pengelolaan data batu bata, proses klasifikasi kualitas batu bata, visualisasi *decision tree*, serta penyajian hasil klasifikasi dan aturan (*rules*) secara otomatis [16][24]. Implementasi sistem bertujuan untuk mempermudah pengguna dalam melakukan pengendalian mutu batu bata secara cepat, objektif, dan konsisten berdasarkan hasil klasifikasi yang diperoleh.

3.7 Pengujian dan Evaluasi Model

Fase pengujian dilaksanakan guna mengukur efektivitas serta kapabilitas performa model klasifikasi yang telah dibangun. Untuk mengevaluasi keandalan model, digunakan pendekatan *Confusion Matrix* melalui perbandingan antara label prediksi sistem dengan kategori aktual di lapangan [25]. Parameter kinerja yang dijadikan acuan mencakup metrik *Accuracy*, *Precision*, *Recall*, serta *F1-Score*. Melalui analisis metrik-metrik tersebut, dapat diidentifikasi sejauh mana algoritma C4.5 mampu mengklasifikasikan kualitas batu bata secara presisi berdasarkan dataset yang diuji dalam riset ini.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Pengumpulan Data

Riset ini mengandalkan data primer yang diperoleh melalui serangkaian pengujian eksperimental langsung terhadap sampel batu bata. Variabel penelitian yang diamati meliputi durasi pembakaran, kekuatan tekan, daya serap air, dan karakteristik warna fisik, yang secara keseluruhan menghasilkan 300 observasi data. Untuk keperluan pemodelan, dataset ini dibagi menjadi dua bagian: 210 entitas sebagai *training set* untuk proses pembelajaran model, dan 90 entitas sebagai *testing set* untuk tahapan validasi. Seluruh informasi data disusun ke dalam format terstruktur dengan label target berupa kategori kualitas yang terdiri atas tiga kelas: *Baik*, *Sedang*, dan *Buruk*. Ringkasan karakteristik data hasil pengujian batu bata tersebut disajikan pada Tabel 3 berikut ini:

Tabel 3. Data Hasil Pengujian Kualitas Batu Bata

Atribut	Kelas	Baik	Sedang	Buruk	Jumlah	Total
Durasi Pembakaran (Jam)	48	11	26	14	51	210
	60	7	26	17	50	
	72	8	20	32	60	
	84	7	20	22	49	
Kekuatan Tekan (Mpa)	<= 10	0	0	28	28	210
	10 - 13	0	38	17	55	
	13 - 15	0	32	10	42	
	> =15	33	22	30	85	
Daya Serap Air (%)	< =10%	12	14	7	33	210
	10 - 15%	21	23	8	52	
	15 - 20%	0	55	7	62	
	>=20	0	0	63	63	
Warna Fisik Batu Bata	Merah Tua	11	28	30	69	210
	Merah Muda	10	27	34	71	
	Hitam	12	37	21	70	

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4.2 Hasil Preprocessing Data

Tahap *preprocessing* dilakukan untuk memastikan kualitas data sebelum proses klasifikasi menggunakan algoritma C4.5. Proses ini meliputi pemeriksaan kelengkapan data, identifikasi data duplikat, dan validasi konsistensi data pada seluruh atribut penelitian [26]. Berdasarkan hasil preprocessing terhadap 300 data batu bata, tidak ditemukan data kosong (*missing value*), data duplikat, maupun data yang tidak konsisten. Dengan demikian, seluruh data dinyatakan valid dan dapat digunakan pada proses pembentukan model klasifikasi. Ringkasan hasil *preprocessing* data disajikan pada Tabel 4.

Tabel 4. Hasil *Preprocessing Data*

Keterangan	Hasil
Jumlah data awal	300
Data kosong	0
Data duplikat	0
Jumlah data akhir	300

4.3 Perhitungan Algortima C4.5

Implementasi algoritma C4.5 dijalankan melalui serangkaian prosedur sistematis guna mengekstraksi aturan klasifikasi dari 210 *training data*. Tahapan dimulai dengan kalkulasi nilai *entropy* total untuk mengukur tingkat homogenitas dataset, yang kemudian diikuti dengan penentuan *information gain*, *split information*, serta *gain ratio* bagi setiap atribut. Atribut yang menunjukkan nilai *gain ratio* paling optimal ditetapkan sebagai *root node* (akar percabangan), mengingat efektivitasnya dalam memisahkan entitas data ke dalam kelas-kelas tujuan. Prosedur ini dijalankan secara rekursif hingga dihasilkan model *decision tree* yang utuh dan komprehensif.

$$Entropy(S) = -\sum_{i=1}^{n} p_i \log_2(p_i) \quad (5)$$

$$Entropy(S) = -\left(\frac{33}{210}\log_2\frac{33}{210}\right) - \left(\frac{92}{210}\log_2\frac{92}{210}\right) - \left(\frac{85}{210}\log_2\frac{85}{210}\right) = 1.469$$

$$Entropy(Daya\ Serap\ Air\ (< 15)) = -\left(\frac{12}{33}\log_2\frac{12}{33}\right) - \left(\frac{14}{33}\log_2\frac{14}{33}\right) - \left(\frac{7}{33}\log_2\frac{7}{33}\right) = 1.530$$

$$Entropy(Daya\ Serap\ Air\ (10 - 15)) = -\left(\frac{21}{52}\log_2\frac{21}{52}\right) - \left(\frac{23}{52}\log_2\frac{23}{52}\right) - \left(\frac{8}{52}\log_2\frac{8}{52}\right) = 1.464$$

$$Entropy(Daya\ Serap\ Air\ (15 - 20)) = -\left(\frac{0}{62}\log_2\frac{0}{62}\right) - \left(\frac{55}{62}\log_2\frac{55}{62}\right) - \left(\frac{7}{62}\log_2\frac{7}{62}\right) = 0.509$$

$$Entropy(Daya\ Serap\ Air\ (> 20)) = -\left(\frac{0}{63}\log_2\frac{0}{63}\right) - \left(\frac{0}{63}\log_2\frac{0}{63}\right) - \left(\frac{63}{63}\log_2\frac{63}{63}\right) = 0$$

Setelah estimasi nilai *entropy* untuk setiap kategori pada atribut daya serap air ditentukan, tahap berikutnya adalah mengukur nilai *Gain* dari atribut tersebut. Kalkulasi *Gain* dilakukan untuk mengidentifikasi tingkat efektivitas atribut dalam membedakan kelas mutu batu bata secara signifikan. Secara matematis, nilai *Gain* merupakan selisih antara *entropy* total dataset dengan akumulasi *entropy* setiap kategori yang telah dibobotkan berdasarkan proporsi distribusi data di dalamnya. Atribut yang menghasilkan nilai *Gain* tertinggi secara otomatis ditetapkan sebagai *root node* (simpul akar) pada struktur pohon keputusan, sebab parameter ini merepresentasikan variabel dengan pengaruh dominan dalam proses klasifikasi.

$$Gain(S - A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} Entropy(S_i) \quad (6)$$

$$Gain(Total\ -\ Daya\ Serap\ Air)$$
$$= 1.469\ -\left(\left(\frac{33}{210}\times 1{,}530\right) + \left(\frac{52}{210}\times 1{,}464\right) + \left(\frac{62}{210}\times 0.509\right) + \left(\frac{63}{210}\times 0\right)\right)$$
$$= 0.716$$

Tahap berikutnya melibatkan perhitungan *SplitInfo* serta *Gain Ratio* untuk atribut daya serap air guna mengevaluasi sejauh mana efektivitas atribut tersebut dalam mengklasifikasikan data ke dalam tingkatan mutu batu bata. Nilai *Gain Ratio* sendiri didapatkan melalui normalisasi *Information Gain* dengan cara membaginya terhadap *SplitInfo*. Dalam kerangka pembentukan pohon keputusan, atribut yang menghasilkan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

skor *Gain Ratio* paling tinggi akan dipilih sebagai atribut kunci karena memiliki performa pemisahan data yang paling optimal.

$$SplitInfo(S,A) = -\sum_{i=1}^{n} \frac{|S_i|}{|S|} \log_2 \left(\frac{|S_i|}{|S|}\right) \quad (7)$$

$$SplitInfo(Daya\ Serap\ Air) = -(\frac{33}{210}\log_2\frac{33}{210} + \frac{52}{210}\log_2\frac{52}{210} + \frac{62}{210}\log_2\frac{62}{210} + \frac{63}{210}\log_2\frac{63}{210} = 1.959$$

$$GainRatio(S,A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (8)$$

$$GainRatio(Daya\ Serap\ Air) = \frac{0.716}{1.959} = 0.365$$

Proses yang sama dilakukan pada seluruh atribut penelitian, yaitu durasi pembakaran, kekuatan tekan, daya serap air, dan warna fisik batu bata. Hasil perhitungan nilai *entropy, information gain, split information*, dan *gain ratio* dari setiap atribut disajikan pada Tabel 5.

Tabel 5. Hasil Perhitungan Algoritma C4.5

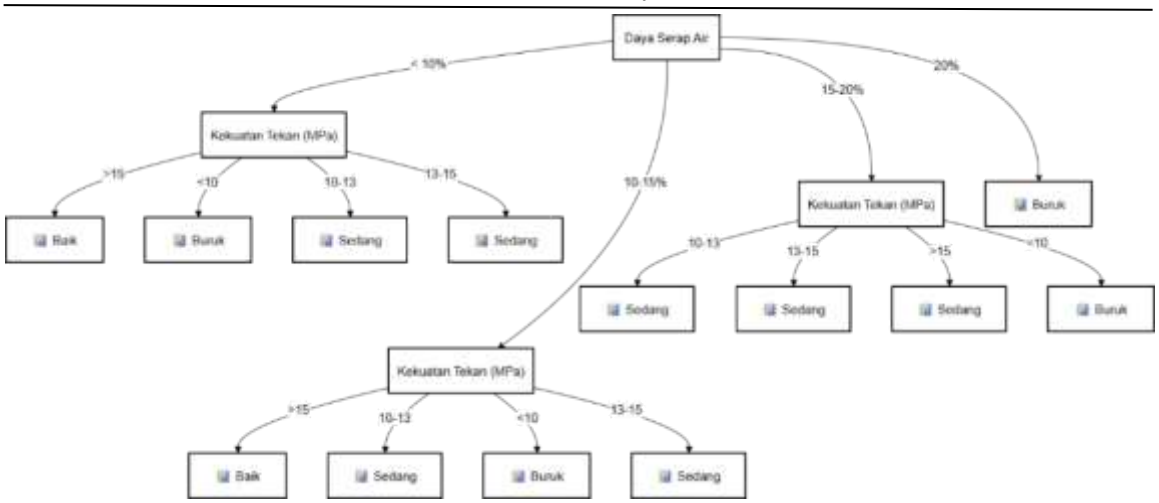
Atribut	Kelas	Jumlah	Baik	Sedang	Buruk	Entropy	Gain	SplitInfo	Gain Ratio
		210	33	92	85	1.469	0.034	1.993	0.017
Durasi (Jam)	48	51	11	26	14	1.485			
	60	50	7	26	17	1.417			
	72	60	8	20	32	1.400			
	84	49	7	20	22	1.447			
Kekuatan Tekan (Mpa)	< 10	28	0	0	28	0.000	0.444	1.886	0.235
	10 - 13	55	0	38	17	0.892			
	13 - 15	42	0	32	10	0.792			
	> 15	85	33	22	30	1.565			
	< 10	28	0	0	28	0.000			
Daya Serap Air	< 10%	33	12	14	7	1.530	0.716	1.959	0.3656
	10 - 15%	52	21	23	8	1.464			
	15 - 20%	62	0	55	7	0.509			
	>20	63	0	0	63	0.000			
Warna	Merah Tua	69	11	28	30	1.659	0.144	1.585	0.091
	Merah Muda	70	10	27	34	1.649			
	Hitam	71	12	37	21	1.533			

Merujuk pada hasil kalkulasi yang disajikan dalam Tabel 3, atribut Daya Serap Air ditetapkan sebagai *root node* karena mencatatkan nilai *gain ratio* tertinggi, yakni sebesar 0,3656. Angka ini mengindikasikan bahwa atribut tersebut memiliki daya pisah yang paling optimal dalam mengelompokkan data ke dalam tingkatan kualitas batu bata. Selanjutnya, atribut Kekuatan Tekan terpilih sebagai simpul pada level berikutnya dengan perolehan *gain ratio* sebesar 0,235. Di sisi lain, atribut warna fisik dan durasi pembakaran tidak terakomodasi dalam struktur *decision tree* yang terbentuk, dikarenakan nilai *gain ratio* yang dihasilkan relatif lebih rendah dibandingkan variabel lainnya.

4.4 Pembentukan *Decision Tree*

Melalui penerapan kalkulasi algoritma C4.5, berhasil dikonstruksi sebuah model *decision tree* (pohon keputusan) yang berfungsi untuk mengklasifikasikan mutu batu bata ke dalam tiga tingkatan: *Baik*, *Sedang*, dan *Buruk*. Penentuan struktur pohon dimulai dengan menempatkan atribut bernilai *gain ratio* tertinggi sebagai *root node* (akar), sementara atribut lainnya disusun secara hierarkis membentuk cabang-cabang klasifikasi berdasarkan hasil komputasi. Model yang terbentuk kemudian menjadi landasan logis dalam merumuskan aturan klasifikasi (*classification rules*) untuk penentuan kualitas batu bata secara sistematis. Visualisasi struktur pohon keputusan yang dihasilkan disajikan pada Gambar 2.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 2. Pohon Keputusan Klasifikasi Kualitas Batu Bata

Secara keseluruhan, kualitas batu bata dipengaruhi oleh kombinasi nilai daya serap air dan kekuatan tekan. Batu bata dengan daya serap air rendah (<15%) dan kekuatan tekan tinggi (>15 MPa) cenderung menghasilkan kualitas yang baik. Sebaliknya, daya serap air yang tinggi (>20%) menyebabkan kualitas batu bata cenderung masuk kategori buruk karena menunjukkan tingkat porositas yang tinggi.

4.5 Implementasi Sistem

Setelah dilakukan perhitungan algoritma C4.5 dan pembentukan decision tree, model klasifikasi yang dihasilkan diimplementasikan ke dalam sistem berbasis *web*. *Implementasi* ini bertujuan untuk menerapkan aturan klasifikasi yang terbentuk sehingga dapat digunakan untuk menentukan kualitas batu bata secara otomatis. Sistem menyediakan fitur pengelolaan data, proses klasifikasi, visualisasi decision tree, serta penyajian hasil klasifikasi. Selanjutnya akan dijelaskan tampilan dan fungsi dari setiap halaman pada sistem yang telah dikembangkan.

4.5.1 Halaman Login

Halaman *Login* merupakan halaman awal yang digunakan untuk melakukan proses autentikasi pengguna pada sistem Klasifikasi Kualitas Batu Bata. Pengguna diwajibkan memasukkan Username dan Password yang telah terdaftar untuk memperoleh hak akses ke dalam sistem. Desain antarmuka halaman dibuat sederhana dan intuitif guna memudahkan pengguna dalam melakukan proses autentikasi. Tampilan halaman *Login* dapat dilihat pada Gambar 3.

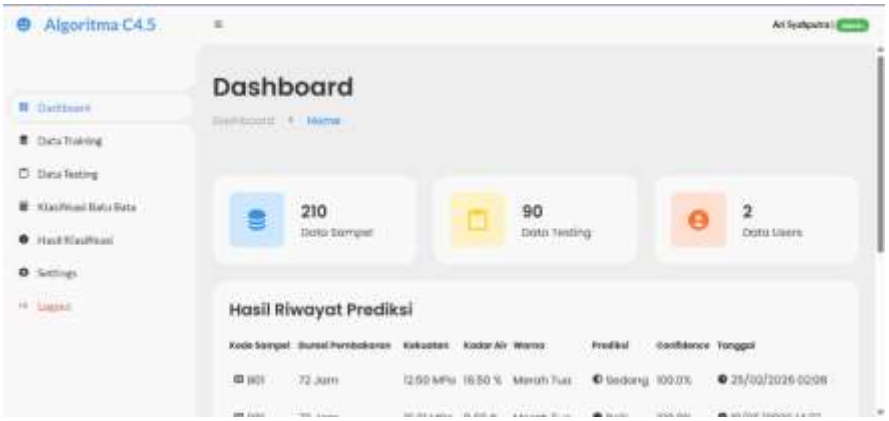


Gambar 3. Tampilan Halaman *Login*

4.5.2 Halaman *Dashboard*

Halaman *Dashboard* merupakan halaman utama yang ditampilkan setelah pengguna berhasil melakukan proses login pada sistem klasifikasi kualitas batu bata berbasis algoritma C4.5. Tampilan halaman Dashboard yang dikembangkan dalam sistem ini dapat dilihat pada Gambar 4. Halaman ini menyajikan ringkasan informasi yang berkaitan dengan data dan hasil klasifikasi kualitas batu bata.

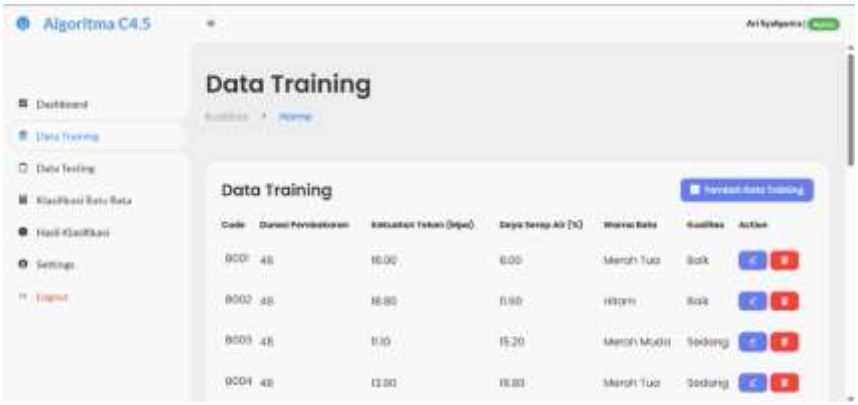
p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 4. Tampilan Halaman *Dashboard*

4.5.3 Halaman Data *Training*

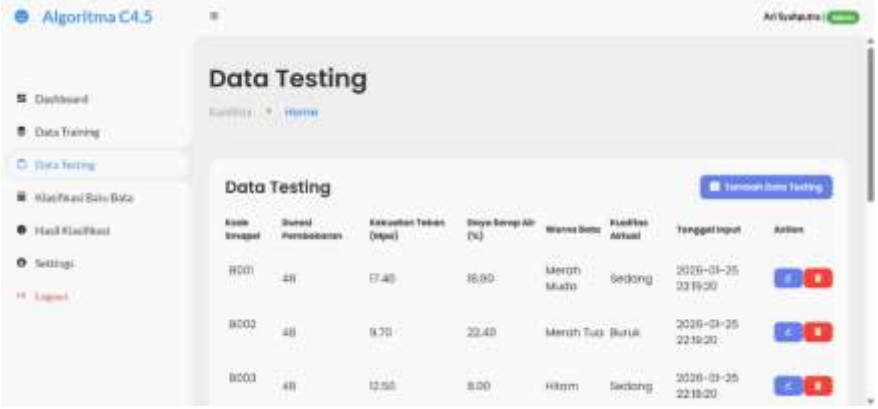
Halaman *Training* berfungsi sebagai media untuk mengelola dan menampilkan data pelatihan yang digunakan dalam proses pembentukan model klasifikasi kualitas batu bata menggunakan algoritma C4.5. Data pada halaman ini terdiri dari atribut-atribut yang menjadi dasar dalam proses perhitungan dan pembentukan decision tree. Tampilan halaman *Data Training* yang digunakan dalam penelitian ini ditunjukkan pada Gambar 5.



Gambar 5. Tampilan Halaman Data *Training*

4.5.4 Halaman Data *Testing*

Halaman Data *Testing* digunakan untuk menampilkan data uji yang berperan dalam proses evaluasi hasil klasifikasi menggunakan algoritma C4.5. Halaman ini menyajikan informasi atribut batu bata, kualitas aktual, tanggal input, serta fitur pengelolaan data. Tampilan halaman Data *Testing* dapat dilihat pada Gambar 6.

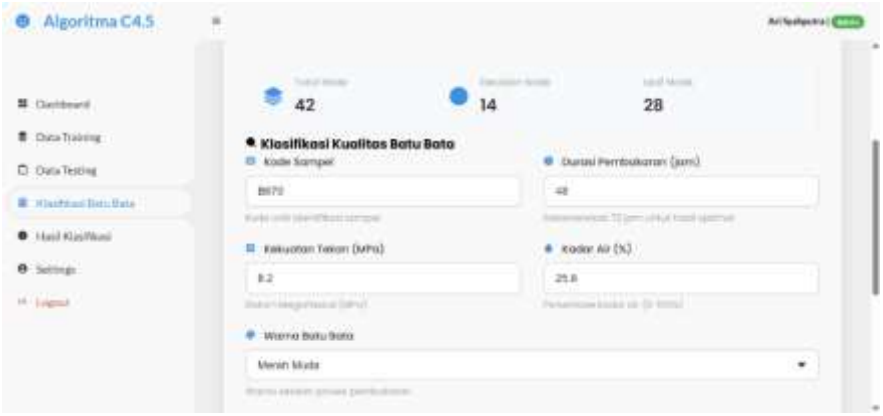


Gambar 6. Tampilan Halaman Data *Testing*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4.5.5 Halaman Klasifikasi

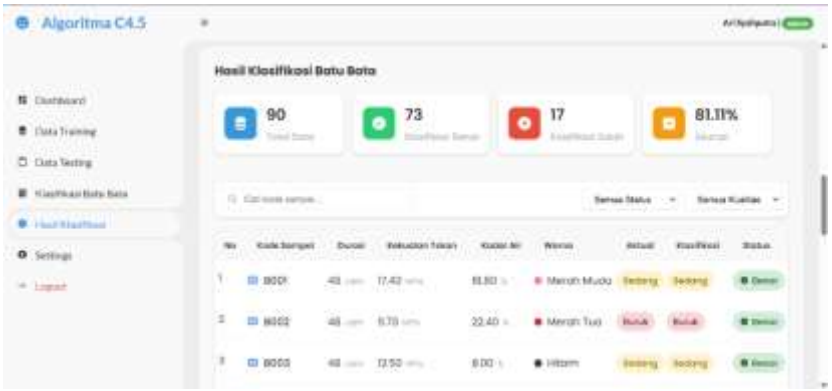
Halaman Klasifikasi C4.5 merupakan halaman yang digunakan untuk menjalankan proses klasifikasi kualitas batu bata berdasarkan model yang telah dibentuk menggunakan algoritma C4.5. Pada halaman ini tersedia fitur untuk membangun dan menampilkan decision tree, melakukan klasifikasi terhadap data pengujian, serta melakukan prediksi kualitas batu bata berdasarkan atribut yang diinputkan pengguna. Tampilan halaman Klasifikasi C4.5 ditunjukkan pada Gambar 7.



Gambar 7. Tampilan Halaman Klasifikasi

4.5.6 Halaman Hasil Klasifikasi

Halaman Hasil Klasifikasi C4.5 digunakan untuk menyajikan hasil evaluasi model klasifikasi berdasarkan data pengujian (*Testing* data). Halaman ini menampilkan informasi ringkasan yang meliputi jumlah total data uji, jumlah prediksi yang benar, jumlah prediksi yang salah, serta nilai akurasi yang diperoleh oleh sistem. Selain itu, tersedia tabel detail yang memuat informasi kode sampel, atribut pengujian, kelas aktual, hasil prediksi, nilai confidence, dan status klasifikasi. Penyajian informasi tersebut bertujuan untuk memudahkan pengguna dalam menganalisis dan mengevaluasi kinerja model klasifikasi yang dihasilkan oleh algoritma C4.5. Tampilan halaman Hasil Klasifikasi C4.5 ditunjukkan pada Gambar 8.



Gambar 8. Tampilan Halaman Hasil Klasifikasi

4.6 Hasil Pengujian dan Evaluasi Model

Pada tahap ini dilakukan pengujian terhadap model klasifikasi menggunakan algoritma C4.5 untuk mengetahui kemampuan model dalam mengklasifikasikan kualitas batu bata. Pengujian dilakukan terhadap data uji untuk menilai performa model yang telah dibentuk pada tahap sebelumnya. Sebagai gambaran hasil prediksi model terhadap data pengujian, beberapa contoh hasil klasifikasi ditampilkan pada Tabel 6. Tabel tersebut menunjukkan perbandingan antara kelas aktual dan hasil prediksi yang dihasilkan oleh algoritma C4.5.

Kolom Aktual menunjukkan kelas sebenarnya, sedangkan kolom Hasil menunjukkan hasil prediksi model. Kolom Status digunakan untuk menunjukkan kesesuaian antara data aktual dan hasil prediksi. Dari 90 data uji yang digunakan, sebanyak 73 data berhasil diklasifikasikan dengan benar dan 17 data mengalami kesalahan klasifikasi. Untuk mengetahui kinerja model secara lebih lanjut, dilakukan evaluasi menggunakan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

confusion matrix, *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil *confusion matrix* ditunjukkan pada gambar 9 berikut.

Tabel 6. Hasil Klasifikasi Data *Testing*

No	Durasi Pembakaran	Kekuatan Tekan	Daya Serap Air	Warna Fisik	Aktual	Hasil	Status
1	48	17.4	18.8	Merah Muda	Sedang	Sedang	Benar
2	60	16.7	8.9	Hitam	Baik	Buruk	Salah
3	72	12.9	29.4	Merah Tua	Buruk	Sedang	Salah
4	84	16.7	17.9	Hitam	Baik	Sedang	Salah



Gambar 9. *Confusion Matrix*

Merujuk pada Gambar 9, model klasifikasi C4.5 menunjukkan kemampuan yang cukup baik dalam mengklasifikasikan kualitas batu bata ke dalam tiga kelas target, yaitu Baik, Sedang, dan Buruk. Berdasarkan hasil prediksi, kelas Sedang mencatat 33 klasifikasi benar, diikuti kelas Buruk sebanyak 38 klasifikasi benar, sedangkan kelas Baik hanya menghasilkan 2 klasifikasi benar. Selain itu, terdapat 6 data kelas Baik yang diprediksi sebagai kelas Sedang dan 10 data kelas Sedang yang diprediksi sebagai kelas Baik. Hasil tersebut menunjukkan adanya kemiripan karakteristik antara kelas Baik dan Sedang yang menyebabkan model mengalami kesulitan dalam membedakan kedua kelas tersebut. Kondisi ini mengindikasikan bahwa performa model belum merata pada seluruh kelas meskipun secara keseluruhan mampu mengidentifikasi sebagian besar data dengan benar.



Gambar 10. *Metrik Evaluasi*

Merujuk pada hasil evaluasi yang disajikan pada Gambar 10, model klasifikasi berbasis algoritma C4.5 memperoleh tingkat akurasi sebesar 81,11%. Nilai tersebut menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data uji dengan benar. Namun, hasil evaluasi per kelas menunjukkan adanya perbedaan performa yang cukup signifikan. Kelas Baik memperoleh nilai *precision* sebesar 16,67%, recall 25,00%, dan F1-score 20,00%, yang mengindikasikan bahwa model masih mengalami kesulitan dalam mengenali sampel pada kategori tersebut. Rendahnya performa kelas Baik diduga dipengaruhi oleh ketidakseimbangan distribusi data (*class imbalance*) serta kemiripan karakteristik atribut dengan kelas Sedang.

Akibatnya, beberapa data yang termasuk kelas Baik diprediksi sebagai kelas Sedang sehingga menurunkan nilai *precision, recall*, dan F1-*score*. Sementara itu, kelas Sedang menunjukkan performa yang lebih stabil dengan nilai *precision* sebesar 82,50%, recall 76,74%, dan F1-*score* 79,52%. Hasil ini menunjukkan bahwa

p-ISSN : 2803-1507 e-ISSN : 2803-1531

model cukup mampu mengenali pola yang terdapat pada kelas Sedang meskipun masih ditemukan sejumlah kesalahan klasifikasi dengan kelas lainnya.

Performa terbaik ditunjukkan oleh kelas Buruk dengan nilai precision sebesar 100,00%, recall 97,44%, dan F1-score 98,70%. Tingginya nilai tersebut menunjukkan bahwa karakteristik data pada kelas Buruk lebih konsisten dan mudah dibedakan dibandingkan kelas lainnya. Hal ini didukung oleh hasil pembentukan *decision tree* yang menunjukkan bahwa atribut daya serap air menjadi akar pohon keputusan (*root node*), sehingga memiliki pengaruh paling dominan dalam proses klasifikasi. Sampel pada kelas Buruk umumnya memiliki nilai daya serap air yang tinggi dan kuat tekan yang rendah, sehingga dapat dipisahkan dengan lebih jelas oleh model. Secara keseluruhan, hasil pengujian menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan kualitas batu bata dengan cukup baik. Meskipun demikian, performa model belum merata pada seluruh kelas, terutama pada kelas Baik yang masih memiliki nilai *precision, recall,* dan F1-*score* yang rendah. Oleh karena itu, diperlukan peningkatan kualitas dan keseimbangan distribusi data pada penelitian selanjutnya untuk meningkatkan kemampuan model dalam mengenali seluruh kategori kualitas secara lebih optimal.

4.7 Diskusi

Temuan penelitian ini menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan mutu batu bata ke dalam tiga kelas, yaitu Baik, Sedang, dan Buruk, dengan tingkat akurasi sebesar 81,11%. Hasil tersebut menunjukkan bahwa model dapat mengenali pola hubungan antara atribut fisik batu bata, meliputi durasi pembakaran, kuat tekan, daya serap air, dan warna fisik, terhadap kualitas produk yang dihasilkan. Dengan kemampuan menghasilkan aturan klasifikasi yang mudah dipahami, algoritma C4.5 berpotensi digunakan sebagai alat bantu dalam proses penilaian kualitas batu bata secara lebih objektif. Dalam proses pembentukan pohon keputusan, atribut daya serap air teridentifikasi sebagai akar pohon keputusan (*root node*) karena memiliki nilai *gain ratio* tertinggi.

Hal ini menunjukkan bahwa daya serap air merupakan atribut yang paling berpengaruh dalam menentukan kualitas batu bata. Secara teknis, daya serap air yang rendah mengindikasikan struktur pori yang lebih rapat sehingga batu bata memiliki kekuatan dan ketahanan yang lebih baik. Sebaliknya, daya serap air yang tinggi menunjukkan tingkat porositas yang besar sehingga berpotensi menurunkan kualitas batu bata. Selain daya serap air, atribut durasi pembakaran, kuat tekan, dan warna fisik juga berkontribusi dalam proses klasifikasi. Durasi pembakaran yang optimal berpengaruh terhadap tingkat kematangan material, sedangkan kuat tekan mencerminkan kemampuan batu bata dalam menahan beban. Sementara itu, warna fisik dapat digunakan sebagai indikator visual untuk menilai kualitas proses pembakaran. Kombinasi keempat atribut tersebut memungkinkan algoritma C4.5 membentuk aturan keputusan yang mampu membedakan kualitas batu bata berdasarkan karakteristik fisiknya.

Hasil evaluasi menunjukkan bahwa performa model belum merata pada seluruh kelas. Kelas Buruk memperoleh nilai *precision, recall*, dan *F1-score* yang sangat tinggi karena memiliki karakteristik yang lebih konsisten dan mudah dibedakan dari kelas lainnya. Dominannya atribut daya serap air sebagai root node menunjukkan bahwa sebagian besar sampel pada kelas Buruk memiliki nilai daya serap air yang tinggi sehingga dapat dikenali dengan baik oleh model. Sebaliknya, kelas Baik memperoleh nilai *precision* sebesar 16,67%, *recall* 25,00%, dan F1-*score* 20,00%, yang menunjukkan bahwa model masih mengalami kesulitan dalam mengidentifikasi sampel pada kategori tersebut. Kondisi ini diduga disebabkan oleh kemiripan karakteristik antara kelas Baik dan Sedang, terutama pada atribut kuat tekan dan daya serap air yang memiliki rentang nilai yang berdekatan. Selain itu, ketidakseimbangan distribusi data (*class imbalance*) juga dapat memengaruhi kemampuan model dalam mengenali sampel pada kelas Baik secara optimal.

Temuan penelitian ini sejalan dengan penelitian terdahulu yang menunjukkan bahwa algoritma C4.5 memiliki kemampuan yang baik dalam menyelesaikan permasalahan klasifikasi melalui model pohon keputusan yang mudah diinterpretasikan. Keunggulan utama algoritma ini terletak pada kemampuannya untuk menghasilkan aturan klasifikasi yang jelas serta mengidentifikasi atribut yang paling berpengaruh dalam proses pengambilan keputusan. Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan, terutama pada jumlah *dataset* yang digunakan dan jumlah atribut yang dianalisis. Selain itu, penilaian warna fisik masih dilakukan secara visual sehingga berpotensi menimbulkan subjektivitas dalam proses pengamatan. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar, menambahkan atribut yang relevan, serta mengintegrasikan teknologi pengolahan citra (*image processing*) untuk meningkatkan objektivitas dan akurasi klasifikasi kualitas batu bata.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini menunjukkan bahwa algoritma C4.5 dapat digunakan untuk mengklasifikasikan kualitas batu bata ke dalam tiga kategori, yaitu Baik, Sedang, dan Buruk. Melalui proses perhitungan *entropy, information gain, split info,* dan *gain ratio*, model berhasil membentuk *decision tree* yang digunakan sebagai dasar klasifikasi. Hasil pengujian terhadap 90 data uji menghasilkan tingkat akurasi sebesar 81,11%, yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Hasil analisis menunjukkan bahwa atribut daya serap air menjadi akar pohon keputusan (*root node*) dan merupakan faktor yang paling berpengaruh dalam proses klasifikasi kualitas batu bata. Selain itu, atribut durasi pembakaran, kuat tekan, dan warna fisik juga berkontribusi dalam pembentukan aturan klasifikasi. Meskipun demikian, performa model belum merata pada seluruh kelas, terutama pada kelas Baik yang masih memiliki nilai *precision, recall*, dan *F1-score* yang rendah. Oleh karena itu, peningkatan kualitas data dan keseimbangan distribusi kelas perlu menjadi perhatian pada pengembangan penelitian selanjutnya. Implementasi model ke dalam sistem berbasis *web* menunjukkan bahwa proses klasifikasi kualitas batu bata dapat dilakukan secara lebih cepat, konsisten, dan terstruktur dibandingkan proses penilaian manual.

5.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa rekomendasi untuk pengembangan penelitian di masa mendatang agar model klasifikasi menjadi lebih andal dan komprehensif. Pertama, disarankan untuk meningkatkan kuantitas serta keragaman dataset guna memperkuat kemampuan generalisasi dan akurasi model klasifikasi. Kedua, efektivitas model dapat ditingkatkan dengan mengeksplorasi atribut lain yang memengaruhi kualitas batu bata, seperti kadar air bahan baku, tingkat kepadatan, serta komposisi tanah liat. Selain itu, perlu dilakukan studi komparatif dengan menguji algoritma lain seperti *Random Forest*, *Naïve Bayes*, *Support Vector Machine* (SVM), atau *K-Nearest Neighbor* (KNN) untuk mengidentifikasi metode yang memberikan hasil paling optimal. Dari aspek teknis pengembangan sistem, pengintegrasian teknologi *image processing* sangat disarankan untuk mengotomatisasi analisis warna dan kondisi permukaan, sehingga penilaian menjadi lebih objektif dan efisien. Terakhir, pengembangan sistem ke platform seluler atau integrasi dengan teknologi *Internet of Things* (IoT) berpotensi besar untuk mendukung pengawasan serta pengendalian mutu batu bata secara lebih fleksibel dan *real-time*.

DAFTAR PUSTAKA

[1] P. Tiwan, and P. G. Tamelan, “Kajian Proses dan Kualitas Bata Merah Hasil Pembakaran di Desa Muruona Kecamatan Ile Ape Kabupaten Lembata Study of The Process and Quality of Red Bricks Produced from Burning in Muruona Village, Ile Ape District, Lembata Regency,” 2025.

[2] H. Gustono and Hikmah, “Kualitas Bata Merah Nangalili Kecamatan Lembor Selatan, Kabupaten Manggarai Barat sebagai Bahan Bangunan,” Jurnal Batakarang, Vol. 2, No. 1, 2021.

[3] S. Frapanti, R. Efrida, I. Dewi, S. Asfiati, and F. V. Riza, “Analisis Standar Mutu Batu Bata Merah Tradisional di Deli Serdang dengan Indikator Sni 15-2094-2000,” Teras Jurnal : Jurnal Teknik Sipil, Vol. 13, No. 1, Pp. 163–172, Mar. 2023, Doi: 10.29103/Tj.V13i1.852.

[4] J. Khatib Sulaiman, M. Baru Puput Fadhilah Rizki, “Pendekatan Supervised Learning Decision Tree C4.5 untuk Klasifikasi Calon,” Indonesian Journal Of Computer Science Attribution-Sharealike, Vol. 12, No. 6, Pp. 2023–4228.

[5] A. H. Nasrullah, “Implementasi Algoritma Decision Tree untuk Klasifikasi Produk Laris,” Vol. 7, No. 2, 2021, [Online]. Available: Http://Ejournal.Fikom-Unasman.Ac.Id

[6] A. Diah Kuswanto, A. Fauzi, “Optimasi Kinerja Decision Tree C4.5 dengan Metode Backward Elimination pada Sistem Penilaian Kredit,” 2026.

[7] R. Nursyahfitri et al., “ Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree,” Jip (Jurnal Informatika Polinema) Halaman| 53.

[8] K. Ferdinan Leo Simanjuntak, A. Carolina Br Barus, “Implementasi Metode Decision Tree dan Algoritma C4.5 untuk Klasifikasi Kepribadian Masyarakat,” Joisie Journal Of Information System And Informatics Engineering, Vol. 5, No. 1, Pp. 51–59, 2021.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[9] S. F. Damanik, A. Wanto, and I. Gunawan, "Penerapan Algoritma Decision Tree C4.5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok." [Online]. Available: Https://Ejournal.Catuspata.Com/Index.Php/Jkdn/Index

[10] P. Widia Sari, G. Giwangga, N. Dian Fitriyanti, A. Arlisdia Putra, and H. Irawan, "Implementasi Algoritma C4.5 untuk Klasifikasi Data Insentif Karyawan Pt. Adhi Cakra Utama Mulia," Jurnal Komputer Antartika, Vol. 2, Pp. 95–101, 2024, [Online]. Available: Https://Ejournal.Mediaantartika.Id/Index.Php/Jka

[11] R. Syahputra and J. Siahaan, "Pengaruh Penambahan Natrium Silikat (Na2sio3) terhadap Daya Serap Air Dan Kuat Tekan pada Batako", Doi: 10.54123/Vorteks.V4i1.267.

[12] M. Arief and N. Syah, "Kualitas Batu Bata Berdasarkan Area Pembakaran (Suatu Studi Komperatif)," 2021.

[13] X. Shu and Y. Ye, "Knowledge Discovery: Methods from Data Mining and Machine Learning," Soc. Sci. Res., Vol. 110, Feb. 2023, Doi: 10.1016/J.Ssresearch.2022.102817.

[14] A. A. Alaskar, H. A. Almesned, N. N. Almuqati, and M. M. Hassan, "Big Data Mining and its Challenges: A Review."

[15] D. P. S. Sinaga, Rini Marwati, and Bambang Avip Priatna Martadiputra, "Aplikasi Web Prediksi Dampak Gempa di Indonesia menggunakan Metode Decision Tree dengan Algoritma C4.5," Jmt : Jurnal Matematika Dan Terapan, Vol. 5, No. 2, Pp. 97–108, Aug. 2023, Doi: 10.21009/Jmt.5.2.5.

[16] O. Maha et al., "Prediksi Paket Return menggunakan Metode Decision Tree Menerapkan Algoritma C4.5 Berbasis Website."

[17] V. N. Kristanto, I. Riadi, and Y. Prayudi, "Forensic Analysis of Faces on Low-Quality Images using Detection and Recognition Methods," Jurnal Resti, Vol. 7, No. 2, Pp. 218–225, Apr. 2023, Doi: 10.29207/Resti.V7i2.4630.

[18] W. Musu and A. Ibrahim, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5."

[19] Muhsi and Suprapto, "Modifikasi Gain Ratio pada Algoritma C4.5 dengan Nilai Koefisien Determinasi untuk Prediksi Kelulusan Mahasiswa (Studi Kasus : Universitas Islam Madura)," Jurnal Informatika: Jurnal Pengembangan It, Vol. 10, No. 1, P. 2025, Doi: 10.30591/Jpit.V9ix.Xxx.

[20] E. Oktavianti, M. Agustin, and R. Sari, "Implementasi Algoritma Decision Tree dengan Fitur Seleksi Weight By Information Gain," 2023.

[21] H. Syafputra and H. Latipa Sari, "Klasifikasi Penjualan Perhiasan menggunakan Metode Decision Tree Algoritma C4.5 (Studi Kasus : Toko Emas Berkat Famili)," Jurnal Media Infotama, Vol. 20, No. 2, P. 341139, 2695.

[22] I. K. Nasution, K. Nasution, and Mhd. Z. Siambaton, "Penerapan Algoritma Decision Tree untuk Meningkatkan Ketepatan Klasifikasi Kerusakan Alat Berat Berbasis Website," Sudo Jurnal Teknik Informatika, Vol. 3, No. 4, Pp. 183–200, Jun. 2025, Doi: 10.56211/Sudo.V3i4.741.

[23] J. P. and D. Kong, "Klasifikasi Sistem Penentu Keputusan Penerimaan Bantuan Sosial Covid-19 menggunakan Metode Analisis Hirarki Proses (Ahp) Studi Kasus Rt.04/05 Cengkareng Timur Jakarta Barat."

[24] O. Vian Agesti, F. M. Dewanto, and N. Q. Nada, "Calon Pendonor Darah di PMI Kota Semarang".

[25] P. B. Alamin and Diah Aryani, "Perbandingan Algoritma KNN, Naive Bayes aan SVM untuk Klasifikasi Status Stunting Anak Balita di Puskesmas Sepatan 1," Vol. 4, No. 4, Pp. 360–369, 2025, Doi: 10.55123.

[26] A. Rohman Zarkasi and F. Roma Doni, "Analisis Data Produksi menggunakan Metode Decision Tree (C4.5) Di Pt. Baja Marga Kharisma Utama."