

Data Mining Analysis of Student Perceptions of ChatGPT Usage Using the C4.5 Algorithm

Alifia Maulani¹⁾, Prastyadi Wibawa Rahayu^{2*)}, I Made Dwi Ardiada³⁾

¹⁾²⁾³⁾Program Studi Informatika, Fakultas Teknologi dan Informatika, Universitas Dhyana Pura

^{*)}Correspondence author: prastyadiwibawa@undhirabali.ac.id, Badung, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i1.3484>

Abstract

The rapid advancement of artificial intelligence, particularly ChatGPT, has brought considerable changes to the educational landscape. Despite its benefits, concerns have emerged regarding its potential negative impact on students' critical thinking skills and learning independence. This study aims to classify the perceptions of students at Dhyana Pura University toward the use of ChatGPT in education and to measure the accuracy of the resulting classification model. The method applied is data mining using the Knowledge Discovery in Databases (KDD) approach with the C4.5 algorithm. Data were collected from 350 active students through a Likert scale-based questionnaire covering five indicators: ease of use, perceived benefits, impact on critical thinking, impact on learning independence, and attitude toward ChatGPT usage. The classification process was carried out using Altair AI Studio with an 80:20 training and testing data split. The results indicate that the majority of students hold a neutral perception toward the use of ChatGPT in academic activities. The classification model achieved an accuracy rate of 86.96%, which is categorized as good. This research is expected to serve as a reference for educational institutions in formulating policies on the responsible use of AI in academic settings.

Keywords: ChatGPT, Data Mining, Perception, C4.5 Algorithm

Abstrak

Perkembangan kecerdasan buatan, khususnya ChatGPT, telah membawa perubahan signifikan dalam dunia pendidikan. Meskipun memberikan kemudahan, penggunaan ChatGPT juga memunculkan kekhawatiran terkait penurunan kemampuan berpikir kritis dan kemandirian belajar mahasiswa. Penelitian ini bertujuan untuk mengklasifikasikan persepsi mahasiswa Universitas Dhyana Pura terhadap penggunaan ChatGPT dalam dunia pendidikan serta mengukur tingkat akurasi model klasifikasi yang dihasilkan. Metode yang digunakan adalah *data mining* dengan pendekatan *Knowledge Discovery in Databases* (KDD) dan algoritma C4.5. Data diperoleh dari 350 responden mahasiswa aktif melalui kuesioner berbasis skala Likert yang mencakup lima indikator, yaitu kemudahan penggunaan, manfaat yang dirasakan, dampak terhadap berpikir kritis, dampak terhadap kemandirian belajar, dan sikap terhadap penggunaan ChatGPT. Proses klasifikasi dilakukan menggunakan perangkat lunak *Altair AI Studio* dengan pembagian data *training* dan *testing* sebesar 80:20. Hasil penelitian menunjukkan bahwa sebagian besar mahasiswa memiliki persepsi yang cenderung netral terhadap penggunaan ChatGPT. Model klasifikasi yang dibangun menghasilkan tingkat akurasi sebesar 86,96%, yang termasuk dalam kategori baik. Penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi institusi pendidikan dalam merancang kebijakan pemanfaatan AI yang bijak di lingkungan akademik.

Kata Kunci: ChatGPT, Data Mining, Persepsi, Algoritma C4.5

PENDAHULUAN

Perkembangan teknologi di era modern ini telah membawa banyak perubahan dalam kehidupan manusia sehari-hari, mulai dari cara bekerja, belajar, hingga berkomunikasi satu sama lain (Paidil & Sari, 2025). Salah satu bentuk teknologi yang paling banyak dibicarakan saat ini adalah kecerdasan buatan atau Artificial Intelligence (AI). Teknologi ini mampu memproses informasi dalam jumlah besar, mengenali pola tertentu, dan memberikan jawaban yang tepat secara otomatis. Karen akemmapuan itulah, AI kini sudah digunakan di banyak bidang penting seperti industri, kesehatan, dan juga pendidikan (Zendrato, 2024).

Salah satu produk AI yang paling dikenal luas saat ini adalah ChatGPT, yang dibuat oleh perusahaan OpenAI (Husnaini & Madhani, 2024). ChatGPT menggunakan teknologi yang disebut *Generative Pre-trained Transformer* (GPT), yang membuatnya bisa memahami pertanyaan dalam bahasa sehari-hari dan memberikan jawaban yang jelas dan lengkap (Salmi et al., 2023). Di dunia pendidikan, banyak mahasiswa yang mulai menggunakan ChatGPT sebagai alat bantu belajar, baik untuk memahami materi kuliah, mencari referensi, maupun membantu menyelesaikan tugas akademik dengan lebih cepat (Muhammad Tarmizi & Yahfizham, 2023).

Namun dibalik manfaatnya, penggunaan ChatGPT juga menimbulkan kekhawatiran. Sebagian mahasiswa cenderung langsung menyalin jawaban dari ChatGPT tanpa benar-benar memahami isi atau memeriksa kebenarannya, sehingga memunculkan risiko plagiarisme berbasis AI (Budi Susilo & Tri Widayanti, 2024). Lebih dari itu, kebiasaan mendapatkan jawaban secara instan dapat membuat mahasiswa malas berpikir sendiri dan kurang terlatih dalam memecahkan masalah secara mandiri (Faisal, 2024). Jika dibiarkan, kondisi ini justru bisa bertentangan dengan tujuan utama pendidikan itu sendiri, yaitu membentuk kemampuan berpikir kritis dan kemandirian dalam belajar.

Berdasarkan permasalahan tersebut, penelitian ini dilakukan di Universitas Dhyana Pura untuk mengetahui bagaimana persepsi mahasiswa terhadap penggunaan ChatGPT dalam kegiatan akademik. Metode yang digunakan adalah data mining dengan algoritma C4.5, yaitu sebuah algoritma yang bekerja dengan cara membentuk pohon keputusan untuk mengklasifikasikan data berdasarkan atribut yang paling berpengaruh (Dhea Halimah et al., 2022). Algoritma ini dipilih karena sudah terbukti mampu menghasilkan klasifikasi yang baik

dalam penelitian di bidang pendidikan. Pemanfaatan teknologi informasi dalam pembelajaran merupakan salah satu faktor penting dalam menunjang kualitas pendidikan, sehingga penting untuk memahami sikap mahasiswa terhadap AI secara lebih mendalam (Salmi et al., 2023). Penelitian ini bertujuan untuk mengklasifikasikan persepsi mahasiswa Universitas Dhyana Pura terhadap penggunaan ChatGPT dalam kegiatan akademik, sekaligus mengukur tingkat akurasi model klasifikasi yang dihasilkan menggunakan confusion matrix.

METODE PENELITIAN

Jenis penelitian yang digunakan adalah penelitian kuantitatif dengan pendekatan data mining. Data penelitian diperoleh dari hasil kuesioner yang kemudian diubah ke dalam bentuk data numerik untuk selanjutnya diolah dan dianalisis. Pendekatan data mining digunakan untuk menemukan pola dan informasi yang terkandung dalam data melalui serangkaian tahapan pengolahan data.

Penelitian ini dilakukan di Universitas Dhyana Pura, Bali, dengan pengumpulan data yang berlangsung pada bulan Februari hingga Maret 2026. Populasi dalam penelitian ini adalah seluruh mahasiswa aktif Universitas Dhyana Pura yang berjumlah 2.800 orang. Karena jumlah populasi yang cukup besar, pengambilan sampel dilakukan menggunakan rumus Slovin dengan tingkat kesalahan sebesar 5% agar jumlah sampel yang digunakan dapat mewakili populasi. Berdasarkan hasil perhitungan menggunakan rumus Slovin, jumlah sampel yang digunakan dalam penelitian ini sebanyak 350 mahasiswa. Rumus Slovin yang digunakan adalah sebagai berikut:

$$n = \frac{N}{1 + Ne^2} \quad (1)$$

Keterangan:

n : Jumlah Sampel

N : Jumlah Populasi

e : Tingkat Kesalahan (*margin of error*)

Data penelitian diperoleh melalui penyebaran kuesioner menggunakan Google Form. Instrumen penelitian terdiri dari 20 pernyataan yang disusun berdasarkan lima indikator, yaitu Kemudahan Pengguna, Manfaat yang Dirasakan, Dampak terhadap Berpikir Kritis, Dampak terhadap Kemandirian Belajar, dan Sikap terhadap Penggunaan ChatGPT (Salmi et al., 2023). Setiap pernyataan diukur menggunakan skala likert 5 tingkat, yaitu Sangat Tidak Setuju (1), Tidak Setuju (2), Ragu-Ragu (3), Setuju (4), dan Sangat Setuju (5). Penggunaan skala likert bertujuan untuk mengukur tingkat persepsi responden terhadap penggunaan ChatGPT dalam dunia pendidikan (Gustiana, 2024).

Analisis data pada penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD), yang merupakan sebuah teknik untuk menemukan informasi yang berguna dari kumpulan data yang tersimpan dalam basis data (Atalya Angelus Leza et al., 2024). Metode ini terdiri dari beberapa tahap yaitu:

- a. *Data Selection*: tahap *data selection* dilakukan dengan memilih data yang sesuai dengan kebutuhan penelitian.
- b. *Pre-Processing*: tahap *pre-processing* dilakukan untuk memastikan kualitas data sebelum proses analisis. Pada tahap ini dilakukan pengecekan terhadap kelengkapan data, kesesuaian jawaban responden, serta penghapusan data yang tidak sesuai kriteria penelitian.
- c. *Data Transformation*: tahap *data transformation* dilakukan untuk mengubah data ke dalam format yang siap digunakan pada proses klasifikasi. Pada tahap ini dilakukan perhitungan nilai rata-rata dari setiap responden menggunakan rumus mean (Situmeang et al., 2025).

$$x = \frac{\sum x}{n} \quad (2)$$

Keterangan :

x : nilai rata-rata

$\sum x$: total skor jawaban

n : jumlah item pernyataan

Hasil rata-rata kemudian diklasifikasikan ke dalam tiga label persepsi menggunakan metode interval kelas dengan rumus (Maratussholihah et al., n.d.):

$$Interval\ Kelas = \frac{Nilai\ Skor\ Tertinggi - Nilai\ Skor\ Terendah}{Jumlah\ kategori} \quad (3)$$

Berdasarkan rumus tersebut, perhitungan nilai interval kelas untuk data ini adalah sebagai berikut:

$$Interval\ Kelas = \frac{5 - 1}{3} = 1,33$$

Nilai interval sebesar 1,33 ini kemudian digunakan sebagai dasar penentuan rentang kategori persepsi mahasiswa, yang secara rinci disajikan pada Tabel 1.

Tabel 1. Kategori Persepsi

Rentang Skor	Kategori Persepsi
1,00 – 2,33	Negatif
2,34 – 3,67	Netral
3,68 – 5,00	Positif

- d. Data Mining: tahap data mining dilakukan menggunakan algoritma C4.5 untuk membangun model klasifikasi berupa pohon keputusan (*decision tree*). Proses klasifikasi dilakukan dengan membagi data berdasarkan atribut yang paling dominan melalui perhitungan nilai *Entropy* dan *Information Gain* menggunakan rumus sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (4)$$

Keterangan :

Entropy(S) : nilai entropy

n : jumlah kelas pada atribut

pi : proporsi kelas ke-i

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (5)$$

Keterangan :

Gain(S,A) : nilai gain atribut A

S : himpunan kasus

Si : partisi data

Proses pembentukan model klasifikasi dilakukan menggunakan perangkat lunak Altair AI Studio. Altair AI Studio (sebelumnya RapidMiner) adalah sebuah aplikasi komputer yang digunakan untuk mengolah data dengan memanfaatkan konsep dan metode data mining (Seminar & Matematika, 2021). Pada hasil klasifikasi, nilai akurasi ditampilkan dalam beberapa kategori dengan rentang yang berbeda-beda. Rincian parameter tersebut disajikan pada Tabel 2 (Husaini, 2023).

Tabel 2. Parameter Akurasi Klasifikasi

Nilai Akurasi (%)	Hasil Prediksi
90-100	Sangat Baik
80-90	Baik
70-80	Cukup
60-70	Kurang
50-60	Gagal

- e. *Evaluation*: tahap *evaluation* dilakukan untuk mengukur kinerja model klasifikasi yang dihasilkan. Evaluasi menggunakan *confusion matrix* dengan menghitung nilai *accuracy*, *precision*, dan *recall* (Gusti et al., 2025). Hasil evaluasi digunakan untuk mengetahui tingkat ketepatan model dalam mengklasifikasikan data penelitian.

HASIL DAN PEMBAHASAN

1. Data Selection

Proses pengumpulan data dilakukan melalui penyebaran kuesioner kepada mahasiswa aktif Universitas Dhyana Pura. Dari total responden yang mengisi kuesioner, sebanyak

367 tanggapan responden berhasil terkumpul. Namun, sesuai dengan jumlah sampel yang telah ditetapkan menggunakan rumus Slovin, hanya 350 data yang digunakan dalam analisis ini.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	15	P16	P17	P18	P19	P20
8	3	3	3	3	4	3	4	4	3	3	3	3	4	3	4	4	3	3	4	4
9	5	5	5	4	4	3	4	3	4	2	2	3	4	4	5	3	3	3	4	4
10	5	5	5	5	4	5	5	1	3	1	3	2	3	3	2	1	4	3	3	5
11	5	4	3	3	3	4	4	2	3	4	4	3	5	3	2	2	4	3	3	4
12	5	5	5	5	5	5	3	4	3	3	2	1	1	1	1	1	3	2	3	3
13	4	4	4	4	4	4	4	4	4	3	2	3	3	3	3	3	4	3	4	5
14	5	5	4	5	4	4	5	4	4	4	2	2	4	3	4	3	3	3	4	5
15	4	4	3	5	4	4	3	3	4	3	2	3	2	2	3	3	3	3	3	4
16	5	5	4	3	4	3	4	4	3	3	3	3	3	2	3	3	4	3	3	5
17	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
18	4	4	4	4	4	5	4	5	4	2	2	2	2	2	4	2	4	4	5	5
19	4	3	3	4	3	4	4	4	3	3	2	3	3	2	4	2	3	4	4	4
20	4	4	4	3	4	4	4	4	3	2	2	4	4	4	4	4	4	4	4	4
21	5	5	5	4	4	5	4	4	4	1	2	2	3	2	2	3	4	4	4	5
22	4	3	3	3	4	4	4	2	3	2	1	2	3	3	2	1	1	2	1	3
23	5	4	4	4	4	4	4	4	4	2	1	2	2	4	4	2	4	4	4	4
24	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5
25	5	4	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5
26	5	3	4	3	4	3	4	3	3	2	2	3	2	2	2	2	3	3	3	5
27	4	3	3	3	4	4	4	4	3	3	1	2	2	3	2	1	3	3	3	3
28	5	5	5	5	5	5	4	4	4	4	2	2	2	4	4	2	4	4	4	4
29	4	4	4	3	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4
30	5	4	4	5	3	4	5	4	4	4	5	5	4	5	4	5	4	5	4	4
31	4	5	4	3	2	3	4	3	4	1	1	4	2	2	3	1	3	3	3	5
32	5	4	5	5	5	5	5	5	5	4	3	5	2	2	4	2	4	3	4	4
33	5	5	5	5	5	5	5	4	4	2	2	4	2	4	4	3	4	4	4	4
34	5	5	3	4	4	4	4	4	4	5	3	3	4	4	4	4	3	4	5	4
35	5	5	5	5	4	4	4	4	5	3	3	4	3	4	5	3	4	4	4	4
36	5	5	5	5	5	5	4	5	4	5	5	5	5	5	4	4	4	3	5	5
37	5	5	5	5	5	5	5	3	3	3	2	4	3	3	3	3	4	3	3	4
38	4	4	4	4	3	5	5	5	4	2	2	3	2	2	3	1	3	3	4	5
39	3	5	5	4	4	4	4	3	3	1	1	1	1	1	1	1	4	4	2	5

Gambar 1. *Data Selection*

2. *Pre-Processing*

Berdasarkan hasil pemeriksaan, seluruh 350 data responden telah terisi dengan lengkap tanpa ditemukan nilai kosong. Oleh karena itu, tidak dilakukan penghapusan data apapun, dan keseluruhan data dinyatakan valid untuk digunakan pada tahap berikutnya. Selain pengecekan kelengkapan, pada tahap ini juga dilakukan pengelompokan nilai skala Likert ke dalam bentuk kategori jawaban (Sangat Tidak Setuju, Tidak Setuju, Ragu-Ragu, Setuju, Sangat Setuju) untuk mempermudah proses klasifikasi pada algoritma C4.5. Meski demikian, perhitungan rata-rata pada tahap transformasi data tetap menggunakan nilai numerik skala Likert aslinya.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
2	SS	S	S	SS	SS	S	SS	SS	S	TS	TS	TS	RR	RR	RR	S	S	S	S	S
3	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS
4	SS	RR	RR	SS	SS	SS	RR	RR	RR	SS	SS	RR	TS	TS	TS	RR	RR	SS	SS	SS
5	SS	SS	SS	RR	RR	SS	SS	SS	S	STS	STS	RR	TS	RR	TS	STS	S	S	S	S
6	SS	SS	SS	SS	SS	SS	SS	SS	S	S	RR	RR	RR	STS	TS	TS	S	RR	STS	RR
7	S	SS	SS	SS	SS	S	SS	SS	SS	STS	STS	STS	TS	STS	STS	RR	S	S	SS	SS
8	RR	RR	RR	RR	S	RR	S	S	RR	RR	RR	RR	S	RR	S	S	RR	RR	S	S
9	SS	SS	SS	S	S	RR	S	RR	S	TS	RR	S	S	SS	RR	RR	RR	S	S	S
10	SS	SS	SS	SS	S	SS	SS	STS	RR	STS	RR	TS	RR	RR	TS	STS	S	RR	RR	SS
11	SS	S	RR	RR	RR	S	S	TS	RR	S	S	RR	SS	RR	TS	TS	S	RR	RR	S
12	SS	SS	SS	SS	SS	SS	RR	S	RR	RR	TS	STS	STS	STS	STS	STS	RR	TS	RR	RR
13	S	S	S	S	S	S	S	S	S	RR	TS	RR	RR	RR	RR	RR	S	RR	S	SS
14	SS	SS	S	SS	S	S	SS	S	S	S	TS	TS	S	RR	S	RR	RR	RR	S	SS
15	S	S	RR	SS	S	S	RR	RR	S	RR	TS	RR	TS	TS	RR	RR	RR	RR	RR	S
16	SS	SS	S	RR	S	RR	S	S	RR	RR	RR	RR	TS	RR	RR	S	RR	RR	RR	SS
17	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS
18	S	S	S	S	S	SS	S	SS	S	TS	TS	TS	TS	S	TS	S	S	S	S	SS
19	S	RR	RR	S	RR	S	S	S	RR	RR	TS	RR	RR	TS	S	TS	RR	S	S	S
20	S	S	S	RR	S	S	S	S	RR	TS	TS	S	S	S	S	S	S	S	S	S
21	SS	SS	SS	S	S	SS	S	S	S	STS	TS	TS	RR	TS	TS	RR	S	S	S	SS
22	S	RR	RR	RR	S	S	S	TS	RR	TS	STS	TS	RR	RR	TS	STS	STS	TS	STS	RR
23	SS	S	S	S	S	S	S	S	S	TS	STS	TS	TS	S	S	TS	S	S	S	S
24	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	RR	SS
25	SS	S	SS	SS	SS	SS	SS	SS	SS	S	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS
26	SS	RR	S	RR	S	RR	S	RR	RR	TS	TS	RR	TS	TS	TS	TS	RR	RR	RR	SS
27	S	RR	RR	RR	S	S	S	S	RR	RR	STS	TS	TS	RR	TS	STS	RR	RR	RR	RR
28	SS	SS	SS	SS	SS	SS	S	S	S	S	TS	TS	TS	S	S	TS	S	S	S	S

Gambar 2. Hasil *Pre-Processing*

3. Data Transformation

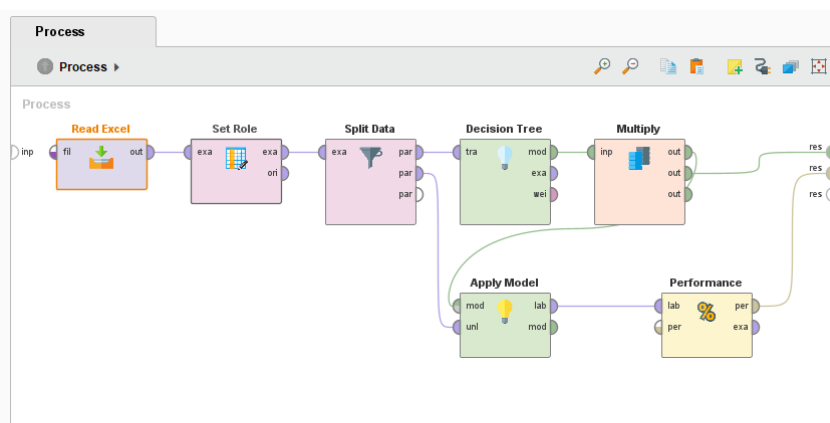
Pada tahap *data transformation*, data yang telah lolos proses *pre-processing* diolah lebih lanjut dengan menjumlahkan nilai dari seluruh atribut pernyataan (P1–P20) untuk setiap responden agar mendapatkan total skor. Total skor tersebut kemudian dihitung nilai rata-ratanya menggunakan rumus mean. Nilai rata-rata yang diperoleh selanjutnya diklasifikasikan ke dalam tiga kategori persepsi berdasarkan rentang interval kelas sebesar 1,33, sehingga menghasilkan label Negatif (skor 1,00–2,33), Netral (2,34–3,67), dan Positif (3,68–5,00). Hasil pelabelan ini menjadi dasar klasifikasi pada tahap berikutnya.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	Total	Rata R	Label
1	5	4	4	5	5	4	5	5	4	2	2	2	3	3	3	4	4	4	4	4	76	3,8	Positif
2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	100	5	Positif
3	5	3	3	5	5	5	3	3	3	5	5	3	2	2	2	3	3	5	5	5	75	3,75	Positif
4	5	5	5	3	3	5	5	5	4	1	1	3	2	3	2	1	4	4	4	4	69	3,45	Netral
5	5	5	5	5	5	5	5	5	4	4	3	3	3	3	1	2	2	4	3	1	73	3,65	Netral
6	4	5	5	5	5	4	5	5	5	1	1	1	1	1	2	1	1	3	4	4	67	3,35	Netral
7	3	3	3	3	4	3	4	4	3	3	3	3	4	3	4	4	3	3	4	4	68	3,4	Netral
8	5	5	5	4	4	3	4	3	4	2	2	3	4	4	5	3	3	3	4	4	74	3,7	Positif
9	5	5	5	5	4	5	5	1	3	1	3	2	3	3	2	1	4	3	3	5	68	3,4	Netral
10	5	4	3	3	3	4	4	2	3	4	4	3	5	3	3	2	2	4	3	4	68	3,4	Netral
11	5	5	5	5	5	5	3	4	3	3	2	1	1	1	1	1	3	2	3	3	61	3,05	Netral
12	4	4	4	4	4	4	4	4	4	3	2	3	3	3	3	3	3	4	3	4	72	3,6	Netral
13	5	5	4	5	4	4	5	4	4	4	2	2	4	3	4	3	3	3	4	5	77	3,85	Positif
14	4	4	3	5	4	4	3	3	4	3	2	3	2	2	3	3	3	3	3	4	65	3,25	Netral
15	5	5	4	3	4	3	4	4	3	3	3	3	3	2	3	3	4	3	3	5	70	3,5	Netral
16	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	100	5	Positif
17	4	4	4	4	4	5	4	5	4	2	2	2	2	2	4	2	4	4	5	5	72	3,6	Netral
18	4	3	3	4	3	4	4	4	3	3	2	3	3	2	4	2	3	4	4	4	66	3,3	Netral
19	4	4	4	3	4	4	4	4	3	2	2	4	4	4	4	4	4	4	4	4	74	3,7	Positif
20	5	5	5	4	4	5	4	4	4	1	2	2	3	2	2	3	4	4	4	5	72	3,6	Netral
21	4	3	3	3	4	4	4	2	3	2	1	2	3	3	2	1	1	2	1	3	51	2,55	Netral
22	5	4	4	4	4	4	4	4	4	2	1	2	2	4	4	2	4	4	4	4	70	3,5	Netral
23	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5	62	3,1	Netral
24	5	4	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	98	4,9	Positif
25	5	3	4	3	4	3	4	3	3	2	2	3	2	2	2	2	3	3	3	5	61	3,05	Netral
26	4	3	3	4	4	4	4	4	3	3	1	2	2	3	2	1	3	3	3	3	58	2,9	Netral
27	5	5	5	5	5	5	4	4	4	4	2	2	2	4	4	2	4	4	4	4	78	3,9	Positif
28	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	78	3,9	Positif
29	5	4	4	5	3	4	5	4	4	4	5	5	4	5	4	5	4	5	4	4	87	4,35	Positif

Gambar 3. Hasil Transformasi

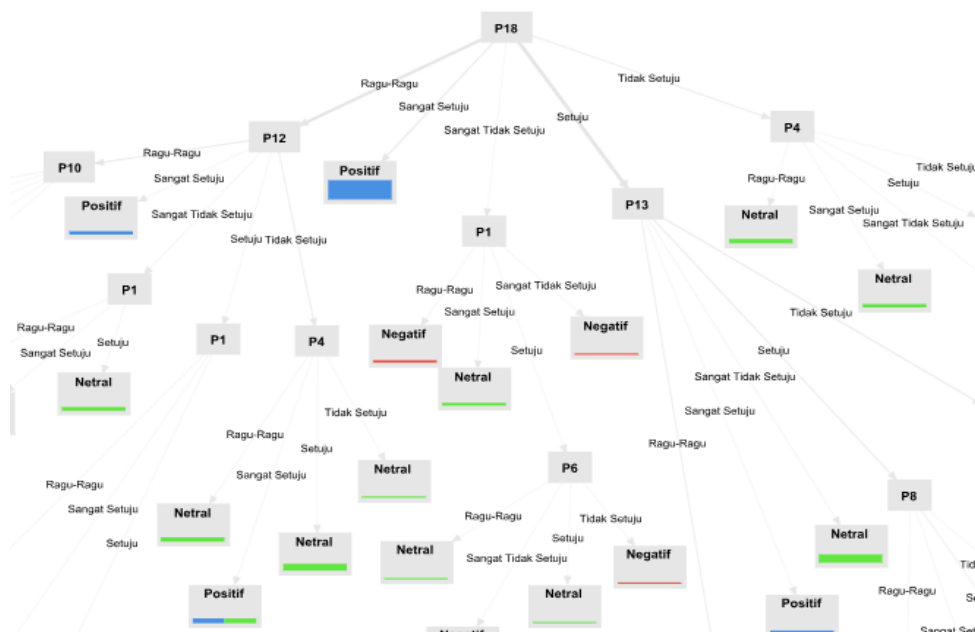
4. Klasifikasi Algoritma C4.5

Data yang telah ditransformasi diimpor pada Altair AI Studio untuk diklasifikasikan dengan algoritma C4.5. Proses ini dilakukan dengan menyusun rangkaian operator secara berurutan, meliputi *Read Excel* (membaca data), *Set Role* (menentukan kelas target), dan *Decision Tree* (metode klasifikasi utama). Alur proses tersebut digunakan untuk mengolah 350 baris data penelitian yang terdiri dari atribut P1 hingga P20 serta kolom label kategori. Pengujian model dilakukan menggunakan metode *split data* dengan rasio 80:20, menghasilkan pembagian proporsi sebanyak 281 data sebagai data *training* dan 69 data sebagai data *testing*.



Gambar 4. Alur Klasifikasi Algoritma C4.5

Kemudian untuk hasil pohon keputusan yang terbentuk dapat dilihat pada Gambar 5 berikut:



Gambar 5. Hasil Pohon Keputusan

Berdasarkan hasil proses klasifikasi, atribut P18 ("Saya akan terus menggunakan ChatGPT untuk mendukung kegiatan akademik saya") terpilih sebagai *root node* karena memiliki nilai *information gain* tertinggi dibanding atribut lainnya. Dari poin utama (P18) ini, pilihan jawaban responden (seperti Sangat Setuju, Setuju, atau Tidak Setuju) akan membentuk cabang-cabang baru. Cabang ini akan terus turun ke bawah sampai semua data mahasiswa selesai dikelompokkan ke dalam tiga hasil akhir: persepsi positif, netral, atau negatif.

5. Evaluation

Evaluasi kinerja model dilakukan menggunakan *confusion matrix* berdasarkan 69 data *testing*. Hasil evaluasi pada *Altair AI Studio* menunjukkan bahwa model C4.5 menghasilkan tingkat akurasi sebesar **86,96%**, yang termasuk dalam kategori **Baik** berdasarkan kategori parameter akurasi klasifikasi. Rincian hasil evaluasi per kelas disajikan pada Gambar berikut.

accuracy: 86.96%

	true Positif	true Netral	true Negatif	class precision
pred. Positif	23	3	0	88.46%
pred. Netral	5	36	1	85.71%
pred. Negatif	0	0	1	100.00%
class recall	82.14%	92.31%	50.00%	

Gambar 6. Hasil *Confusion Matrix*

Berdasarkan hasil evaluasi, kelas netral menunjukkan nilai *recall* tertinggi (92,31%), yang berarti model paling baik dalam mengenali data dengan persepsi netral. Kelas positif memiliki *precision* 88,46% dengan *recall* 82,14%, menandakan performa yang cukup seimbang. Sementara itu, kelas negatif memperoleh *precision* sempurna sebesar 100%, namun *recall*-nya hanya 50,00%. Rendahnya nilai *recall* pada kelas negatif disebabkan oleh jumlah data berkelas negatif yang relatif sedikit dibanding dua kelas lainnya, sehingga model belum cukup terlatih untuk mengenali pola kelas tersebut secara menyeluruh. Secara keseluruhan, algoritma C4.5 terbukti mampu mengklasifikasikan persepsi mahasiswa Universitas Dhyana Pura terhadap penggunaan ChatGPT dengan kinerja yang baik.

KESIMPULAN DAN REKOMENDASI

Berdasarkan analisis hasil penelitian mengenai persepsi mahasiswa Universitas Dhyana Pura terhadap penggunaan ChatGPT dalam dunia pendidikan menggunakan algoritma C4.5, ditemukan bahwa persepsi mahasiswa terbagi ke dalam tiga kategori, yaitu positif, netral, dan negatif. Hasil pengolahan data menunjukkan bahwa sebagian besar mahasiswa memiliki persepsi yang cenderung netral terhadap pemanfaatan teknologi ini. Kondisi tersebut diperkuat oleh terpilihnya atribut P18 yang berkaitan dengan sikap keberlanjutan penggunaan ChatGPT untuk kegiatan akademik sebagai *root node* karena memiliki nilai *information gain* tertinggi, sehingga menjadi faktor utama yang paling berpengaruh dalam menentukan arah persepsi mahasiswa secara keseluruhan.

Hasil evaluasi model menggunakan confusion matrix dengan pembagian data training dan testing sebesar 80:20 menunjukkan bahwa model klasifikasi C4.5 memperoleh

tingkat akurasi sebesar 86,96%. Berdasarkan kategori klasifikasi parameter akurasi, nilai tersebut termasuk dalam kategori baik. Dengan demikian, algoritma C4.5 dinilai mampu mengklasifikasikan persepsi mahasiswa terhadap penggunaan ChatGPT dalam dunia pendidikan dengan baik.

REFERENSI

- Atalya Angelus Leza, M., Widya Utami, N., & Anugrah Cahya Dewi, P. (2024). Prediksi Prestasi Siswa Smas Katolik Santo Yoseph Denpasar Berdasarkan Kedisiplinan Dan Tingkat Ekonomi Orang Tua Menggunakan Metode Knowledge Discovery In Database Dan Algoritma Regresi Linier Berganda. *Jati (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 373–379. <https://doi.org/10.36040/Jati.V8i1.8754>
- Budi Susilo, & Tri Widayanti. (2024). Kecerdasan Buatan: Plagiarisme Dan Perilaku Mandiri Siswa Sekolah Menengah Atas Dalam Penggunaan Chatgpt. *SABER : Jurnal Teknik Informatika, Sains Dan Ilmu Komunikasi*, 2(3), 341–352. <https://doi.org/10.59841/Saber.V2i3.1526>
- Dhea Halimah, Muhammad Ridwan Lubis, & Widodo Saputra. (2022). Algoritma C4.5 Untuk Menentukan Klasifikasi Tingkat Pemahaman Mahasiswa Pada Matakuliah Bahasa Pemrograman. *Jurnal Teknik Mesin, Industri, Elektro Dan Informatika*, 1(3), 24–38. <https://doi.org/10.55606/Jtmei.V1i3.534>
- Faisal, M. (2024). Dampak Kecerdasan Buatan (AI) Terhadap Pola Pikir Cerdas Mahasiswa Di Pontianak. *NUCLEUS*, 5(1), 60–66. <https://doi.org/10.37010/Nuc.V5i1.1684>
- Gusti, N., Dasriani, A., Ahmad, L., Pariandi, G., & Dharma, I. M. Y. (2025). Analisis Sentimen Program Jaminan Kesehatan Nasional Menggunakan Multiclass Support Vector Machine. *Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 6(1), 20–30.
- Gustiana, Z. (2024). *PERFORMANCE EVALUATION ALGORITMA C 4 . 5 PADA KLASIFIKASI DATA*. 5(2), 289–296. <https://doi.org/10.46576/Djtechno>
- Husaini, B. Q. (2023). *Penerapan Algoritma Decision Tree C45 Untuk Klasifikasi Penjurusan Siswa*. 9(1), 455–470.

- Husnaini, M., & Madhani, L. M. (2024). Perspektif Mahasiswa Terhadap Chatgpt Dalam Menyelesaikan Tugas Kuliah. *Journal Of Education Research*, 5(3), 2655–2664. <https://doi.org/10.37985/Jer.V5i3.1047>
- Maratussholihah, Z., Surabaya, U. N., Ahmadi, A., Surabaya, U. N., Raharjo, R. P., & Surabaya, U. N. (N.D.). *Persepsi Siswa Terhadap Pemanfaatan E-Book Sibi Sebagai Sarana Literasi Digital Dalam Pembelajaran Bahasa Indonesia*. 93–104.
- Muhammad Tarmizi, & Yahfizham, Y. (2023). Perspektif Mahasiswa Terhadap Penggunaan Kecerdasan Buatan Chatgpt Dalam Penyusunan Tugas Akhir. *Indiktika : Jurnal Inovasi Pendidikan Matematika*, 6(2), 151–161. <https://doi.org/10.31851/Indiktika.V6i2.15425>
- Paidil, P., & Sari, S. (2025). Peran Teknologi Terbaru Membentuk Kehidupan Di Era Digital. *JUPEIS: Jurnal Pendidikan Dan Ilmu Sosial*, 4(1), 8–15. <https://doi.org/10.57218/Jupeis.Vol4.Iss1.1326>
- Salmi, J., Setiyanti, A. A., Satya Wacana, K., Universitas, D., Satya, K., & Abstract, W. (2023). Persepsi Mahasiswa Terhadap Penggunaan Chatgpt Di Era Pendidikan 4.0[Student Perceptions Of The Use Of Chatgpt In The Era Of Education 4.0]. *Jurnal Ilmiah Wahana Pendidikan*, Oktober, 9(19), 399–406. <https://doi.org/10.5281/Zenodo.8403233>
- Seminar, P., & Matematika, N. (2021). *Akurasi Algoritma Klasifikasi Pada Software Rapidminer Dan Weka*. 4, 493–499.
- Situmeang, S. B., Khairani, A., & Fadillah, R. (2025). *Persepsi Pengguna Terhadap Penggunaan AI Dalam Kehidupan Sehari-Hari Pendekatan Technology Acceptance Model (TAM)*. 4(2), 40–50. <https://doi.org/10.62866/Buai.V4i2.235>
- Zendrato, C. P. (2024). Menyikapi Perkembangan Teknologi AI (Chatgpt) Sesuai Dengan Kebenaran Alkitabiah. *REI MAI: Jurnal Ilmu Teologi Dan Pendidikan Kristen*, 2(1), 23–37. <https://doi.org/10.69748/Jrm.V2i1.105>