

IMPLEMENTASI ALGORITME C4.5 UNTUK KLASIFIKASI KELAYAKAN KENAIKAN KELAS PADA SISWA MENENGAH PERTAMA BERBASIS WEB

Wiwi Risnawati^{1*}, Arief Wibowo²

^{1*,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, DKI Jakarta, Indonesia

Email: ^{1*}wiwirisnawati66@gmail.com, ²arief.wibowo@budiluhur.ac.id

(* : corresponding author)

Abstrak- Penentuan kelayakan kenaikan kelas siswa pada saat kenaikan kelas menjadi hal yang sangat penting untuk melakukan pertimbangan pada hasil belajar siswa dengan diringi penilaian lainnya, yang mana hal tersebut dapat membantu baik itu pihak sekolah maupun pihak siswa untuk mengetahui bagaimana keberhasilan dari pembelajaran serta daya serap siswa dalam menguasai masing-masing materi yang telah disampaikan oleh pihak guru selama masa pembelajaran dilakukan. Masalah umum yang kerap dihadapi oleh siswa menengah pertama yaitu rendahnya nilai ujian maupun nilai kepribadian dan juga nilai kehadiran siswa yang mana hal tersebut dapat mempengaruhi penilaian kenaikan kelas masing-masing siswa. Untuk meminimalisir hal tersebut maka pada penelitian kali ini dibuat sebuah sistem untuk mengklasifikasikan kenaikan kelas siswa dengan menggunakan algoritme C4.5. Perolahan hasil yang didapat pada penelitian kali ini yaitu nilai dari akurasi sebesar 98,246%, nilai presisi sebesar 98,246% dan *recall* sebesar 100% dengan prosentase data 80% data latihan dan 20% data uji.

Kata Kunci: klasifikasi, kenaikan kelas, *decision tree* C4.5

APPLICATION OF C4.5 ALGORITHM FOR WEB-BASED CLASIFICATION GRADE PROMOTION OF HIGH SCHOOL STUDENT

Abstract- Determining the eligibility of students at the time of class promotion is an important thing to do, which can help both of the school and the students to find out how the success of learning and how deep the students in mastering each knowledge that has been taught during the learning period. Common problems that are often faced by junior high school students are the low test scores and personality scores as well as student attendance scores which can affect the assessment of each student's grade increase. To minimize this, in this study a system was created to classify student grade increases using the C4.5 algorithm. The results obtained from this study are an accuracy value of 98.246%, a precision value of 98.246% and a recall of 100% with a percentage of 80% training data and 20% test data.

Keywords: classification, student grade promotion, *decision tree* C4.5

1. PENDAHULUAN

Penentuan kelayakan siswa pada saat kenaikan merupakan hal yang sangat penting untuk dilakukan dalam dunia pendidikan, yang mana hal tersebut dapat membantu baik itu pihak sekolah maupun pihak siswa untuk mengetahui bagaimana keberhasilan dari pembelajaran serta daya serap siswa dalam menguasai masing-masing materi yang telah disampaikan oleh pihak guru selama masa pembelajaran dilakukan. Sehingga nantinya dapat diadakan evaluasi tentang bagaimana cara pembelajaran yang lebih baik atau lebih tepat agar menumbuhkan kembali minat belajar siswa oleh masing-masing pengajar juga untuk meningkatkan penyerapan materi yang telah disampaikan kepada siswa serta menjadikan proses pembelajaran yang lebih efektif bagi siswa.

Penelitian ini juga dilakukan karena adanya kebutuhan akan sistem terkait masalah umum yang kerap dihadapi oleh siswa menengah pertama yaitu rendahnya nilai ujian maupun nilai kepribadian dan juga nilai kehadiran siswa yang mana hal tersebut dapat mempengaruhi penilaian kenaikan kelas masing-masing siswa. Untuk meminimalisir hal tersebut maka penelitian kali ini membuat sebuah system untuk mengklasifikasikan kenaikan kelas siswa dengan menggunakan algoritme C4.5 yang mana nantinya dapat membantu pengajar dalam proses pengambilan keputusan kenaikan kelas siswa yang kemudian pada tahap berikutnya pihak dewan guru

dapat melakukan perundingan dalam mempertimbangkan hasil dari prediksi sistem dengan siswa-siswa yang mendapatkan hasil prediksi tidak naik kelas.

Penggunaan metode *decision tree* dengan algoritme C4.5 dalam penelitian ini karea adanya keunggulan yang terletak pada kemampuannya yang dapat memprediksi dengan hasil akurasi yang cukup ideal untuk diterima dan juga pohon keputusan yang dihasilkan cukup mudah untuk dipahami. Selain itu dalam klasifikasi algoritme *decision tree* C4.5 juga tergolong memiliki tingkat akurasi nilai yang tinggi [1], maka dalam penelitian ini untuk menentukan kelayakan apakah siswa tersebut layak atau tidak layak untuk naik ke tingkat berikutnya penulis memilih menggunakan algoritme C4.5.

Hal yang ingin dicapai dalam penelitian yang dilakukan ini adalah agar dapat menndapatkan prediksi kelayakan siswa untuk kenaikan kelas pada siswa menengah pertama dengan menggunakan metode C4.5 pada aplikasi berbasis web dan juga untuk menerapkan ilmu pengembangan *data mining* terutama pada prediksi kelayakan kenaikan kelas dengan menggunakan algoritme C4.5.

Data mining termasuk kedalam tahapan dari bagian proses *Knowledge Discovery in Database* (KDD). Hal yang dapat dilakukan dengan *data mining* meliputi klasifikasi, prediksi, perkiraan juga memperoleh informasi yang bermanfaat lainnya. Algoritme C4.5 dapat digunakan juga untuk klasifikasi dalam *data mining* disertai dengan phon keputusan dari algoritme C4.5 yang mudah untuk dipaham. [2]. Juga, *data mining* merupakan suatu proses penggabungan antara teknik pencarian pola penting data dan teknik analisis data [3].

CRISP- DM merupakan metode yang digunakan pada penelitian kali ini untuk pemrosesan data. CRISP-DM itu sendiri merupakan salah satu dari beberapa proses pengembangan data yang digunakan pada tahap pemecahan masalah oleh banyak ahli[4].

Pada tahap pengujian model menggunakan metode *confusion matrix* yang mana merupakan pengujian prosentase yang sangat berguna dan juga komprehensif pada kinerja pengklasifikasian[5]. Tahapan dari pengujian yang dilakukan pada penelitian ini yaitu menggunakan metode pengujian *black box* yang mana merupakan sebuah metode dari beberapa metode pengujian yang ada untuk meneliti fungsi yang berjalan pada perangkat lunak tanpa adanya penglihatan pada kinerja aplikasi atau struktur internal[6]. Pengujian *black box* juga dikenal sebagai pengujian *box* yang bersih, *box* pengujian transparan, dan pengujian terstruktur[7].

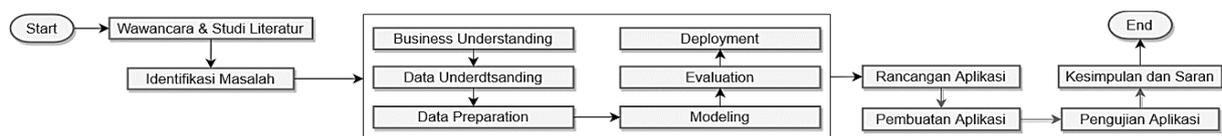
Pada penelitian sebelumnya dilakukan pengembangan untuk prediksi kenaikan kelas siswa menggunakan metode *Analytical Hierarchy Process*[8]. Penelitian selanjutnya juga melakukan penentuan kenaikan kelas siswa menengah atas dengan melakukan pengelompokan nilai akademik menggunakan metode algoritme K-Medoids[9]. Pada penelitian lainnya menggunakan metode TOPSIS untuk melakukan prediksi kenaikan kelas santri dengan menggunakan 42 elemen alternatif juga menggunakan elemen kriteria berjumlah 6 kriteria[10].

Acuan perbedaan yang dapat dilihat pada penelitian kali ini dengan penelitian sebelumnya yaitu penelitian ini menggunakan algoritme C4.5 untuk mengklasifikasikan kelayakan kenaikan kelas siswa tingkat menengah pertama dengan tingkat akurasi dan presisi pada angka 98,246%, dan untuk *recall* didapat nilai sebesar 100%. Penerapan dilakukan dengan membagi data dalam 2 golongan yaitu NAIK dan TIDAK NAIK dengan jumlah keseluruhan data ialah 283 *rows* data

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Pemilihan metode yang digunakan pada tahapan penelitian untuk mengolah data dan menganalisa data ini diputuskan dengan menggunakan metodologi CRISP-DM. Pada tahap ini terbagi kedalam beberapa tahap yaitu seperti yang tertera pada Gambar.



Gambar 1. Tahapan Penelitian

2.2 Wawancara dan Studi Literatur

Studi kepustakaan dan wawancara dilakukan dalam proses pengumpulan data dengan mencari jurnal-jurnal terkait dengan topik penelitian yang diangkat dan wawancara kepada pihak terkait mengenai data kenaikan kelas siswa.

2.3 Identifikasi Masalah

Pada tahap identifikasi masalah ini penulis mengidentifikasi masalah yang berkaitan dengan kelayakan kenaikan kelas siswa berdasarkan acuan yang dibuat dari tahap studi literatur dan tahap wawancara yang telah dilakukan sehingga penelitian dapat berfokus pada pokok permasalahan.

2.4 Analisa

Analisa masalah menjadi hal yang begitu penting untuk dilakukan dalam suatu proses penelitian. Pada tahap ini analisa masalah dilakukan dengan mengacu pada sub tahap dari penerapan metodologi CRISP-DM dimana adanya analisa masalah dilakukan karena adanya kebutuhan akan sistem yang membantu pengajar dalam proses pengambilan keputusan kenaikan kelas siswa. Penjelasan dari setiap tahapan dapat dilihat pada penjabaran dibawah ini.

a. *Business Understanding*

Tahap *business understanding* adalah tahap dimana memahami alur dari proses yang berjalan pada tiap kegiatan yang dilakukan pada penambangan data. Pokok permasalahan yang didapat pada penelitian ini adalah bagaimana mengklasifikasi kelayakan kenaikan kelas siswa. Sehingga penelitian ini memiliki fokus pada data dari suatu kelas atau kondisi kelayakan kenaikan kelas siswa.

b. *Data Understanding*

Proses *data understanding* pertama dilakukan dengan melakukan pengumpulan data kenaikan kelas siswa yang berfokus pada data dari nilai siswa menengah pertama. Setelah mendapatkan data kemudian data dipahami dan dianalisa sehingga mendapatkan beberapa atribut yang dibutuhkan dalam penentuan kelayakan kenaikan kelas siswa dengan jumlah data awal yang didapatkan adalah 283 total data.

c. *Data preparation*

Data preparation memiliki beberapa tahapan *preprocessing* pada data yang telah didapat. Proses *preprocessing* dilakukan secara manual dengan menggunakan Ms.excel dengan proses *data cleaning*, *data transformation* dan *data testing*.

d. *Modelling*

Tahap *modeling* dilakukan dengan melakukan pembagian pada *dataset* menjadi *data testing* dan *data training* kedalam prosentase 80% *data training* dan 20% *data testing*. Dengan keseluruhan jumlah data yang digunakan dalam penelitian ini adalah 283 *rows* data.

e. *Evaluation*

Tahap evaluasi dilakukan dengan penerapan metode *confusion matrix* sehingga mendapat prosentase dari nilai akurasi presisi dan *recall* yang kemudian dari nilai tersebut data dan tahapan dapat dievaluasi Kembali.

f. *Deployment*

Tahap ini merupakan tahap implementasi dari seluruh proses analisa yang telah dilakukan kemudian diterapkan pada sistem dan juga pembuatan laporan.

2.5 Rancangan Aplikasi

Pada tahap ini dilakukan rancangan aplikasi yang akan dibangun dengan penyesuaian terhadap hasil analisa model yang telah dilakukan.

2.6 Pembuatan Aplikasi

Pembuatan aplikasi merupakan tahap pengimplementasian dari analisa dan perancangan yang telah dilakukan sehingga terbangunlah aplikasi yang dibutuhkan

2.7 Pengujian Aplikasi

Metode pengujian *black box* dipilih untuk pengujian pada tahap testing dengan bertujuan agar pengujian ini dapat mengetahui bahwa aplikasi yang dibangun telah melewati masa uji coba sehingga dapat berjalan dengan semestinya.

2.8 Kesimpulan dan Saran

Setelah dilakukannya seluruh tahapan yang ada maka ditraiklah kesimpulan dari hasil yang didapat dan juga saran yang dapat diambil untuk penelitian ini.

2.9 Data Transformation

Proses *data transformation* bertujuan untuk mengelompokkan data kedalam beberapa kategori data. Berikut merupakan hasil dari keseluruhan pengelompokan data dari masing-masing atribut.

Tabel 1. Kategori Atribut Transformasi Data

Variabel	Nilai	Keterangan
pengetahuan	A,B,C,D	Atribut untuk menampung nilai pengetahuan siswa
keterampilan	A,B,C,D	Atribut untuk menampung nilai keterampilan siswa
jumlah_alfa	TINGGI, SEDANG, RENDAH	Atribut yang digunakan untuk jumlah ketidakhadiran siswa tanpa keterangan (alfa)
eskul	SANGAT BAIK, BAIK, KURANG BAIK	Atribut untuk menampung nilai ekstrakurikuler siswa
kepribadian	SANGAT BAIK, BAIK, KURANG BAIK	Atribut yang digunakan untuk nilai kepribadian siswa
her	RENDAH, TINGGI	Atribut yang digunakan untuk nilai dari jumlah mata pelajaran tidak lulus

Pengelompokan untuk nilai keterampilan dan nilai pengetahuan dikategorikan kedalam 4 kategori seperti yang tertera pada table 2.

Tabel 2. Kategori Nilai Keterampilan dan Pengetahuan

Variable	Nilai
100-91	A
90-81	B
80-71	C
>70	D

2.10 Algoritme C4.5

Algoritme C4.5 merupakan salah satu algoritme yang digunakan dalam *data mining* yang berguna untuk membantu dalam pengklasifikasian kelas yang mana algoritme C4.5 adalah algoritme yang dikembangkan dari algoritme ID3 dengan cara kerja untuk menghasilkan keputusan adalah dengan membuat pohon keputusan. Setelah menyiapkan pemilihan atribut yang dapat dihitung dari data latih dalam algoritme C4.5 maka digunakan konsep *entropy*. *Entropy* digunakan untuk menghitung nilai *imurty* dari sekumpulan objek dengan rumus perhitungan seperti yang tertera pada lampiran (1) berikut ini.

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

Keterangan :

S = Himpunan Kasus

n = Jumlah partisi S

p_i = nilai dari probabilitas yang mana dihitung berdasarkan jumlah kelas yang dibagi dengan total kasus yang ada.

Berikutnya setelah nilai *entropy* didapat kemudian dilakukan perhitungan *information gain* yang digunakan untuk pemilihan atribut. Berikut merupakan lampiran (2) yang berisikan rumus dari menghitung nilai *information gain*.

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

Keterangan :

S = Himpunan kasus

A = Atribut

n = Jumlah atribut

|S_i| = Jumlah partisi ke -i

$|S|$ = jumlah kasus dalam S

Setelah didapat nilai dari *entropy* dan *information gain* diatas jika terdapat atribut yang memiliki banyak nilai dari atribut maka hitungan *split information* ini diperlukan. Berikut merupakan lampiran (3) yang berisikan rumus untuk menghitung nilai *split information*.

$$SplitInfo(S, A) = \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (3)$$

Keterangan :

S = ruang (data) sampel yang digunakan untuk *training*.

A = atribut.

S_i = jumlah sampel untuk atribut i.

Kemudian setelah mendapatkan nilai *info gain* dan *split info* maka berikutnya adalah mencari nilai *gain ratio* dengan menggunakan rumus pada lampiran (4) dibawah ini.

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

2.11 Confusion Matrix

Dalam perhitungan tingkat akurasi dalam sebuah sistem *data mining* maka metode *confusion matrix* merupakan salah satu caranya. Berikut merupakan *table confusion matrix*. Tahap selanjutnya setelah memasukan data uji kedalam *confusion matrix* yaitu menghitung nilai yang sudah dimasukan tersebut agar dapat dihitung nilai dari akurasi, presisi dan *recall*. Perhitungan nilai dari akurasi, presisi dan *recall* dapat dihitung melalui rumus berikut ini:

Tabel 3. Confusion Matrix

Kategori	Positive(1)	Negative(0)
Positive(1)	TP	FP
Negative(0)	FN	TN

- a. Nilai akurasi merupakan nilai yang digunakan untuk dapat menghitung keauratan sistem dalam mengklasifikasikan data. Pada lampiran (5) berikut ini merupakan rumus untuk menghitung akurasi.

$$Akurasi = \frac{TP + TN}{TP + FN + FP + TN} * 100\% \quad (5)$$

- b. Nilai presisi merupakan nilai yang digunakan untuk menghitung banyaknya nilai *positive* atau kebenaran yang terdapat pada *dataset* yang memang sebenarnya memiliki nilai *positive* atau benar yang muncul. Rumus perhitungan dari nilai presisi dapat dilihat pada lampiran (6) berikut.

$$Presisi = \frac{TP}{TP + FP} * 100\% \quad (6)$$

- c. Nilai *recall* merupakan nilai yang digunakan dalam menghitung nilai benar dari klasifikasi yang sesuai pada nilai benar yang *real*. Pada lampiran (7) berikut ini merupakan rumus perhitungan untuk mencari nilai *recall*.

$$Recall = \frac{TP}{TP + FN} * 100\% \quad (7)$$

3. HASIL DAN PEMBAHASAN

3.1 Perhitungan C4.5

Proses pembentukan *decision tree* dengan algoritme C4.5 dimulai dengan dilakukannya perhitungan *entropy* total kemudian setelah nilai tersebut didapat maka dihitunglah *entropy* dari setiap atribut yang ada. Kemudian tahap selanjutnya setelah menghitung nilai *entropy* yang diperoleh dari setiap atribut maka hal berikutnya yang dilakukan adalah menghitung nilai *info gain*, *split info* dan *ratio gain* dari setiap atribut.

Perhitungan *entropy* dari masing-masing atribut dan *entropy* total dihitung berdasarkan rumus perhitungan yang terdapat pada lampiran (1)

Berikut merupakan perhitungan *entropy* total dengan jumlah *data training* sebanyak 226 *rows* data dengan jumlah data berlabel Naik adalah 220 data kemudian untuk jumlah data dengan label TIDAK NAIK berjumlah 6 data.

$$\begin{aligned} \text{Entropy total [277,6]} &= -\frac{220}{283} \log_2 \left(\frac{220}{283} \right) - \frac{6}{283} \log_2 \left(\frac{6}{283} \right) \\ &= 1,5624 \end{aligned}$$

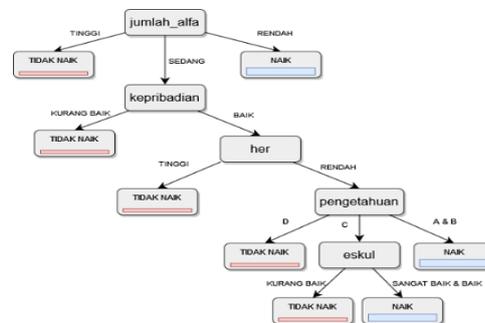
Setelah didapat nilai *entropy* total selanjutnya hal yang harus dilakukan yaitu melakukan perhitungan untuk mendapatkan nilai *entropy* dari masing-masing atribut, dimana dalam perhitungannya masih menggunakan perhitungan yang sama dengan perhitungan *entropy total* yaitu rumus pada lampiran (1), menghitung nilai *info gain* berdasarkan rumus pada lampiran (2), menghitung nilai *split info* berdasarkan lampiran (3) dan *gain ratio* berdasarkan rumus pada lampiran (4). Berikut ini merupakan detail penjabaran dari masing-masing perhitungan. Berikut merupakan tabel hasil dari seluruh perhitungan.

Tabel 4. Hasil Perhitungan

Atribut	Jumlah	Naik	Tidak Naik	Entropy	Info Gain	Split Info	Gain Ratio
PENGETAHUAN					1,4355	0,9220	1,5569
A	1	1	0	0			
B	137	134	3	0,1520			
C	86	85	1	0,0914			
D	2	0	2	0			
KETERAMPILAN					1,4047	0	0
A	0	0	0	0			
B	217	213	4	0,1326			
C	9	7	2	0,7642			
D	0	0	0	0			
HER					1,4343	0,0731	19,6298
RENDAH	224	220	4	0,1292			
TINGGI	2	0	2	0			
ALFA					1,4612	0,2298	6,3594
TINGGI	3	0	3	0			
SEDANG	4	4	0	0			
RENDAH	219	216	3	0,1044			
KEPRIBADIAN					1,4345	0	0
SANGAT BAIK	2	2	0	0			
BAIK	222	218	4	0,1302			
KURANG BAIK	2	0	2	0			
ESKUL					1,4099	0,5929	2,3780
SANGAT BAIK	29	28	1	0,2164			
BAIK	196	192	4	0,1437			
KURANG BAIK	1	0	1	0			

Berikut merupakan pohon keputusan yang dibangun berdasarkan perhitungan yang telah dilakukan. Tabel dibentuk berdasarkan 5 atribut dengan nilai *information gain* tertinggi.

Setelah perhitungan diatas maka didapatlah pohon keputusan yang mana, jika nilai dari atribut jumlah_alfa adalah SEDANG maka keputusan yang dihasilkan adalah naik. Apabila nilai dari atribut jumlah_alfa TINGGI maka keputusan yang dihasilkan adalah TIDAK NAIK. Apabila nilai dari jumlah_alfa adalah SEDANG maka akan dilihat nilai dari atribut berikutnya yaitu atribut kepribadian. Apabila nilai dari atribut kepribadian adalah KURANG BAIK maka keputusan yang dihasilkan adalah TIDAK NAIK. Namun, jika nilai dari atribut kepribadian adalah BAIK maka akan dilanjutkan pada pengecekan atribut berikutnya yaitu atribut her. Jika atribut her bernilai TINGGI maka keputusan yang dihasilkan adalah TIDAK NAIK. Namun, jika nilainya adalah RENDAH maka akan dilanjutkan pada pengecekan nilai atribut berikutnya yaitu atribut pengetahuan. Jika nilai dari atribut pengetahuan adalah D maka keputusan yang dihasilkan adalah TIDAK NAIK dan jika nilai yang didapat adalah A dan atau B maka keputusan yang dihasilkan adalah NAIK. Namun, apabila nilai yang didapat adalah C, maka akan dilakukan pengecekan atribut berikutnya yaitu atribut eskul. Jika nilai dari atribut eskul adalah SANGAT BAIK atau BAIK maka hasil final dari keputusan yang didapat adalah NAIK. Namun, apabila nilai dari atribut yang didapat adalah KURANG BAIK maka hasil final dari keputusan yang didapat adalah TIDAK NAIK. Diagram yang terkait dengan penjelasan pohon keputusan diatas dapat dilihat pada Gambar 2 berikut ini:



Gambar 2. Pohon Keputusan

Dari gambar tersebut didapatlah *rules* pengujian untuk menentukan klasifikasi kelayakan kenaikan kelas yaitu :

```
JUMLAH_ALFA = TINGGI : TIDAK NAIK
JUMLAH_ALFA = RENDAH
|   KEPRIBADIAN = BAIK
|   |   HER = RENDAH
|   |   |   PENGETAHUAN = A: NAIK
|   |   |   PENGETAHUAN = B: NAIK
|   |   |   PENGETAHUAN = C
|   |   |   |   ESKUL = BAIK: NAIK
|   |   |   |   ESKUL = KURANG BAIK: TIDAK NAIK
|   |   |   |   ESKUL = SANGAT BAIK: NAIK
|   |   |   |   PENGETAHUAN = D: TIDAK NAIK
|   |   |   |   HER = TINGGI: TIDAK NAIK
|   |   |   |   KEPRIBADIAN = KURANG BAIK: TIDAK NAIK
JUMLAH_ALFA = SEDANG : NAIK
```

3.2 Perhitungan *Confusion Matrix*

Proses perhitungan *confusion matrix* yang dilakukan bertujuan untuk mencari nilai dari akurasi, presisi dan *recall* pada data uji dimana masing-masing kriteria dari setiap data akan dibandingkan yang kemudian akan dimasukkan kedalam *system* setelah perhitungan tersebut selesai. Berikut merupakan tabel dan penjabaran dari perhitungan *confusion matrix*.

Tabel 5. Perhitungan *Confusion Matrix*

Kategori	NAIK	TIDAK NAIK
NAIK	TP (56)	FP (1)
TIDAK NAIK	FN (0)	TN (0)

Berikut merupakan nilai akurasi, presisi dan *recall* yang didapat dari *table confusion matrix* diatas.

a. Perhitungan akurasi

$$\text{Akurasi} = \frac{TP+TN}{TP+FN+FP+TN} * 100 \%$$

$$= \frac{56+0}{56+0+1+0} * 100 \%$$

$$= \frac{56}{57} * 100 \% = 98,246 \%$$

Nilai dari tingkat akurasi yang didapat berdasarkan perhitungan diatas adalah sebesar 98,246% .

b. Perhitungan presisi

$$\text{Presisi} = \frac{TP}{TP+FP} * 100 \%$$

$$= \frac{56}{56+1} * 100 \%$$

$$= \frac{56}{57} * 100 \%$$

$$= 98,246 \%$$

Nilai dari perhitungan presisi yang didapat berdasarkan perhitungan diatas adalah sebesar 98,246% .

c. Perhitungan recall

$$\text{Recall} = \frac{TP}{TP+FN} * 100 \%$$

$$= \frac{56}{56+0} * 100 \%$$

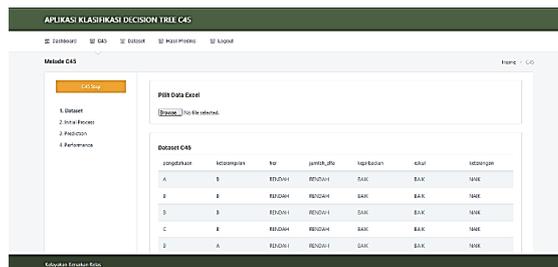
$$= \frac{56}{56} * 100 \%$$

$$= 100\%$$

Nilai dari perhitungan recall yang didapat berdasarkan perhitungan diatas adalah sebesar 100%.

3.3 Implementasi

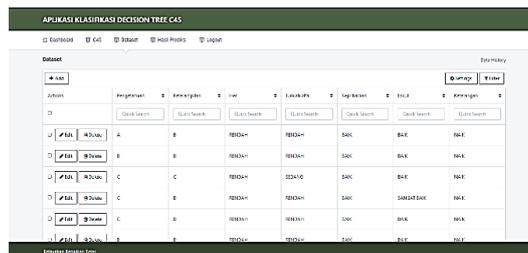
Halaman C4.5 merupakan halaman yang berisikan proses data dengan Algoritme C4.5 dimana pada halaman ini terdapat beberapa submenu untuk mengolah *dataset* dan menguji tingkat akurasi dari data tersebut. Pada halaman ini langsung menampilkan submenu *dataset* yang mana pada submenu *dataset* ini terdapat fungsi *upload file excel dataset* yang kemudian ditampilkan dalam bentuk *table*.



pengukuran	kelembapan	tar	jumlah_uji	kegagalan	kelas	hasil
A	B	RE2D41	RE2D41	SAK	BAK	NIK
B	B	RE2D41	RE2D41	SAK	BAK	NIK
B	B	RE2D41	RE2D41	SAK	BAK	NIK
C	B	RE2D41	RE2D41	SAK	BAK	NIK
D	A	RE2D41	RE2D41	SAK	BAK	NIK

Gambar 3. Halaman C45

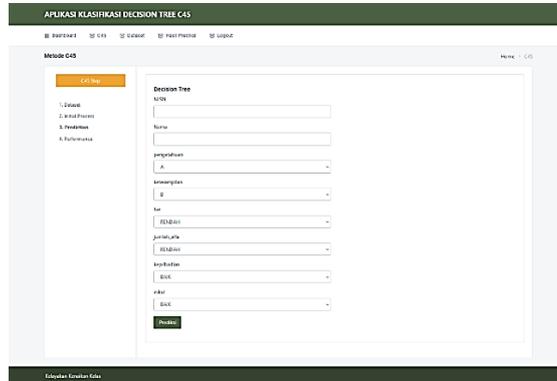
Tampilan dari salah satu halaman yaitu layar halaman *dataset* yang mana halaman ini difungsikan untuk menampilkan hasil *import* data dengan format *.xlsx* yang kemudian dimunculkan ke tampilan dalam bentuk tabel.



Jumlah	Kelembapan	Tar	Jumlah Uji	Kegagalan	Kelas	Hasil
A	B	RE2D41	RE2D41	SAK	BAK	NIK
B	B	RE2D41	RE2D41	SAK	BAK	NIK
C	C	RE2D41	RE2D41	SAK	BAK	NIK
C	B	RE2D41	RE2D41	SAK	SAK2D41SAK	NIK
C	B	RE2D41	RE2D41	SAK	BAK	NIK
D	B	RE2D41	RE2D41	SAK	BAK	NIK

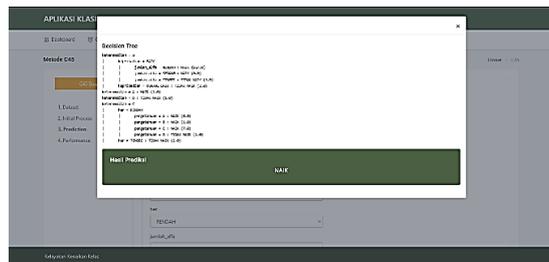
Gambar 4. Menu Dataset

Terdapat *input* data yang terjadi pada halaman prediksi yang kemudian akan ditampilkan hasil prediksi dari data yang telah dimasukkan tersebut apakah hasil prediksinya NAIK atau TIDAK NAIK.



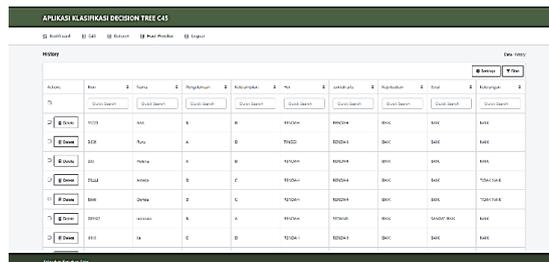
Gambar 5. Submenu *Prediction*

Berikut ini merupakan halaman yang ditampilkan setelah mengklik *button* prediksi, seperti yang dapat dilihat pada tampilan prediksi yang ditampilkan seperti *popup* atau *alert*.



Gambar 6. Hasil *Input* Prediksi

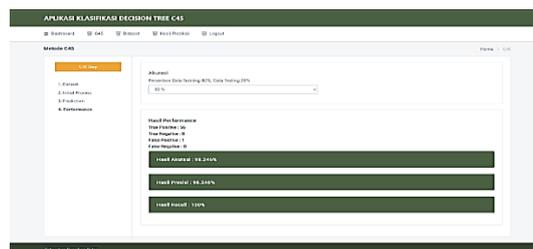
Tampilan layar halaman menu hasil prediksi merupakan halaman yang menampilkan data dari hasil prediksi pada submenu prediksi yang kemudian ditampilkan kedalam bentuk tabel.



Aktiva	Bar	S	Nama	Penghasilan	Kewarganegaraan	No	JumlahAnak	ApakahMantan	JenisKelamin	Status
75	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
8	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
20	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
22	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
23	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
24	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
25	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
26	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
27	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
28	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
29	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah
30	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah	Daerah

Gambar 7. Menu Hasil Prediksi

Halaman *performance* ini berfungsi untuk melakukan perhitungan akurasi, presisi dan *recall* yang kemudian pada halaman ini juga ditampilkan hasil dari proses perhitungan akurasi, presisi dan *recall* dari data sesuai dengan prosentase data yang dipilih.



Gambar 8. Submenu *Performance*

Pengujian *performance* dilakukan dengan menggunakan metode *confusion matrix* yang menggunakan jumlah data uji sebanyak 57 data yang kemudian dari data tersebut didapatlah nilai *True Positive* (TP) sebanyak 56 data, *True Negative* sebanyak 0 data, *False Positive* sebanyak 1 data dan *False Negative* sebanyak 0 data. Data tersebut dihitung berdasarkan hitungan pada poin sebelumnya yaitu poin 3.2 Perhitungan *Confusion Matrix*. Kemudian menghasilkan nilai akurasi sebesar 98,246%. Detail dari hasil perhitungan nilai *performance* dapat dilihat pada Tabel 6 berikut ini:

Tabel 6. Nilai Hasil Performance

Akurasi	Presisi	Recall
98,246%	98,246%	100%

4. KESIMPULAN

Berlandaskan pada penelitian yang dilakukan ini, kemudian sampai lah pada penarikan kesimpulan yang mana berisikan bahwa sistem berbasis web yang telah dibangun ini dapat menentukan kelayakan kenaikan kelas siswa menengah pertama dengan algoritme *decision tree* C4.5 yang dipilih sebagai metode yang digunakan untuk pemecahan masalah pada penelitian ini. Pada kasusnya, banyaknya data juga menentukan keberagaman nilai dari akurasi yang akan diperoleh. Pada penelitian ini dengan menggunakan data berjumlah 283 *rows* dengan prosentase *data training* dan *data testing* sebesar 80% dan 20% maka diperoleh nilai akurasi sebesar 98,246%, nilai presisi sebesar 98,246% dan juga *recall* sebesar 100%. Dengan adanya penelitian ini yang menghasilkan aplikasi klasifikasi kenaikan kelas berbasis web, diharapkan dapat memudahkan pada saat penentuan kenaikan kelas siswa pada tingkat menengah pertama.

DAFTAR PUSTAKA

- [1] M. B. Hanif, H. Arga, D. Rani, A. Rifai, and G. Guntoro, "Komparasi Metode Naive Bayes dan C4 . 5 Pada Klasifikasi Persalinan Prematur," *Joined J.*, vol. 5, no. 1, pp. 54–65, 2022.
- [2] Yuli Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [3] I. A. Shah and S. Rajper, "Using ML and Data-Mining Techniques in Automatic Vulnerability Software Discovery," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 3, pp. 2109–2126, 2021, doi: 10.30534/ijatcse/2021/871032021.
- [4] D. Astuti, "Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 60–72, 2019, doi: 10.20895/inista.v1i2.71.
- [5] D. Krstinić, M. Braović, L. Šerić, and D. Božić-Štulić, "Multi-label Classifier Performance Evaluation with Confusion Matrix," pp. 01–14, 2020, doi: 10.5121/csit.2020.100801.
- [6] A. Fahrezi, F. N. Salam, G. M. Ibrahim, R. R. Syaiful, and A. Saifudin, "Pengujian Black Box Testing pada Aplikasi Inventori Barang Berbasis Web di PT. AINO Indonesia," *J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 1, pp. 1–5, 2022, [Online]. Available: <https://journal.mediapublikasi.id/index.php/logic>
- [7] S. R. Yulistina, T. Nurmala, R. M. A. T. Supriawan, S. H. I. Juni, and A. Saifudin, "Penerapan Teknik Boundary Value Analysis untuk Pengujian Aplikasi Penjualan Menggunakan Metode Black Box Testing," *J. Inform. Univ. Pamulang*, vol. 5, no. 2, p. 129, 2020, doi: 10.32493/informatika.v5i2.5366.
- [8] A. Aminah and M. N. Dasaprawira, "Pengembangan Sistem Pendukung Keputusan Dengan Metode Ahp Di Sekolah.," *Joinics*, vol. 1, no. 2, pp. 21–27, 2021, [Online]. Available: <https://jurnal.unugba.ac.id/index.php/jnc/article/view/265>
- [9] S. Silitonga, E. Irawan, S. Saifullah, M. R. Lubis, and I. Parlina, "Pengelompokan Nilai Akademik untuk Menentukan Kenaikan Kelas Menggunakan Algoritma K-Medoids," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 740, 2019, doi: 10.30645/senaris.v1i0.80.
- [10] B. Aziz and Andrianingsih, "Sistem Pendukung Keputusan Penentuan Home industry Menggunakan Metode Topsis," *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 3, pp. 470–478, 2022, doi: 10.35870/jtik.v6i3.529.