

# KLASIFIKASI HASIL BELAJAR SISWA MENGUNAKAN METODE C4.5 BERDASARKAN RIWAYAT AKADEMIK DI SMP XYZ

Bryan Daniel Pinenda Pasaribu <sup>1)</sup> Tri Sutrisno <sup>2)</sup> Bagus Mulyawan <sup>3)</sup>

<sup>1)</sup> Teknik Informatika, FTI, Universitas Tarumanagara  
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia  
email : bryandanielp2332@gmail.com

<sup>2)</sup> Teknik Informatika, FTI, Universitas Tarumanagara  
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia  
email : tris@fti.untar.ac.id

<sup>3)</sup> Teknik Informatika, FTI, Universitas Tarumanagara  
Jl. Letjen S Parman no 1, Jakarta 11440, Indonesia  
email : bagus@fti.untar.ac.id

## ABSTRACT

*This research was conducted to classify student learning outcomes at XYZ Middle School based on academic history during learning and student learning interests. This study aims to provide information to students and teaching staff regarding student learning outcomes. With the available information, it is hoped that teaching staff can develop methods for conveying material in order to obtain better results. The method used in solving this problem is the C4.5 algorithm method. Starting from collecting data consisting of assignment scores, daily tests, UTS, and UAS. Then the formation of a decision system as initial data that has condition and decision attribute values. Then calculate the entropy value of each attribute. Calculating the highest gain value which will then be used as a node. Then, determine the decision from the results of the decision tree process by starting from the highest root to the lowest root to determine the decision criteria.*

## Key words

*Algoritma C4.5, Learning History, Outcomes*

## 1. Pendahuluan

Hasil belajar siswa merupakan salah satu faktor yang cukup berpengaruh untuk mengukur kualitas belajar. Belajar dikatakan berhasil jika dapat menumbuhkan sikap, tingkah laku dan cara berfikir dalam memecahkan permasalahan - permasalahan yang dihadapi.[1]

Penelitian ini bertujuan untuk membuat aplikasi berbasis web untuk mengklasifikasikan hasil belajar siswa berdasarkan riwayat akademik yang terdiri dari nilai tugas, ulangan harian, UTS dan UAS dari setiap mata pelajaran, serta mempermudah tenaga pengajar untuk memantau perkembangan hasil belajar siswa dan memberikan informasi kepada siswa untuk lebih

meningkatkan kualitas belajar

Metode yang digunakan dalam penelitian ini adalah Algoritma C4.5. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Decision Tree). Pohon keputusan berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Dengan menggunakan metode ini, data yang tercatat akan diproses untuk mendapatkan pola berupa *rule* yang akan menjadi landasan dalam melakukan prediksi hasil belajar siswa.

Atribut dari data yang akan dikelola meliputi riwayat akademik, yang meliputi data nilai Tugas, Ulangan Harian, UTS, dan UAS. Berdasarkan uraian di atas, penelitian ini menggunakan metode data mining C4.5 untuk mengklasifikasikan hasil belajar siswa dengan memanfaatkan data yang telah dimiliki.

## 2. Metode Perancangan

### 2.1 Algoritma C4.5

Algoritma C4.5 adalah rangkaian algoritma untuk masalah klasifikasi dalam pembelajaran mesin dan penggalian data. Algoritma C4.5 merupakan pengembangan dari algoritma ID3. Algoritma C4.5 dan ID3 diciptakan oleh seorang peneliti dibidang kecerdasan buatan bernama J. Rose Quinlan pada akhir tahun 1970-an. Algoritma C4.5 membuat pohon keputusan dari atas ke bawah, dimana atribut paling atas merupakan akar, dan paling bawah dinamakan daun.[2] Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses pada setiap cabang [3]

Perhitungan rumus untuk mencari nilai gain adalah sebagai berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=0}^n Entropy(S_i) \dots\dots(1)$$

Keterangan:

S : himpunan kasus  
A : atribut  
N : jumlah partisi atribut A  
[Si] : jumlah kasus pada partisi ke-i  
[S] : jumlah kasus pada S

Untuk mencari nilai entropy menggunakan rumus sebagai berikut

$$Entropy(S) = \sum_{i=0}^n -p_i * \log_2 p_i \dots\dots(2)$$

Keterangan:

S : himpunan kasus  
n : jumlah partisi S  
pi : proporsi Si terhadap S [4]

Berikut merupakan hasil perhitungan Entropy dan Gain dari seluruh atribut

$$(-(\frac{8}{12}) \times \log_2(\frac{8}{12})) + (-(\frac{4}{12}) \times \log_2(\frac{4}{12})) = 0.918295834$$

## 2.2 Data Siswa

Dataset yang akan digunakan pada rancangan ini adalah data histori nilai akademik yang terdiri dari nilai tugas, ulangan harian, UTS dan UAS dari setiap mata pelajaran. Data diperoleh dengan cara meminta izin dari sekolah tempat penelitian. Data yang diperoleh kemudian di kategorisasikan berdasarkan bobot nilai yang berlaku. Contoh data yang digunakan dapat dilihat pada Gambar 1.

Nama Siswa	Kelas	Nomor Induk	Tahun Ajaran	Mata pelajaran	PTS	PA	RPH	A	NRB
ERIKA MAYORI	VII	212207010	2021/2022	Matematika	86	98	90	93	87
FARISI WINDRAYA RAVANDRY	VII	171807005	2021/2022	Matematika	75	92	92	80	81
FEUSA SALSABILAH	VII	212207011	2021/2022	Matematika	90	98	95	78	91
FIANA	VII	171807006	2021/2022	Matematika	65	82	85	75	78
FREDI TIA MARWAN	VII	171807007	2021/2022	Matematika	87	96	95	75	91

Gambar 1. Tampilan Data Siswa

Data yang telah diperoleh diproses melalui aplikasi yang telah dibuat untuk mengkategorisasikan nilai tersebut. Nantinya nilai yang sudah diperoleh, dapat ditampilkan dalam bentuk tabel di aplikasi. Selain itu, pengguna juga dapat dilihat perkembangan nilai dari setiap murid berdasarkan matapelajaran dan kelas dalam bentuk grafik.

## 2.3 Bahasa Pemrograman

Pada pembuatan aplikasi sistem klasifikasi nilai siswa, bahasa pemrograman yang digunakan adalah Python. Library yang digunakan untuk metode C4.5 adalah Sklearn. Library ini berfungsi untuk melakukan proses klasifikasi nilai siswa.

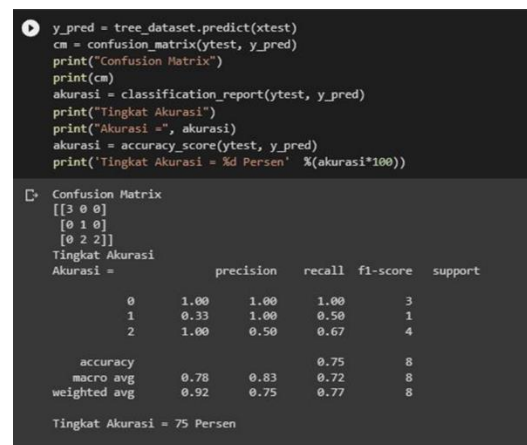
Selanjutnya, data yang telah diklasifikasi ditampilkan dalam bentuk pohon keputusan. Klasifikasi data nilai

siswa dilakukan secara otomatis berdasarkan data yang telah diunggah kedalam aplikasi

## 3. Hasil Percobaan

### 3.1 Pengujian Metode

Pada tahapan pengujian ini bertujuan untuk mengetahui nilai akurasi dari metode C4.5 yang digunakan. Pengujian ini dilakukan dengan melihat tingkat akurasi dan confusion matrix. Pengujian dilakukan menggunakan 4 atribut yaitu Tugas, Ulangan, UTS, dan UAS. Selanjutnya diambil data sampel untuk pengujian sebanyak 40 data. Data tersebut dihitung menggunakan tool Google Colaborator dengan pemrograman python. Tampilan confusion matrix dan tingkat akurasi metode dapat dilihat pada Gambar 2.



Gambar 2 Confusion Matrix

Pengujian metode juga dilakukan dengan menampilkan pohon keputusan untuk proses klasifikasi nilai siswa. Pada penelitian ini, pohon keputusan berperan dalam menentukan kategorisasi nilai siswa. Penentuan kategori nilai ditetapkan berdasarkan bobot yang diberlakukan di sekolah. Rentang nilai dapat dilihat pada Tabel 1.

Tabel 1 Kategori Nilai

Interval Nilai	Predikat	Keterangan
93-100	A	Sangat Baik
84-92	B	Baik
75-83	C	Cukup
<75	D	Kurang

Untuk tampilan dari decision tree dari penelitian ini dapat dilihat pada Gambar 3.

Gambar 3 Decision Tree

### 3.2 Pengujian Tampilan

Pada tahapan pengujian ini bertujuan untuk memastikan apakah aplikasi yang telah dibuat dapat menampilkan hasil yang sesuai dengan tujuan rancangan. Setelah dilakukan percobaan, aplikasi dapat menampilkan kategorisasi nilai secara otomatis. Pengujian aplikasi yang pertama dilakukan adalah memasukan data nilai siswa ke dalam aplikasi. Data yang digunakan berupa data nilai berbentuk excel. Tampilan dari data yang digunakan dapat dilihat pada Gambar 4.

Nama Siswa	Kelas	Nomor Induk	Tahun Ajaran	Mata pelajaran	PTS	PA	RPH	A
ASIAN LAZ MUSA	VI	212207003	2021/2022	Pendidikan Agama	82	94	81	78
ALICIA GISELLE GUNASETIAWAN	VI	212207005	2021/2022	Pendidikan Agama	83	89	96	95
BANGKA VERGE SUEHAL	VI	212207006	2021/2022	Pendidikan Agama	84	83	95	85
BENI AURUS	VI	212207007	2021/2022	Pendidikan Agama	89	89	89	89
CHRISTIAN JAMES GERAR	VI	212207008	2021/2022	Pendidikan Agama	71	84	81	86
CHRISTIAN YOSLIA HUMIRAN PAKPHAN	VI	212207009	2021/2022	Pendidikan Agama	86	87	94	81
ERIKY ANTONY	VI	212207010	2021/2022	Pendidikan Agama	92	70	70	80
ELISA SALSABIH	VI	212207011	2021/2022	Pendidikan Agama	80	90	89	85
GABRIEL JOKRI DAVIDA TIMBULENG	VI	212207012	2021/2022	Pendidikan Agama	72	70	70	85
ITAN ASHA KASHI	VI	212207013	2021/2022	Pendidikan Agama	80	84	85	70
JANIS BASTIARA	VI	212207014	2021/2022	Pendidikan Agama	80	84	85	70
JASON BEN MATTHEW PARUBAK	VI	212207015	2021/2022	Pendidikan Agama	81	73	65	68
JESSICA CAUSTIA	VI	212207016	2021/2022	Pendidikan Agama	88	80	90	76
JOCELYN WILMA LARISSA T	VI	212207017	2021/2022	Pendidikan Agama	87	76	82	80
JUAN CARLO C	VI	212207018	2021/2022	Pendidikan Agama	83	83	83	83
JUVITA PUSPITA SARI	VI	212207019	2021/2022	Pendidikan Agama	88	79	89	76
LINA ANINDYA PUTRI	VI	212207020	2021/2022	Pendidikan Agama	80	87	87	82
LUKY KURNIAWAN	VI	212207021	2021/2022	Pendidikan Agama	79	88	90	85
MARCEL DAUD	VI	212207022	2021/2022	Pendidikan Agama	83	80	80	80
MARVEL ADWIN WIJAYA S	VI	212207023	2021/2022	Pendidikan Agama	83	80	71	70
MATTHEW RODINSKI TANJUNG	VI	212207024	2021/2022	Pendidikan Agama	88	54	72	46
MUHAMMAD NADIA M	VI	212207025	2021/2022	Pendidikan Agama	87	85	85	85
RACHEL WISE SACHA	VI	212207026	2021/2022	Pendidikan Agama	84	91	97	94

Gambar 4. Data Yang Digunakan

### Keterangan

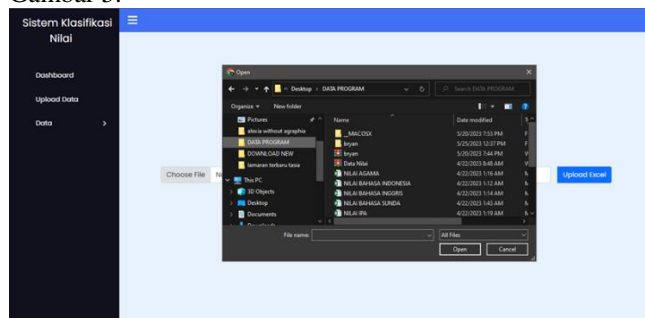
PTS : Nilai Tengah Semester

PA : Nilai Rata-Rata PR dan PS

RPH : Nilai Rata-Rata Akhir Penilaian Harian

A : Penilaian Akhir Semester

Tampilan dari proses upload data dapat dilihat pada Gambar 5.



Gambar 5. Proses Unggah Data

Selanjutnya, data akan yang telah diunggah dapat dilihat pada menu Tabel Data Nilai. Tampilan data nilai yang telah dikategorisasi dapat dilihat pada Gambar 5.

Gambar 6. Tabel Data Nilai

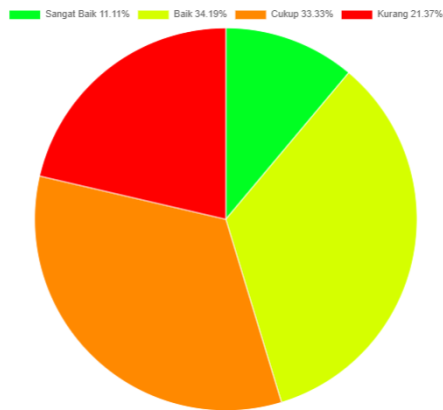
Pada tampilan tabel data nilai, terdapat kolom keterangan yang berisi kategori nilai siswa. Kategori ini merupakan hasil dari klasifikasi menggunakan metode C4.5. Untuk menentukan kategori nilai menggunakan library python sklearn. Hasil metode berupa kategori nilai dapat dilihat pada Gambar 5.



Gambar 7. Kategori Nilai

Pada tabel data nilai, data yang ditampilkan dapat diurutkan berdasarkan mata pelajaran. Selain itu, terdapat menu pencarian untuk mencari data siswa tertentu. Fitur ini dapat mencari data berdasarkan nama dan nomor induk siswa.

Berdasarkan data yang ditampilkan di tabel, jumlah data dari masing masing kategori dapat ditampilkan dalam bentuk diagram lingkaran, sehingga user dapat lebih mudah dalam melakukan perhitungan. Data yang ditampilkan pada diagram lingkaran merupakan data nilai dengan kategori Sangat Baik, Baik, Cukup, dan Kurang. Diagram lingkaran juga menampilkan jumlah data dan persentase dari masing masing kategori nilai. Tampilan dari diagram lingkaran dapat dilihat pada Gambar 8.



Gambar 8. Diagram Lingkaran

Pada aplikasi ini, seluruh data yang telah di unggah dapat dilihat perkembangan nilai rata rata siswa dalam bentuk diagram batang. Nilai yang ditampilkan terpisah berdasarkan urutan data diunggah. Tampilan diagram nilai dapat dilihat pada Gambar 9.



Gambar 9. Diagram Batang

Selanjutnya, user dapat melihat data klasifikasi nilai yang sudah dipisahkan persiswa pada menu Data Individu. Tujuannya agar lebih memudahkan user dalam melihat siswa yang memperoleh nilai yang kurang baik pada mata pelajaran tertentu. Selain itu, user dapat melihat peningkatan maupun penurunan nilai rata rata siswa berdasarkan kelas. Terdapat juga tombol pencarian guna memudahkan pengguna untuk mencari siswa tertentu berdasarkan nama dan nomor induk siswa. Tampilan data individu dapat dilihat pada Gambar 10.

Nama Siswa	Nomor Induk	View Details
AISAH LAZ MUSA	202207003	View
AUCIA GISELLE GUNASETIAWAN	202207005	View
BIANCA VERGE SUMAEL	202207006	View
CHELSEI AURELLA	202207007	View
CHRISTIAN JAMES GERARD	202207008	View

Gambar 10. Data Individu

Pada menu ini terdapat tombol view. Tombol ini berfungsi untuk menampilkan tabel keterangan nilai yang berisi semua data pelajaran yang telah di upload. Tabel ini

akan menampilkan nilai individu siswa dari setiap kelas. Data yang ditampilkan berupa Mata Pelajaran, Kelas, Nilai, dan Keterangan. Tampilan dari tabel keterangan nilai dapat dilihat pada Gambar 11.

Mata Pelajaran	Kelas	Nilai	Keterangan
Bahasa Inggris	VII	85	Baik
Matematika	VII	82	Cukup
Pendidikan Agama	VII	92	Baik

Gambar 11. Tabel Keterangan Nilai

## 4. Kesimpulan

### 4.1 Kesimpulan

Berdasarkan hasil pembuatan aplikasi berbasis web untuk klasifikasi hasil belajar siswa di SMP XYZ, diperoleh kesimpulan sebagai berikut:

1. Metode Algoritma C4.5 dapat digunakan untuk mengklasifikasikan hasil belajar siswa di SMP XYZ berdasarkan riwayat akademik dengan tingkat akurasi sebesar 75%.
2. Data yang digunakan merupakan data nilai siswa tahun pertama yang terdiri dari nilai tugas, nilai ulangan harian, nilai UTS dan nilai UAS.
3. Aplikasi yang dibuat dapat menampilkan hasil berupa klasifikasi hasil belajar siswa berdasarkan nilai akademik siswa dan dapat menampilkan kategori dengan akurasi yang baik.
4. Aplikasi dapat berjalan sesuai dengan rancangan yang telah dibuat dan menampilkan hasil yang sesuai.

### 4.2 Saran

Berikut saran pengembangan dari aplikasi yang telah dibuat

1. Jumlah data yang digunakan perlu ditambah agar tingkat akurasi dari metode yang digunakan dapat memperoleh hasil yang lebih baik dan akurat
2. Pertimbangan menggunakan metode klasifikasi lain untuk membandingkan metode mana yang lebih akurat

## REFERENSI

- [1] Fanny Irnanda, Khairunisa ; Hartama, Dedy; Windarto, Agus Perdana. Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi, JURNAL MEDIA INFORMATIKA BUDIDARMA, Vol 4. No.1 2020
- [2] Sembiring, Muhammad Ardiansyah; Sibuea, Mustika Fitri Larasati; Sapta, Andy. ANALISA KINERJA ALGORITMA C.45 DALAM MEMPREDIKSI HASIL BELAJAR, JOURNAL OF SCIENCE AND SOCIAL RESEARCH Vol 1, No. 1. 2019.
- [3] Faishal; Kukuh, Hilmi. ANALISA TINGKAT KEMAMPUAN BELAJAR SISWA MENGGUNAKAN METODE C4.5 Studi Kasus : SMPN 2 Keresi TA

2020/2021, [https://eprints.amikom.ac.id/view/creators/Faisal=3AHilmy\\_Kukuh=3A=3A.html](https://eprints.amikom.ac.id/view/creators/Faisal=3AHilmy_Kukuh=3A=3A.html), 2022.

- [4] Novianti; Rismawan, Tedy; Bahri, Syamsul. IMPLEMENTASI DATA MINING DENGAN ALGORITMA C4.5 UNTUK PENJURUSAN SISWA (STUDI KASUS: SMA NEGERI 1 PONTIANAK), <https://jurnal.untan.ac.id/index.php/jcskommipa/article/view/17034>, Vol 4, no. 1. 2016

**Bryan Daniel Pinenda Pasaribu**, Mahasiswa S1, Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Tarumanagara

**Bagus Mulyawan S.Kom., M.M.**, Dosen Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Tarumanagara

**Tri Sutrisno , S.Si., M.Sc** Dosen Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Tarumanagara