



Klasifikasi Siswa Slow Learner Menggunakan Algoritma C4.5 Dalam Optimalisasi Pembelajaran di Sekolah Menengah Pertama

Bela Priska Gloria, Rahmi Hidayati*, Hirzen Hasfani

Fakultas Matematika dan Ilmu Pengetahuan Alam, Program Studi Rekayasa Sistem Komputer, Universitas Tanjungpura, Pontianak
Jl. Prof. Dr. H JI. Profesor Dokter H. Hadari Nawawi, Bansir Laut, Kec. Pontianak Tenggara, Kota Pontianak, Kalimantan Barat, Indonesia

Email: ¹belagloria99@student.untan.ac.id, ^{2,*}rahmihidayati@siskom.untan.ac.id, ³hirzen.hasfani@siskom.untan.ac.id

Email Penulis Korespondensi: rahmihidayati@siskom.untan.ac.id

Submitted: 06/02/2025; Accepted: 14/04/2025; Published: 15/04/2025

Abstrak—Pendidikan merupakan proses pembelajaran yang bertujuan untuk meningkatkan kemampuan siswa di lingkungan sekolah. SMP Negeri 12 Sungai Ambawang merupakan salah satu institusi pendidikan yang berlokasi di Kecamatan Sungai Ambawang. Dalam proses pembelajaran, setiap siswa memiliki tingkat pemahaman yang berbeda terhadap materi yang diajarkan. Beberapa siswa mengalami kesulitan dalam memahami pelajaran dengan kecepatan yang sama seperti teman sebayanya, sehingga dikategorikan sebagai slow learner. Ketidaktahuan terhadap siswa slow learner dapat menghambat proses belajar mengajar karena guru harus mengulang materi secara berulang. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan siswa slow learner menggunakan algoritma pembelajaran mesin C4.5 agar sekolah dapat merancang strategi pembelajaran yang lebih baik dan efektif. Klasifikasi siswa slow learner dibagi menjadi empat kelas, yaitu normal, ringan, sedang, dan berat. Dataset yang digunakan terdiri dari 135 data dengan 81 data latih dan 54 data uji. Atribut yang digunakan meliputi nilai mata pelajaran PKN, Bahasa Indonesia, Matematika, IPA, IPS, dan Bahasa Inggris. Algoritma C4.5 menghasilkan pohon keputusan dengan node akar pada atribut IPA. Pengujian menggunakan Confusion Matrix menunjukkan akurasi sebesar 91%, presisi 54%, recall 68%, dan tingkat error 9%. Persentase hasil klasifikasi menunjukkan bahwa 9% siswa berada di kelas normal, 24% kelas ringan, 62% kelas sedang, dan 0% kelas berat. Hasil ini menunjukkan bahwa algoritma C4.5 efektif dalam mengklasifikasikan siswa slow learner.

Kata Kunci: Algoritma C4.5; Data Mining; Klasifikasi; Pendidikan; Slow Learner

Abstract—Education is a learning process aimed at enhancing students' abilities in the school environment. SMP Negeri 12 Sungai Ambawang is one of the educational institutions located in Sungai Ambawang District. In the learning process, each student has a different level of understanding of the material being taught. Some students struggle to grasp lessons at the same pace as their peers, which categorizes them as slow learners. Lack of awareness about slow learners can hinder the teaching and learning process, as teachers must repeatedly explain the material. Therefore, this study aims to identify and classify slow learners using the C4.5 machine learning algorithm to help schools design more effective and adaptive learning strategies. The classification of slow learners is divided into four categories: normal, mild, moderate, and severe. The dataset consists of 135 data points, including 81 training samples and 54 testing samples. The attributes used include scores from subjects such as Civics (PKN), Indonesian Language, Mathematics, Natural Sciences (IPA), Social Sciences (IPS), and English. The C4.5 algorithm generates a decision tree with Natural Sciences (IPA) as the root node attribute. Testing using the Confusion Matrix shows an accuracy of 91%, precision of 54%, recall of 68%, and an error rate of 9%. The classification results indicate that 9% of students fall into the normal category, 24% into the mild category, 62% into the moderate category, and 0% into the severe category. These results demonstrate that the C4.5 algorithm is effective in classifying slow learners.

Keywords: C4.5 Algorithm; Classification; Data Mining; Education; Slow Learner

1. PENDAHULUAN

Pendidikan merupakan hal paling penting yang harus didapatkan oleh setiap manusia. Pendidikan menjadi sarana yang digunakan untuk manusia agar dapat mengembangkan potensi diri melalui proses pembelajaran yang didapat. Dengan adanya pendidikan diharapkan dapat melahirkan generasi penerus bangsa yang cerdas dan berkualitas [1]. Salah satu jalur pendidikan yang dapat ditempuh adalah pendidikan formal. Pendidikan formal [2] adalah lembaga yang disebut dengan sekolah yang merupakan bagian dari pendidikan yang berjenjang dan berkesinambungan. Dalam proses belajar mengajar disekolah, kemampuan berfikir siswa mempengaruhi tingkat keberhasilan belajar siswa disekolah tidak semua siswa mampu menerima materi pembelajaran yang disampaikan oleh guru dengan baik. Siswa dengan kondisi lambat dalam belajar dapat dikategorikan dalam siswa slow learner [3].

Slow learner adalah kondisi dimana siswa belajar lebih lambat dalam memahami pelajaran dari teman sekelasnya namun tidak memiliki disabilitas yang membutuhkan pendidikan khusus. Siswa dikatakan slow learner karena siswa tidak mampu mencapai apa yang diharapkan dari kelompok seusianya [4]. Dalam beberapa hal siswa lambat belajar mengalami hambatan atau keterlambatan berpikir, merespon rangsangan dan kemampuan untuk beradaptasi, tetapi lebih baik dibandingkan dengan tunagrahita. Siswa slow learner yang tidak diketahui akan menyebabkan memperlambat kegiatan belajar dan mengajar karena guru harus mengulang penjelasan materi kepada siswa slow learner [5]. Namun, cara mengetahui siswa mengalami slow learner merupakan tantangan yang cukup rumit bagi institusi Pendidikan. Untuk membantu pihak sekolah dalam melakukan pemantauan hasil belajar siswa disekolah, adapun cara untuk mengetahui apakah anak mengalami slow learner [6] adalah dengan menggunakan teknik data mining. Salah satu metode data mining yang dapat digunakan adalah Algoritma C4.5. Pada Algoritma C4.5 pemilihan atribut yang menjadi node akar dalam pohon keputusan sangat penting, karena

atribut yang terpilih akan menjadi langkah yang mempengaruhi proses pengambilan keputusan pada pohon keputusan. Dengan menerapkan Algoritma C4.5[7] akan dibangun pohon keputusan (decision tree) untuk membantu mengklasifikasikan siswa slow learner[8]. Dimana Algoritma C4.5 merupakan algoritma[9] klasifikasi data dengan teknik pohon keputusan dan akan menghasilkan aturan-aturan yang mudah diinterpretasikan dan menjadi algoritma tercepat diantara algoritma-algoritma lain. Hasil klasifikasi dari sistem akan digunakan untuk pengujian dengan confusion matrix[10] dimana hasil pengujian akan menunjukkan tingkat keberhasilan sistem dalam mengklasifikasikan siswa slow learner yang dapat dilihat dari tingkat akurasi yang dihasilkan oleh klasifikasi dengan Algoritma C4.5.

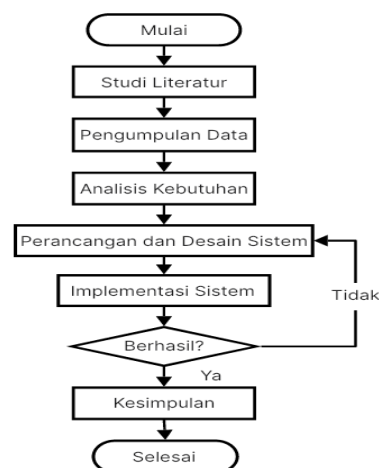
Penelitian terkait klasifikasi siswa slow learner telah pernah dilakukan sebelumnya oleh [11] dengan menggunakan Algoritma Naives Bayes Classifier dengan data yang digunakan sebanyak 71 data training dan 38 data testing yang dikategorikan kedalam 2 kelas klasifikasi yaitu slow learner dan non slow learner menghasilkan akurasi sebesar 92.86% precision sebesar 94.73%, recall sebesar 97.29%, dan F-Measure sebesar 96%. Selanjutnya, penelitian oleh [12] yang bertujuan untuk mengklasifikasi faktor penyebab menurunnya prestasi belajar mahasiswa pada masa pandemi dengan menggunakan teknik data mining yaitu Algoritma C4.5. Atribut yang digunakan adalah cara belajar, waktu belajar, pemahaman materi, pemberian tugas, dan lingkungan. Pengujian yang dilakukan menggunakan software Rapidminer diperoleh akurasi 97,5%. Terdapat pula penelitian oleh [13] yang menggunakan Algoritma C4.5 dalam mengklasifikasikan siswa yang berpotensi untuk di drop out dan tidak berpotensi drop out, yang menghasilkan 9 rules pada pohon keputusan dan akurasi sebesar 96,15%. Selanjutnya penelitian oleh [14] yang menggunakan data hasil rekam medis pasien diabetes melitus sebagai data klasifikasi penyebab penyakit diabetes menggunakan Algoritma C4.5 menghasilkan nilai tingkat akurasi yaitu rata-rata sebesar 98,5%.

Berdasarkan dari uraian yang telah dijabarkan sebelumnya, penelitian ini akan berfokus pada klasifikasi siswa slow learner dengan membuat sebuah sistem yang dapat digunakan untuk mengklasifikasikan siswa slow learner menggunakan Algoritma C4.5 yaitu untuk mengetahui atribut yang paling berpengaruh dalam menentukan kelas slow learner siswa, mengetahui persentasi kelas klasifikasi siswa slow learner dan mengetahui tingkat akurasi yang dihasilkan saat pengujian sistem. Hasil klasifikasi tersebut dapat digunakan sebagai dasar pengambilan keputusan untuk menangani siswa yang masuk kedalam kelas slow learner.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Terdapat 7 tahapan penelitian yang digunakan yaitu Studi Literatur dengan mengumpulkan sumber-sumber terkait penelitian dari buku, jurnal, dan majalah sebagai referensi yang berhubungan dengan Algoritma C4.5[15] dan klasifikasi siswa slow learner. Tahap kedua yaitu Pengumpulan Data, dalam penelitian ini data yang peneliti gunakan didapatkan dari SMP N 12 Sungai Ambawang. Tahap ketiga Analisis Kebutuhan, yaitu kebutuhan perangkat keras dan perangkat lunak yang dibutuhkan dalam penelitian. Tahap keempat Perancangan dan Desain Sistem, yaitu tahap yang menggambarkan rancangan sistem, membuat flowchart, Data Flow Diagram (DFD), basis data, relasi tabel, dan desain tampilan website yang akan digunakan dalam penelitian ini. Tahap kelima Implementasi, yaitu proses pembuatan sistem yang sudah dirancang dengan menggunakan bahasa pemrograman PHP, database yang digunakan menggunakan MySQL, sistem yang dibuat berbasis website. Tahap keenam Pengujian Sistem, pengujian dilakukan menggunakan confusion matrix dilakukan untuk mengukur tingkat accuracy, precision, recall, dan laju error. Tahap keenam Kesimpulan, yaitu tahap menuliskan hasil keseluruhan penelitian yang telah dilakukan. Diagram alir penelitian ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data yang diambil dari SMP Negeri 12 Sungai Ambawang. Data yang digunakan terdapat 6 kategori atribut yaitu atribut mata pelajaran Pendidikan Kewarganegaraan, Bahasa Indonesia, Matematika, Ilmu Pengetahuan Alam, Ilmu Pengetahuan Sosial, dan Bahasa Inggris. Data yang digunakan sebanyak 135 data dengan 81 data latih dan 54 data uji.

2.3 Algoritma C4.5

Algoritma C4.5 adalah algoritma yang difungsikan sebagai algoritma pembentuk pohon keputusan (Decision Tree)[16] dengan melakukan klasifikasi atau pengelompokan pada dataset. Metode dengan pohon keputusan akan mengubah data fakta data yang besar menjadi bentuk tatanan pohon keputusan yang merepresentasikan aturan (rule). Aturan keputusan yang dihasilkan dari Algoritma C4.5[17] dapat menjadi mudah dimengerti dengan bahasa alami dan dapat juga diubahkan kedalam bentuk bahasa yang basis data yaitu seperti SQL untuk mencari data pada kategori tertentu [18]. Algoritma C4.5 memiliki langkah-langkah yang digunakan untuk penyelesaian perhitungan, yaitu:

1. Menyiapkan data latih.
2. Menghitung nilai Entropy. Untuk mencari nilai Entropy dapat dihitung dengan Persamaan 1.

$$\text{Entropy}(S) = \sum_{i=1}^n (-p_i * \log^2 p_i) \tag{1}$$

3. Menghitung nilai dari information gain (Gain) untuk semua atribut data. Information gain (Gain) dapat dihitung menggunakan Persamaan 2.

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

4. Atribut yang mempunyai nilai Gain tertinggi akan menjadi node akar (root) dan atribut yang mempunyai nilai Gain lebih kecil dari akar (root) akan dipilih menjadi cabang (branches),
5. Menghitung lagi untuk nilai Entropy dan nilai Gain setiap atribut yang dengan tidak mengikut sertakan atribut yang dipilih menjadi akar (root) ditahap selanjutnya,
6. Atribut yang mempunyai nilai Gain paling tinggi lalu dipilih untuk menjadi node cabang (branches), Selanjutnya mengulang proses dari Langkah ke-4 dan Langkah ke-5 hingga pada akhir proses nilai Gain = 0 untuk semua data atribut yang masih tersisa.

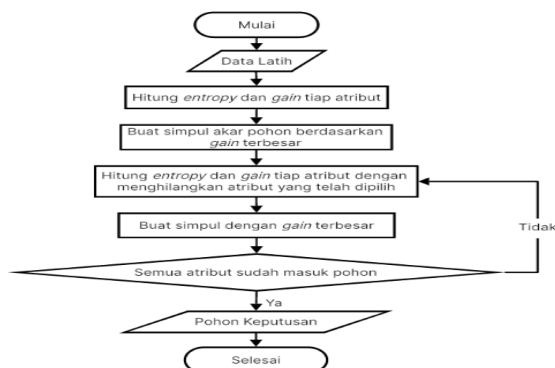
2.4 Perancangan Sistem dan Pengujian Sistem

Sistem yang akan dibangun adalah sistem klasifikasi siswa slow learner dengan menggunakan Algoritma C4.5. Perancangan pertama adalah diagram blok sistem yang menjadi gambaran umum pada sistem yang akan bekerja pada aplikasi siswa slow learner. Pada diagram blok sistem terdiri dari memasukkan data nilai siswa, data nilai selanjutnya diproses menggunakan Algoritma C4.5 dan menghasilkan keluaran berupa klasifikasi siswa slow learner. Diagram blok sistem ditunjukkan pada Gambar 2.



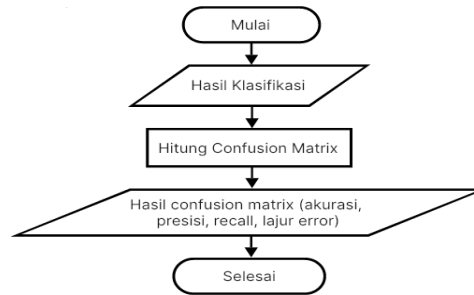
Gambar 2. Diagram Blok Sistem

Selanjutnya adalah diagram alir pelatihan proses dari Algoritma C4.5 Proses dimulai dengan melakukan input data latih, yang selanjutnya dilakukan perhitungan nilai entropy dan nilai gain setiap atribut, dari hasil tersebut dibuat simpul akar pohon berdasarkan dari nilai gain terbesar. Selanjutnya menghitung nilai entropy dan nilai gain setiap atribut dengan menghapus atribut yang sudah dipilih, lalu buat simpul pohon dengan gain terbesar. Lalu terakhir adalah pohon keputusan. Diagram alir pelatihan pada Algoritma C4.5 ditunjukkan pada Gambar 3.



Gambar 3. Diagram Alir Pelatihan Algoritma C4.5

Tahap pengujian pada Algoritma C4.5 pertama yaitu memasukkan hasil klasifikasi siswa slow learner. Selanjutnya hasil klasifikasi akan diuji dengan confusion matrix untuk menghasilkan nilai akurasi, presisi, recall dan laju error. Diagram alir pengujian Algoritma C4.5 ditunjukkan pada Gambar 4.



Gambar 4. Diagram Alir Pengujian Algoritma C4.5

Confusion matrix merupakan tabel pencatatan untuk hasil kerja pada klasifikasi. Kinerja umum dari confusion matrix untuk mengevaluasi menggunakan data dalam bentuk matrix [19]. Model klasifikasi multiclass digunakan untuk variabel yang memiliki lebih dari dua kelas berfungsi untuk menghasilkan nilai kinerja berupa TP, TN, FP, FN. Matrix yang digunakan adalah matrix N×N dimana baris akan mewakili kelas sebenarnya dan kolom akan mewakili kelas yang diprediksi [20]. Pengukuran kinerja confusion matrix 4×4 ditunjukkan pada Tabel 1.

Tabel 1. Confusion Matrix 4x4

Confusion Matrix	Prediksi				
	A	B	C	D	
Benar	A	AA	AB	AC	AD
	B	BA	BB	BC	BD
	C	CA	CB	CC	CD
	D	DA	DB	DC	DD

Berdasarkan hasil pada Tabel 1. Confusion Matrix, maka akan dihitung beberapa parameter kinerja sebagai berikut:

1. Accuracy (Akurasi)

Akurasi adalah nilai yang akan digunakan untuk mengukur keakuratan dari model dalam mengklasifikasikan dengan benar. Nilai akurasi didapatkan dari jumlah data dengan nilai positif yang diprediksikan positif dan data bernilai negatif yang diprediksi menjadi negatif dibagi dengan jumlah seluruh data didalam dataset. Rumus perhitungan akurasi ditunjukkan pada Persamaan 3.

$$\text{Akurasi} = \frac{AA+BB+CC+DD}{AA+AB+AC+AD+BA+BB+BC+BD+CA+CB+CC+CD+DA+DB+DC+DD} \times 100\% \tag{3}$$

2. Precision (Ketepatan)

Precision adalah nilai peluang dari kasus yang diprediksikan menjadi nilai positif yang pada faktanya adalah nilai dengan kasus kategori positif. Precision menggambarkan akurasi di antara data yang diminta dengan data dari hasil prediksi yang diberikan oleh model. Rumus perhitungan precision (ketepatan) ditunjukkan pada Persamaan 4.

$$\text{Precision/ketepatan} = \frac{\frac{AA}{AA+BA+CA+DA} + \frac{BB}{AB+BB+CB+DB} + \frac{CC}{AC+BC+CC+DC} + \frac{DD}{AD+BD+CD+DD}}{4} \times 100\% \tag{4}$$

3. Recall

Recall merupakan nilai dari peluang kasus kategori positif yang dengan tepat diprediksi menjadi positif. Recall akan memperlihatkan keberhasilan dari model dalam menemukan kembali sebuah data informasi. Rumus perhitungan recall ditunjukkan pada Persamaan 5.

$$\text{Recall} = \frac{\frac{AA}{AA+AB+AC+AD} + \frac{BB}{BA+BB+BC+BD} + \frac{CC}{CA+CB+CC+CD} + \frac{DD}{DA+DB+DC+DD}}{4} \times 100\% \tag{5}$$

4. Laju Error

Laju Error adalah pengukuran yang digunakan untuk mengetahui tingkat ketidakakuratan dari sistem dalam klasifikasi data secara benar. Rumus perhitungan laju error ditunjukkan pada Persamaan 6.

$$\text{Laju Error} = \frac{\text{jumlah data diidentifikasi salah}}{\text{jumlah seluruh data}} \times 100\% \tag{6}$$

3. HASIL DAN PEMBAHASAN

3.1 Perhitungan Dengan Algoritma C4.5

Pada tahap ini untuk mendapatkan hasil penelitian dilakukan pengujian terlebih dahulu pada data latih. Dari total data 135 dibagi menjadi 81 data latih dan 54 data uji. Perhitungan dilakukan dengan menggunakan metode Algoritma C4.5 yang menghitung dan mencari nilai entropy dan gain tiap atribut. Perhitungan dilakukan dengan menyiapkan data latih terlebih dahulu, data yang digunakan sebanyak 81 data. Lalu dilakukan perhitungan entropy dengan menggunakan Persamaan (1). Perhitungan entropy total sebagai berikut.

$$entropy(S) = \left(\left(-\frac{7}{81}\right) \log_2 \left(\frac{7}{81}\right) + \left(-\frac{18}{81}\right) \log_2 \left(\frac{18}{81}\right) + \left(-\frac{45}{81}\right) \log_2 \left(\frac{45}{81}\right) + \left(-\frac{11}{81}\right) \log_2 \left(\frac{11}{81}\right) \right) = 1,64976$$

Perhitungan entropy dari atribut Pendidikan Kewarganegaraan (A_1).

$$entropy(S_1) = \left(\left(-\frac{3}{4}\right) \log_2 \left(\frac{3}{4}\right) + \left(-\frac{1}{4}\right) \log_2 \left(\frac{1}{4}\right) + 0 + 0 \right) = 0,811278$$

$$entropy(S_2) = \left(\left(-\frac{4}{49}\right) \log_2 \left(\frac{4}{49}\right) + \left(-\frac{16}{49}\right) \log_2 \left(\frac{16}{49}\right) + \left(-\frac{29}{49}\right) \log_2 \left(\frac{29}{49}\right) + 0 \right) = 1,27019$$

$$entropy(S_3) = \left(0 + \left(-\frac{1}{27}\right) \log_2 \left(\frac{1}{27}\right) + \left(-\frac{16}{27}\right) \log_2 \left(\frac{16}{27}\right) + \left(-\frac{10}{27}\right) \log_2 \left(\frac{10}{27}\right) \right) = 1,154173$$

$$entropy(S_4) = \left(0 + 0 + 0 + \left(-\frac{1}{1}\right) \log_2 \left(\frac{1}{1}\right) \right) = 0$$

Selanjutnya dilakukan perhitungan gain dengan menggunakan Persamaan (2) sebagai berikut.

Perhitungan gain dari atribut Pendidikan Kewarganegaraan (A_1).

$$gain(S, A_1) = 1,64976 - \left(\left(\frac{4}{81} \cdot 0,811278\right) + \left(\frac{49}{81} \cdot 1,27019\right) + \left(\frac{27}{81} \cdot 1,154173\right) + \left(\frac{1}{81} \cdot 0\right) \right) = 0,456586$$

Perhitungan dilanjutkan kesemua atribut yang digunakan. Lalu tahap selanjutnya, adalah memilih nilai gain tertinggi yang akan menjadi node akar (root). Perhitungan node 1 dapat dilihat seperti pada Tabel 2.

Tabel 2. Perhitungan Node 1

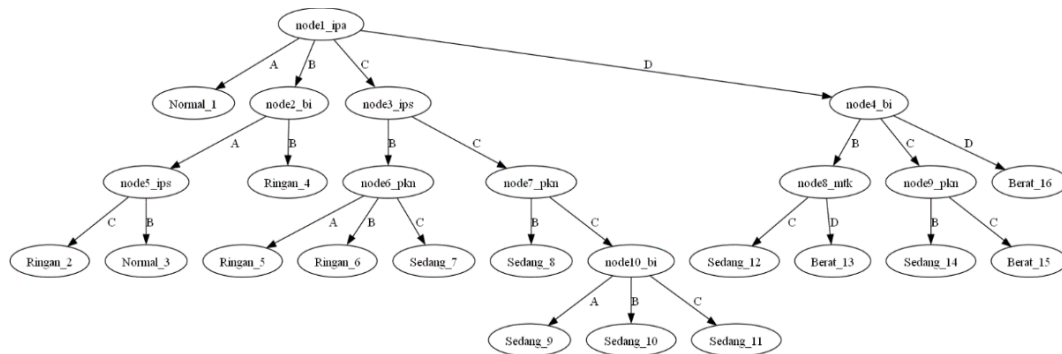
Atribut (A_i)	Kasus (S_i)	Jumlah Kasus	Kategori Nilai				Entropy	Gain
			Normal	Ringan	Sedang	Berat		
Total (S)		81	7	18	45	11	1,64976	
Pendidikan Kewarganegaraan (A_1)								0,456586
A (S_1)	4	4	3	1	0	0	0,811278	
B (S_2)	49	49	4	16	29	0	1,27019	
C (S_3)	27	27	0	1	16	10	1,154173	
D (S_4)	1	1	0	0	0	1	0	
Bahasa Indonesia (A_2)								0,682484
A (S_1)	12	12	7	3	2	0	1,384432	
B (S_2)	57	57	0	15	41	1	0,951083	
C (S_3)	11	11	0	0	2	9	0,684038	
D (S_4)	1	1	0	0	0	1	0	
Matematika (A_3)								0,280181
A (S_1)	0	0	0	0	0	0	0	
B (S_2)	10	10	6	4	0	0	0,970951	
C (S_3)	69	69	1	14	23	9	1,46705	
D (S_4)	2	2	0	0	0	2	0	
Ilmu Pengetahuan Alam (A_4)								0,814999
A (S_1)	2	2	2	0	0	0	0	
B (S_2)	14	14	5	9	0	0	0,940286	
C (S_3)	42	42	0	9	33	0	0,749595	
D (S_4)	23	23	0	0	12	11	0,998636	
Ilmu Pengetahuan Sosial (A_5)								0,539206
A (S_1)	0	0	0	0	0	0	0	
B (S_2)	22	22	7	13	2	0	1,28865	
C (S_3)	58	58	0	5	43	10	1,062148	
D (S_4)	1	1	0	0	0	1	0	
Bahasa Inggris (A_6)								0,128274
A (S_1)	0	0	0	0	0	0	0	
B (S_2)	2	2	2	0	0	0	0	
C (S_3)	78	78	5	18	45	10	1,580005	
D (S_4)	1	1	0	0	0	1	0	

Berdasarkan Tabel 2 atribut yang memiliki nilai gain tertinggi adalah atribut Ilmu Pengetahuan Alam (IPA) dengan nilai gain sebesar 0,814999, maka atribut Ilmu Pengetahuan Alam (IPA) menjadi node akar (node 1). Dari atribut Ilmu Pengetahuan Alam (IPA) terdapat 4 cabang untuk node 1 yaitu A, B, C, dan D. Atribut Ilmu Pengetahuan Alam (IPA) terdapat 3 atribut yaitu B, C, dan D, yang akan masih dilakukan lagi perhitungan selanjutnya yang untuk menghasilkan keputusan. Atribut A sudah menghasilkan kelas, yaitu Normal.

Perhitungan untuk node 2 yaitu atribut Ilmu Pengetahuan Alam (IPA) atribut B menggunakan total data sebanyak 14 data. Dari 14 data pada atribut B dilakukan perhitungan mencari nilai entropy dan gain pada setiap atribut dengan mengecualikan atribut dengan nilai gain tertinggi yaitu Pengetahuan Alam (IPA). Lalu perhitungan akan dilakukan sampai semua data memiliki kelas klasifikasi yaitu sampai pada 11 node yang hasilnya dapat dilihat pada Gambar 5. Pohon Keputusan. Pada perhitungan node 2, atribut yang memiliki nilai gain tertinggi adalah Bahasa Indonesia (BI) sebesar 0,508726 sehingga atribut Bahasa Indonesia (BI) menjadi node cabang. Dari atribut Bahasa Indonesia (BI) masih terdapat kasus yang dilakukan perhitungan yaitu kasus A. Lalu dilanjutkan perhitungan dengan menggunakan kasus A didapatkan atribut dengan gain tertinggi yaitu Ilmu Pengetahuan Sosial (IPS), dari atribut Ilmu Pengetahuan Sosial (IPS) semua data sudah tidak mempunyai data yang dapat dilakukan perhitungan kembali, maka perhitungan node 2 berhenti di atribut Ilmu Pengetahuan Sosial (IPS).

Perhitungan selanjutnya adalah perhitungan node 3 yaitu atribut Ilmu Pengetahuan Alam (IPA) kasus C. Perhitungan atribut Ilmu Pengetahuan Alam (IPA) kasus C dilakukan dengan menggunakan 42 data. Dari 42 data dilakukan perhitungan mencari hasil nilai entropy tiap kasus atribut yang tersedia dan menghitung nilai gain tiap atribut dengan mengecualikan atribut Ilmu Pengetahuan Alam (IPA) yang menjadi node akar. Pada perhitungan node 3 atribut yang mendapatkan nilai gain tertinggi adalah atribut Ilmu Pengetahuan Sosial (IPS), yaitu sebesar 0,487825. Dalam atribut Ilmu Pengetahuan Sosial (IPS) masih terdapat kasus yang dilakukan perhitungan lanjutan yaitu kasus B dan C. Perhitungan atribut Ilmu Pengetahuan Sosial dilanjutkan ke perhitungan node 6 untuk kasus B dan node 7 untuk kasus C. Didapatkan hasil gain tertinggi pada perhitungan node 6 adalah atribut Bahasa Indonesia. Pada perhitungan node 7 tidak terdapat kasus yang dapat dilakukan perhitungan lagi semua kasus didalam atribut sudah menghasilkan kelas klasifikasi.

Perhitungan selanjutnya adalah perhitungan node 4 yaitu atribut Ilmu Pengetahuan Alam (IPA) kasus D dengan menggunakan 23 data. Dari 23 data dilakukan perhitungan mencari hasil nilai entropy dan gain dengan mengecualikan atribut Ilmu Pengetahuan Alam (IPA) yang menjadi node akar. Pada perhitungan node 4 didapatkan atribut dengan nilai gain tertinggi adalah Bahasa Indonesia (BI) dengan nilai gain sebesar 0,57882. Dalam atribut Bahasa Indonesia (BI) masih terdapat kasus yang dilakukan perhitungan yaitu kasus B dan kasus C. Perhitungan kasus B dan kasus C dilanjutkan ke perhitungan node 8 dan node 9. Dari perhitungan node 8 didapatkan nilai gain tertinggi adalah atribut Matematika (MTK) dan perhitungan dari node 9 didapatkan nilai gain tertinggi adalah atribut Pendidikan Kewarganegaraan (PKN). Dengan hasil perhitungan yang telah dilakukan terbentuk pohon keputusan yang dapat digunakan untuk mengklasifikasikan data uji. Pohon keputusan klasifikasi siswa slow learner yang terbentuk dapat dilihat pada Gambar 5.



Gambar 5. Pohon Keputusan

Pada Gambar 5 pohon keputusan klasifikasi siswa slow learner atribut yang paling berpengaruh adalah atribut Ilmu Pengetahuan Alam (IPA) yaitu node akar dari pohon keputusan. Dalam pohon keputusan terdapat 16 keputusan klasifikasi yang dapat digunakan dalam klasifikasi data uji. Dengan menggunakan pohon klasifikasi dapat dibentuk aturan (rule) yang dapat digunakan untuk mengklasifikasikan data uji kedalam kelas klasifikasi. Aturan (rule) yang terbentuk adalah sebagai berikut.

1. If IPA = A then Normal
2. If IPA = B and BI = A and IPS = B then Normal
3. If IPA = B and BI = A and IPS = C then Ringan
4. If IPA = B and BI = B then Ringan
5. If IPA = C and IPS = B and PKN = A then Ringan
6. If IPA = C and IPS = B and PKN = B then Ringan
7. If IPA = C and IPS = B and PKN = C then Sedang
8. If IPA = C and IPS = C and PKN = B then Sedang

9. If IPA = C and IPS = C and PKN = C and BI = A then Sedang
10. If IPA = C and IPS = C and PKN = C and BI = B then Sedang
11. If IPA = C and IPS = C and PKN = C and BI = C then Sedang
12. If IPA = D and BI = B and MTK = C then Sedang
13. If IPA = D and BI = B and MTK = D then Berat
14. If IPA = D and BI = C and PKN = B then Sedang
15. If IPA = D and BI = C and PKN = C then Berat
16. If IPA = D and BI = D then Berat

3.2 Pengujian Data

Pengujian data dilakukan menggunakan confusion matrix pada 135 data nilai siswa. Data diuji menggunakan confusion matrix untuk mendapatkan hasil akurasi, precision, recall, dan laju error. Perbandingan data latih banding data uji yang digunakan ada 4 yaitu 60%:40%, 70%:30%, 80%:20%, dan 90%:10%. Confusion matrix perbandingan 60%:40% dapat dilihat pada Tabel 3.

Tabel 3. Confusion Matrix 60%:40%

Confusion matrix	Prediksi			
	Normal	Ringan	Sedang	Berat
Benar Normal	1	0	0	0
Benar Ringan	4	13	1	0
Benar Sedang	0	0	35	0
Benar Berat	0	0	0	0

Berdasarkan Tabel 3 dilakukan perhitungan pengujian dengan confusion matrix mencari nilai akurasi, presisi, recall, dan laju error sebagai berikut.

$$\text{Akurasi} = \frac{1+13+35+0}{1+0+0+0+4+13+1+0+0+0+0+35+0+0+0+0+0} \times 100\% = \frac{49}{54} \times 100\% = 91\%$$

$$\text{Presisi} = \frac{\left(\frac{1}{1+4+0+0} + \frac{13}{0+13+0+0} + \frac{35}{0+0+35+0} + \frac{0}{0+0+0+0}\right)}{4} \times 100\% = \frac{0,2 + 1 + 0,97 + 0}{4} \times 100\% = 54\%$$

$$\text{Recall} = \frac{\left(\frac{1}{1+0+0+0} + \frac{13}{4+13+1+0} + \frac{35}{0+0+35+0} + \frac{0}{0+0+0+0}\right)}{4} \times 100\% = \frac{1+0,72+1+0}{4} \times 100\% = 68\%$$

$$\text{Laju error} = \frac{0+0+0+4+1+0+0+0+0+0+0+0}{54} \times 100\% = 9\%$$

Pada perhitungan confusion matrix perbandingan 60%:40% didapatkan hasil akurasi sebesar 91%, nilai presisi sebesar 54%, nilai recall sebesar 68%, dan laju error 9%. Selanjutnya dilakukan perhitungan confusion matrix untuk perbandingan data uji dan latih 70%:30%, 80%:20%, dan 90%:10%. Hasil perhitungan dari confusion matrix ditunjukkan pada Tabel 4.

Tabel 4. Hasil Perhitungan Confusion Matrix

Perbandingan data latih: data uji	Akurasi %	Presisi %	Recall %	Laju Error %
60%:40%	91	54	68	9
70%:30%	88	54	67	13
80%:20%	81	49	39	19
90%:10%	77	42	35	23

3.3 Implementasi

Implementasi sistem dari aplikasi klasifikasi siswa slow learner menggunakan Algoritma C4.5 telah dibuat menggunakan bahasa pemrograman PHP dan database yang digunakan yaitu MySQL. Pada aplikasi yang dibuat data yang masuk akan diproses untuk menghasilkan output yaitu kelas Normal, Ringan, Sedang, dan Berat. Proses klasifikasi pada aplikasi dapat dilihat pada Gambar 6.



Gambar 6. Implementasi Antarmuka Halaman Klasifikasi



Pada Gambar 6 halaman ini pengguna dan admin dapat melakukan klasifikasi siswa dengan memasukkan data yang sesuai field pada form. Setelah data dimasukan dengan sesuai akan di hasilkan keluaran kelas siswa sesuai dengan rule yang terbentuk.

3.4 Pembahasan

Pada sistem yang dibuat ada dua user yang dapat menggunakan sistem yaitu admin dan pengguna, pengguna dapat langsung masuk ke sistem tanpa melakukan login terlebih dahulu dan dapat melakukan klasifikasi siswa pada halaman klasifikasi. Admin harus melakukan login ke dalam sistem dengan memasukkan email dan password kedalam halaman login, di dashboard admin ada halaman data klasifikasi dimana halaman tersebut berfungsi untuk menambah data klasifikasi dan melakukan proses pengujian dengan memilih perbandingan data untuk diuji. Pada saat data baru masuk pada halaman pengujian data yang dipilih admin harus menekan tombol refresh data agar data baru yang masuk mendapatkan kelas klasifikasi dari sistem. Proses refresh data difungsikan untuk memulai ulang metode klasifikasi dari sistem yang dibuat.

Penerapan Algoritma C4.5 pada penelitian ini menggunakan perbandingan 60%:40%. Penelitian ini menggunakan data nilai siswa SMP Negeri 12 Sungai Ambawang, data siswa yang digunakan adalah data pada tahun 2022/2023 kelas 7, kelas 8, dan kelas 9. Total data yang digunakan pada penelitian ini adalah 135 data serta menghasilkan data latih sebanyak 81 dan data uji sebanyak 54. Atribut penentu yang digunakan sebanyak 6 atribut yang terdiri dari nilai mata pelajaran Pendidikan Kewarganegaraan, nilai Bahasa Indonesia, nilai Matematika, nilai Ilmu Pengetahuan Alam (IPA), nilai Ilmu Pengetahuan Sosial (IPS), dan nilai Bahasa Inggris. Kelas klasifikasi yang digunakan ada 4 yaitu normal, ringan, sedang, dan berat. Dalam sistem yang telah dibuat pengguna dapat mendapatkan hasil klasifikasi pada sistem dengan memasukkan data nilai pengguna kedalam aplikasi, yang selanjutnya akan diproses dengan algoritma C4.5. Hasil keluaran dari sistem akan menunjukkan kelas klasifikasi dari pengguna.

Pada perbandingan data 60%:40% perhitungan dilakukan sampai pada node ke 11 pada perhitungan node ke 11 semua data sudah memiliki kelas klasifikasi. Perhitungan yang dilakukan menghasilkan rule sebanyak 16 rule yang digunakan dalam proses pengujian data uji dalam perbandingan data tersebut. Berdasarkan hasil pohon keputusan (decision tree) perbandingan 60%:40% atribut yang paling berpengaruh adalah atribut Ilmu Pengetahuan Alam (IPA) yaitu atribut yang menjadi simpul akar dari pohon keputusan. Hasil pengujian sistem yang dilakukan menggunakan confusion matrix dengan perbandingan 60%:40% menghasilkan akurasi sebesar 91%, presisi 54%, recall 68% dan laju error 9%. Perbandingan 70%:30% menghasilkan akurasi sebesar 88%, presisi 54%, recall 67%, dan laju error 13%. Perbandingan 80%:20% menghasilkan akurasi sebesar 74%, presisi 43%, recall 35%, dan laju error 26%. selanjutnya perbandingan 90%:10% menghasilkan akurasi sebesar 85%, presisi 44%, recall 41%, dan laju error 15%. Berdasarkan hasil didalam tabel pengujian menggunakan confusion matrix perbandingan data yang menghasilkan hasil akurasi yang optimal adalah perbandingan 60%:40%, karena perbandingan 60%:40% mendapatkan tingkat akurasi paling tinggi. Dilihat dari hasil pengujian perbandingan yang telah dilakukan, sistem ini berhasil mengklasifikasikan siswa slow learner menggunakan Algoritma C4.5.

4. KESIMPULAN

Berdasarkan pada penelitian yang telah dilakukan peneliti, kesimpulan yang dapat diambil dari keseluruhan penelitian yang telah dilakukan didapatkan kesimpulan yaitu atribut atau mata pelajaran yang paling berpengaruh dalam menentukan kelas klasifikasi siswa slow learner menggunakan Algoritma C4.5 adalah atribut Ilmu Pengetahuan Alam (IPA). Atribut tersebut merupakan atribut dari node tertinggi (node akar) pohon keputusan Algoritma C4.5. Berdasarkan hasil pengujian yang telah dilakukan menggunakan Algoritma C4.5, terdapat empat akurasi dengan hasil nilai yang berbeda. Perbedaan hasil akurasi yang didapatkan disebabkan karena penggunaan data latih dan data uji yang berbeda. Akurasi yang memiliki nilai tertinggi menggunakan data latih sebanyak 81 data dan data uji sebanyak 54 data menghasilkan nilai akurasi sebesar 91%. Dari hasil pengujian yang telah dilakukan dengan perbandingan data 60%:40% didapatkan tingkat persentase hasil klasifikasi dari 4 kelas yang digunakan. Pada kelas normal didapatkan persentase sebesar 9%, kelas ringan didapatkan persentase sebesar 24%, untuk kelas sedang didapatkan persentase sebesar 62%, dan kelas berat sebesar 0%. Untuk penelitian selanjutnya dapat dilakukan penambahan atribut yang digunakan dalam penelitian agar dapat menghasilkan rule lebih banyak. Atribut yang dapat ditambahkan, meliputi mata pelajaran Agama, mata pelajaran Seni, ataupun mata pelajaran Olahraga dan menggunakan jumlah jumlah total data penelitian yang lebih banyak dengan tujuan agar mendapatkan tingkat akurasi yang lebih baik.

REFERENCES

- [1] S. F. N. Fitri, "Problematika Kualitas Pendidikan di Indonesia," *J. Pendidik. Tambusai*, vol. 5, no. 1, pp. 1617–1620, 2021.
- [2] A. Ahmad, A. J. Tiara Sari, A. H. Wardana, M. N. I. Rosyid, E. Widiyanto, and A. Rasyad, "Tren Perkembangan Pendidikan Non-Formal," *J. Pendidik. (Teori dan Prakt.)*, vol. 7, no. 2, pp. 76–82, 2023, doi: 10.26740/jp.v7n2.p76-82.
- [3] A. Anggraeni, "Rancangan Program Pengembangan Pendidikan Individual untuk Siswa Lambat Belajar (Slow Learner),"



- Nathiqiyah, vol. 5, no. 2, pp. 48–55, 2022, doi: 10.46781/nathiqiyah.v5i2.423.
- [4] S. Nurfadhillah et al., “Analisis Strategi Guru dalam Mengajar Siswa Lambat Belajar atau Slow Learner di Kelas II SDN Kunciran Indah 7,” *Masaliq*, vol. 2, no. 1, pp. 53–63, 2022, doi: 10.58578/masaliq.v2i1.92.
- [5] N. M. Haryanti, “Layanan Pendidikan Anak Lamban Belajar (Slow Learner) di Sekolah,” *Al-Madrasah J. Pendidik. Madrasah Ibtidaiyah*, vol. 6, no. 2, p. 437, 2022, doi: 10.35931/am.v6i2.1006.
- [6] A. Anggraeni, “Gambaran Kognitif Pada Siswa Yang Mengalami Lambat Belajar (Slow Learner),” *J. RASI*, vol. 3, no. 1, pp. 1–8, 2021, doi: 10.52496/rasi.v3i1.116.
- [7] N. Azwanti and E. Elisa, “Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5,” *Pros. Semin. Nas. Ilmu Sos. dan Teknol.*, no. 3, pp. 126–131, 2020.
- [8] R. N. Putri, P. Nurani, and B. Amsal, “Gambaran Kemampuan Kreativitas Siswa Slow Learner (Lambat Belajar),” *J-Bkpi*, vol. 03, no. 02, pp. 67–79, 2023.
- [9] R. R. Andarista and A. Jananto, “Penerapan Data Mining Algoritma C4.5 Untuk Klasifikasi Hasil Pengujian Kendaraan Bermotor,” *J. Tekno Kompak*, vol. 16, no. 2, p. 29, 2022, doi: 10.33365/jtk.v16i2.1525.
- [10] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, “Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix,” *J. Inform. Upgris*, vol. 6, no. 2, pp. 66–75, 2021, doi: 10.26877/jiu.v6i2.6552.
- [11] A. H. Wicaksono et al., “Klasifikasi Siswa Slow Learner Untuk Mendukung Sekolah Algoritma Naïve Bayes Classification of Slow Learner Students To Support Schools in Improving Student Understanding Using the Naïve Bayes,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 589–596, 2022, doi: 10.25126/jtiik.202295509.
- [12] K. F. Irmada, D. Hartama, and A. P. Windarto, “Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi,” *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 327, 2021, doi: 10.30865/mib.v5i1.2763.
- [13] M. Muhamad, A. P. Windarto, and S. Suhada, “Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 1–8, 2019, doi: 10.30865/komik.v3i1.1688.
- [14] D. R. Ente, S. A. Thamrin, S. Arifin, H. Kuswanto, and A. Andreza, “Klasifikasi Faktor-Faktor Penyebab Penyakit Diabetes Melitus Di Rumah Sakit Unhas Menggunakan Algoritma C4.5,” *Indones. J. Stat. Its Appl.*, vol. 4, no. 1, pp. 80–88, 2020, doi: 10.29244/ijsa.v4i1.330.
- [15] T. Widiastuti, K. Karsa, and C. Juliane, “Evaluasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Klasifikasi Algoritma C4.5,” *Technomedia J.*, vol. 7, no. 3, pp. 364–380, 2022, doi: 10.33050/tmj.v7i3.1932.
- [16] A. Prasetyo, “Simulasi Penerapan Metode Decision Tree (C4.5) Pada Penentuan Status Gizi Balita,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 4, no. 3, pp. 209–214, 2021, doi: 10.32672/jnkti.v4i3.2983.
- [17] S. Febriani and H. Sulistiani, “Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5,” *89Jurnal Teknol. dan Sist. Inf.*, vol. 2, no. 4, pp. 89–95, 2021.
- [18] S. U. Putri, E. Irawan, and F. Rizky, “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5,” *J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 2, no. 1, pp. 39–46, 2021.
- [19] A. J. Larner, *The 2x2 Matrix: Contingency, Confusion, and the Metrics*. 2021. doi: 10.1007/978-3-030-74920-0.
- [20] A. Mishra, *Machine Learning in the AWS Cloud: Add Intelligence to Applications with Amazon SageMaker and Amazon Rekognition*. wiley, 2019.