

ALGORITMA C4.5 UNTUK PREDIKSI BIMBINGAN SISWA BERDASARKAN TIPOLOGI *HIPPOCRATES-GALENUS*

Boy Sandy Dwi Nugraha.H¹⁾, Sarjon Defit²⁾, Gunadi Widi Nurcahyo³⁾

Informatika

Fakultas Ilmu Komputer

Universitas Putra Indonesia YPTK Padang

E-mail : aboysandy@gmail.com¹⁾

The type of personality possessed by a student belived affect their behavior, whether positively or negatively, and if left unattended, it will harm the student. Student guidance is necessary to provide appropriate guidance for the student. This study aims to predict student guidance based on personality by using student data at SMP 1 Negeri Tembilahan. The data collection process was obtained from the BK teacher at SMPN 1 Tembilahan for grade 8 and grade 7. Grade 8 will be used as training data and grade 7 will be used as testing data. 5 parameters were selected for the prediction process and 1 label as the target class. The method used is the C4.5 algorithm to build a decision tree and obtain prediction rules. The results of the study were obtained using Confussion Matrix testing with a prediction accuracy rate of 70%. The ultimate goal of the student guidance prediction process is to have a higher percentage of "Yes" (need guidance) than "No" (don't need guidance) in the prediction results. Therefore, it can be stated that the prediction process model with the C4.5 algorithm is suitable for determining good decision-making results in terms of prediction, and the researcher hopes that after obtaining these results, the BK teacher at SMPN 1 Tembilahan can provide guidance as soon as possible and provide necessary guidance to students who need it.

Keywords : C4.5, Confussion Matrix, Data Mining, Personality, Prediction, Students

Intisari

Jenis kepribadian yang dimiliki seorang siswa diyakini akan mempengaruhi prilakunya baik dalam hal positif atau negatif yang apabila dibiarkan akan merugikan siswa tersebut. Bimbingan siswa perlu dilakukan untuk memberikan arahan yang sesuai untuk siswa tersebut. Penelitian ini bertujuan untuk melakukan prediksi bimbingan siswa berdasarkan kepribadian dengan menggunakan data siswa di SMP 1 Negeri Tembilahan, proses pengumpulan data didapatkan dari guru BK di SMPN 1 Tembilahan untuk kelas 8 dan kelas 7. Kelas 8 akan digunakan sebagai data latih dan kelas 7 digunakan sebagai data uji. 5 parameter dipilih untuk proses prediksi dan 1 label sebagai kelas target. Metode yang digunakan adalah Alogritma C4.5 untuk membangun pohon keputusan dan mendapatkan rule-rule prediksi, Hasil penelitian dilakukan menggunakan pengujian Confussion Matrix dengan tingkat akurasi prediksi sebesar 70%. Tujuan akhir dari proses prediksi bimbingan siswa adalah hasil prediksi dengan persentasi Ya(perlu bimbingan) lebih besar dari Tidak(tidak perlu bimbingan), maka dapat dinyatakan model proses prediksi dengan Alogitma C4.5 layak digunakan sebagai penentuan hasil keputusan yang baik dalam hal prediksi dan peneliti berharap setelah mendapatkan hasil tersebut diharapkan pihak Guru BK di SMPN 1 Tembilahan dapat melakukan proses bimbingan secepat mungkin dan memberikan bimbingan yang diperlukan oleh siswa yang membutuhkan.

Kata Kunci : C4.5, Confussion Matrix, Data Mining, Kepribadian, Prediksi, Siswa

1. PENDAHULUAN

Knowledge Discovery in Database (KDD) adalah proses yang berfokus pada penggalian atau pencarian informasi yang berguna dari kumpulan data yang besar, prosesnya melibatkan pembersihan data, pengambilan sampel, penskalaan, pengelompokan dan lainnya

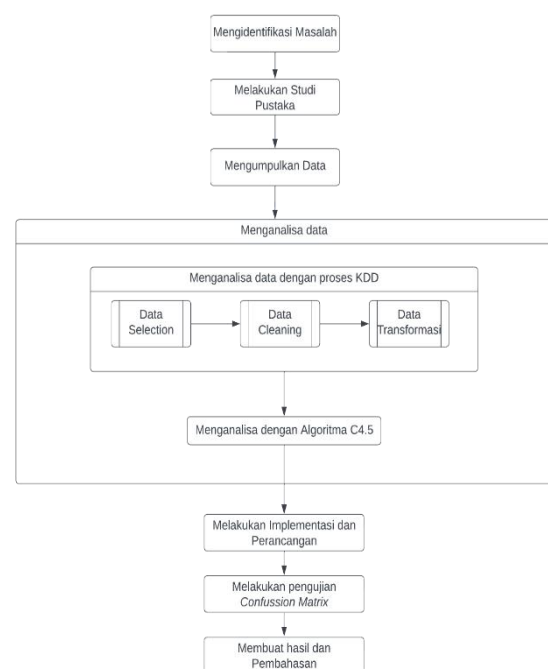
untuk mendapatkan wawasan yang bermakna[1]. KDD memiliki tahapan di dalamnya yaitu Data Mining, kegiatan yang meliputi pengumpulan dan pemakaian data historis dengan tujuan menemukan hubungan dalam set data berukuran besar[2]. Salah satu teknik pada Data mining adalah Prediksi. Prediksi sendiri memiliki arti peramalan atau perkiraan, dan menurut KBBI prediksi adalah

sebuah kegiatan untuk memprediksi, meramal, atau mempekirakan sesuatu dan prediksi bisa dipengaruhi oleh metode ilmiah atau hanya subjektif saja[3]. C4.5 adalah bagian dari kelompok algoritma *Decision Tree* Algoritma ini mempunyai input berupa *training samples* dan *samples*. Training samples berupa data contoh yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field – field data yang nantinya akan digunakan sebagai parameter dalam melakukan prediksi data dalam penelitian ini[4]. Tipologi *hippocrates-galenus* merupakan suatu pembelajaran ilmu tentang kepribadian, dalam tipologi tersebut dijelaskan bahwa karakteristik kepribadian yang dimiliki manusia ada empat yaitu melankolis, sanguinis, plegmatis dan koleris[5]. Siswa SMP memiliki sifat dan kepribadian yang masih perlu bimbingan dalam proses pendewasaannya [6] dan apabila siswa tidak diberikan arahan yang sesuai maka siswa tersebut dapat merugikan dirinya sendiri di kemudian hari nanti, dalam lingkungan sekolah hal ini perlu ditangani khusus oleh tim sekolah atau guru yang bisa menangani hal tersebut seperti contohnya guru Bimbingan Konseling (BK)[7]. Untuk menangani hal ini data siswa dan kepribadiannya diperlukan untuk melakukan prediksi bimbingan siswa, dan kepribadian siswa dapat ditentukan dengan memberikan beberapa pertanyaan psikologis pada siswa[8]. Beberapa penelitian mengenai metode C4.5 telah dilakukan sebelumnya, di antaranya yaitu penelitian Yee dan Ma Ma (2020), penelitian ini menggunakan C4.5 untuk pengidentifikasian racun[9]. Penelitian oleh Elacio (2020) untuk mengukur kepuasan karyawan berdasarkan teori Herzberg menggunakan algoritma C4.5[10]. Penelitian yang dilakukan oleh Shamrat, 2021 melakukan penelitian untuk membandingkan performa dari beberapa algoritma Prediksi Data Mining salah satunya C4.5[11]. Penelitian Meilana (2021) untuk mengklasifikasikan kepribadian *Hipocrates-Galenus*[5]. Penelitian oleh Xuanyuan dan Yue (2020) menggunakan algoritma C4.5 untuk pelayanan asuransi dan finansial[12], penelitian yang dilakukan oleh Dony Fahrudi mengklasifikasikan kepribadian dengan C4.5 dan naïve bayes [13], Penelitian tentang pengoptimasian entrophy pada Algoritma C4.5 oleh Sekhar Reddy[14], prediksi penyakit diabetes dengan Algoritma C4.5 oleh Sanni Ucha[15], lalu penelitian oleh

Panji Bimo mengklasifikasikan pohon keputusan dengan algoritma C4.5[16], dan penelitian tentang klasifikasi gizi balita dengan algoritma C4.5 dan *Swarm Optimization* oleh Amanhy Akhyar[17], penelitian yang dilakukan oleh Aprianto Tumangor untuk memprediksi tingkat kemampuan anak[18], lalu penelitian untuk memprediksi kerusakan mesin ATM oleh Dahri Yani Hakim[19]. Setelah mendapatkan hasil penelitian ini pihak Guru BK di SMPN 1 Tembilahan dapat melakukan proses bimbingan secepat mungkin dan memberikan bimbingan yang diperlukan oleh siswa yang membutuhkan.

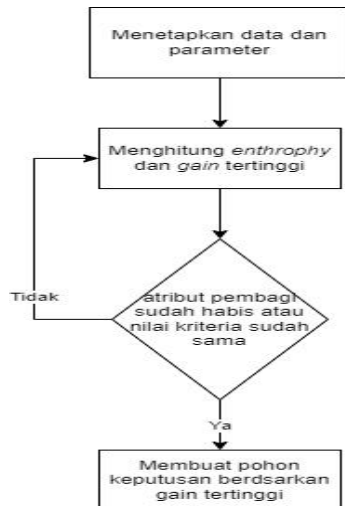
2. METODOLOGI

Dalam penelitian ini akan dibentuk dahulu suatu kerangka kerja penelitian. Kerangka kerja akan membantu dan memberikan proses apa saja dan apa yang akan dilakukan peneliti pada penelitiannya, Kerangka kerja penelitian ini dapat dilihat pada Gambar 1



Gambar 1. Kerangka Kerja Penelitian

Langkah-langkah penerapan metode C4.5 untuk memprediksi bisa dilihat pada Gambar 2 flowchart alur C4.5 berikut beserta tahapan yang akan dilakukan pada penelitian ini.



Gambar 2. Alur Algoritma C4.5

1. Menetapkan data dan parameter

Data yang akan diprediksi adalah data yang sudah dibersihkan bebas dari data yang rusak, tidak konsisten, data yang kosong atau berbagai macam kerusakan pada data dan siap digunakan untuk proses prediksi. Selanjutnya data siswa ditetapkan parameter yang akan digunakan untuk prediksi serta parameter label atau kelas tujuan dari proses prediksi dan nilai-nilai dari atribut tersebut data siswa dan nilai kriteria yang akan digunakan dapat dilihat pada tabel 1.

Tabel 1. Parameter Data Siswa

No	Nama Parameter	Kriteria
1	Jenis Kelamin	Laki-Laki Perempuan
2	Usia	12 Tahun 13 Tahun 14 Tahun 15 Tahun 16 Tahun
3	Finansial	Mampu Tidak Mampu
4	Anggota Keluarga	Lengkap Tidak Lengkap
5	Kepribadian	Sanguinis Koleris Melankolis Plegmatis
6	Bimbingan	Ya Tidak

2. Menghitung *Entropy* dan *Gain* tertinggi

Proses pencarian *Entropy* dan *Gain* dilakukan untuk membentuk pohon keputusan dan mendapatkan *rule* yang akan digunakan untuk memprediksi. Berikut adalah rumus mencari *entropy* dan *gain*.

$$Entropy(s) = - \sum_{i=1}^m p(\omega_i|s) \log_2 p(\omega_i|s) \quad (1)$$

dimana,

$p(\omega_i|s)$ = proporsi kelas ke- i dalam data yang diproses di *node s*

m = jumlah nilai berbeda dalam data

untuk menghitung *gain*

$$Gain(s, j) = E(s) - \sum_{i=1}^n p(v_i|s) \times E(s_i) \quad (2)$$

dengan ketentuan:

$p(v_i|s)$ = proporsi nilai v muncul pada kelas dalam *node s*

$E(s_i)$ = *entropy* komposisi nilai v dari kelas ke- j dalam data ke- i di *node s*

n = jumlah nilai berbeda dalam *node*

- Memastikan Parameter tidak ada lagi atau sama.

Dalam prosesnya untuk mengerahui proses perhitungan entrophy dan gain itu sudah selesai atau berakhir adalah dengan beberapa tanda seperti parameter yang sudah habis atau nilai dari kriteria yang sudah sama berikut adalah contoh gambar 3 dan 4 dari proses pencarian yang sudah selesai.

sanguinis perempuan usia 13					
	Jumlah	YA	TIDAK	Entrophy	Gain
Total	6	4	2	0.9183	
Finansial					0.31669
Mampu	5	4	1	0.72193	
Tidak Mampu	1	0	1	0	
Anggota Keluarga					0
Lengkap	6	4	2	0.9183	
Tidak Lengkap	0	0	0	0	

sanguinis perempuan usia 13 finansial mampu					
	Jumlah	YA	TIDAK	Entrophy	Gain
Total		4	1	0.72193	
Anggota Keluarga					0
Lengkap	5	4	1	0.72193	
Tidak Lengkap	0	0	0	0	

Parameter masih ada

Parameter tinggal 1 sebagai parameter akhir

Gambar 3. Parameter tidak ada lagi

Gambar parameter yang terakhir dan digunakan sebagai node akhir dari pohon keputusan

Kepribadian Koleris					
	Jumlah	YA	TIDAK	Entrophy	Gain
Total	4	2	2	1	
Jenis Kelamin					1
Laki-Laki	3	2	1	0	
Perempuan	1	0	1	0	
Usia					0.5
12 Tahun	0	0	0	0	
13 Tahun	2	1	1	1	

Nilai kriteria sudah sama

Gambar 4. Nilai Kriteria Sudah Sama

- Kembali melakukan proses No.3 jika tahap No.4 belum terpenuhi.
- Jika proses No.4 sudah terpenuhi dilanjutkan dengan membuat pohon keputusan berdasarkan *Gain* tertinggi.
- Menetapkan cabang, akar dan daun dari pohon
- Mengekstrak rule berdasarkan pohon keputusan yang sudah terbentuk.

3. HASIL DAN PEMBAHASAN

1. Menghitung *Gain* dan *Entropy*

Penelitian ini menggunakan data siswa kelas 7 dan kelas 8 pada tahun 2022, data yang sudah didapatkan akan dilakukan proses KDD yaitu *cleaning* setelah data siap data akan digunakan membangun pohon keputusan. Tahapan awal adalah dengan melakukan perhitungan untuk menentukan node pertama pada pohon keputusan

Tabel 2. Perhitungan Node Awal

	Entropy	Gain
Total	0.881290	
Jenis Kelamin		0.028084
Laki-Laki	0.650022	
Perempuan	0.940285	
Usia		0.073610
12 Tahun	0	
13 Tahun	0.845350	
14 Tahun	0.970950	
15 Tahun	1	
16 Tahun	0	
Finansial		0.074488
Mampu	0.787126	
Tidak Mampu	0.918295	
Anggota Keluarga		0.000662
Lengkap	0.873981	
Tidak Lengkap	0.918295	
Kepribadian		0.161177
Sanguinis	0.945660	
Koleris	1	
Melankolis	0	
Plegmatis	0	

Gain terbesar terdapat pada Kepribadian sehingga kepribadian akan menjadi node pertama, selanjutnya kita akan mencari node 2 yaitu kepribadian dengan cabang dari Sanguinis

Tabel 3. Node Kepribadian Sanguinis

	Entropy	Gain
Total	0.945660	
Jenis Kelamin		0.218387
Laki-Laki	0	
Perempuan	1	

Usia		0.173821
12 Tahun	0	
13 Tahun	0.811278	
14 Tahun	0	
15 Tahun	1	
16 Tahun	0	
Finansial		0.14448
Mampu	0.881290	
Tidak Mampu	0	
Anggota Keluarga		0.062977
Lengkap	0.970950	
Tidak Lengkap	0	

Gain terbesar terdapat pada jenis kelamin sehingga jenis kelamin akan menjadi node selanjutnya.

Tabel 4. Kepribadian Sanguinis Perempuan

	Enthrophy	Gain
Total	1	
Usia		0.311278
12 Tahun	0	
13 Tahun	0.91829	
14 Tahun	0	
15 Tahun	0	
16 Tahun	0	
Finansial		0.137925
Mampu	0.98522	
Tidak Mampu	0	
Anggota Keluarga		0
Lengkap	1	
Tidak Lengkap	0	

Gain terbesar terdapat pada usia sehingga usia akan menjadi node selanjutnya.

Tabel 5. Sanguinis Perempuan Usia 13

	Enthrophy	Gain
Total	0.918296	
Finansial		0.316689
Mampu	0.721928	
Tidak Mampu	0	
Anggota Keluarga		0
Lengkap	0.918296	
Tidak Lengkap	0	

Gain tertinggi adalah finansial dan finansial akan menjadi node selanjutnya.

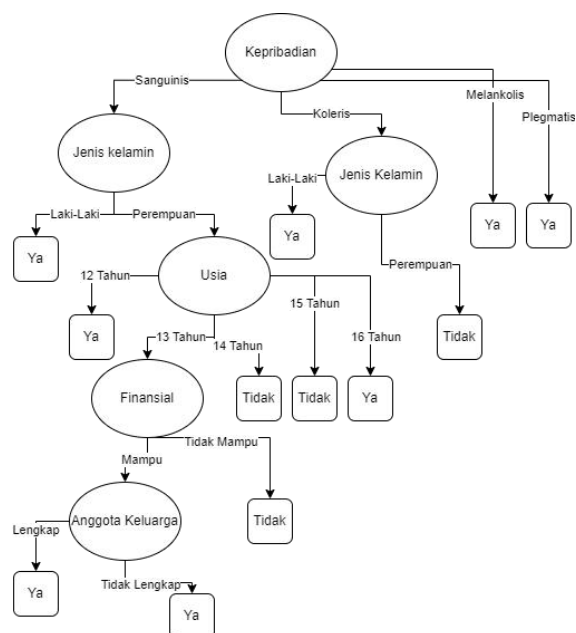
Tabel 6. Usia 13 tahun Finansial

	Enthrophy	Gain
Total	0.721928	
Anggota Keluarga		0
Lengkap	0.721928	
Tidak Lengkap	0	

Anggota keluarga menjadi Node terakhir pada pencarian pohon keputusan cabang Sanguinis.

2. Membuat pohon keputusan

Pohon keputusan dibuat berdasarkan hasil nilai gain tertinggi dari setiap node yang dihasilkan, perlu diingat disaat kita mencari gain tertinggi diberapapun node tanpa kita sadari kita juga sudah menyelesaikan beberapa cabang dari pohon keputusan sehingga tidak perlu dilakukan perhitungan ulang apabila cabang yang ingin dicari sudah selesai, pada gambar 5 di bawah ini merupakan pohon keputusan yang dibuat berdasarkan hasil dari node dan cabang yang sudah dicari, sebagai berikut:



Gambar 5. Pohon Keputusan

3. Rule Model

Setelah pohon keputusan berhasil dibuat maka selanjutnya menentukan *rule* model yang didapatkan berdasarkan pohon keputusan, berikut adalah *rule* model yang di hasilkan:

- if kepribadian **Sanguinis** and Jenis Kelamin **Laki-Laki** then bimbingan **ya**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **12 Tahun** then bimbingan **ya**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **13 Tahun** and Finansial **Mampu** and Anggota Keluarga **Lengkap** then bimbingan **ya**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **13 Tahun** and Finansial **Mampu** and Anggota Keluarga **Tidak Lengkap** then bimbingan **ya**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **13 Tahun** and Finansial **Tidak Mampu** then bimbingan **Tidak**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **14 Tahun** then bimbingan **tidak**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **15 Tahun** then bimbingan **tidak**
- if kepribadian **Sanguinis** and Jenis Kelamin **Perempuan** and Usia **16 Tahun** then bimbingan **ya**
- if kepribadian **Koleris** and Jenis Kelamin **Laki-Laki** then bimbingan **ya**
- if kepribadian **Koleris** and Jenis Kelamin **Perempuan** then bimbingan **tidak**
- if kepribadian **Melankolis** then bimbingan **ya**
- if kepribadian **Plegmatis** then bimbingan **ya**

4. Hasil Prediksi *Confussion Matrix*

Untuk menguji akurasi *rule* dari pohon keputusan yang sudah dibentuk digunakan pengujian *confusion matrix* untuk menentukan akurasi prediksi, tabel 7 menampilkan hasil pengujian sebagai berikut:

Tabel 7. Hasil Prediksi

Nama	Rule	Label	Prediksi
Keisha	Rule c	Ya	Ya
Azilla			
zahara			

Kenza izzatunnisa	Rule b	Ya	Ya
Saputri Agustia	Rule i	Tidak	Ya
Galia safira	Rule k	Ya	Ya
Malika Humaira	Rule i	Ya	Ya
M.davvi putra	Rule i	Ya	Ya
Jelita	Rule f	Ya	Tidak
Kanaya Nur Yazuwa	Rule b	Tidak	Ya
Regina Vatricia Huang	Rule k	Ya	Ya
Sugeng Hardyanto	Rule e	Tidak	Tidak

Dari hasil prediksi 10 data uji siswa didapatkan hasil prediksi 7 data siswa yang sesuai label dan hasil prediksinya. Tujuan akhir dari proses prediksi bimbingan siswa adalah hasil prediksi dengan persentasi Ya(perlu bimbingan) lebih besar dari Tidak(tidak perlu bimbingan), maka dapat dinyatakan model proses prediksi dengan Alogtima C4.5 layak digunakan sebagai penententuan hasil keputusan yang baik dalam hal prediksi berikut tabel 8 hasil perhitungan dari *Confusion Matrix*.

Tabel 8. Perhitungan *Confussion Matrix*

<i>f_{ij}</i>	Kelas hasil prediksi (<i>j</i>)	
	Kelas = Ya	Kelas = Tidak
Kelas asli (<i>i</i>)	Kelas = 6 Ya	1
	Kelas = 2 Tidak	1

$$\text{Akurasi} = \frac{6+1}{6+1+2+1} = \frac{7}{10} = 0,7\%$$

$$\text{Laju Error} = \frac{1+2}{6+1+2+1} = \frac{3}{10} = 0,3\%$$

Hasil pengujian akurasi metode dengan menggunakan confusion matrix menyimpulkan bahwa algoritma C4.5 dalam memprediksi bimbingan siswa mendapatkan akurasi sebesar 0,7 atau 70% dan laju error 0,3 atau 30%.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil analisa yang telah didapatkan, maka dihasilkanlah suatu pengetahuan (*knowledge*) bahwa kepribadian menjadi parameter utama yang menentukan prediksi bimbingan siswa, dari hasil pembuatan *rule* model didapatkan 12 *rule* yang akan dijadikan sebagai pedoman untuk memprediksi dan akurasi prediksi berhasil mendapat nilai sebesar 70% dengan melakukan pengujian dengan *Confusion Matrix*.

DAFTAR PUSTAKA

- [1] L. A. Briceno-Mena, M. Nnadili, M. G. Benton, and J. A. Romagnoli, "Data mining and knowledge discovery in chemical processes: Effect of alternative processing techniques," *Data-Centric Eng.*, vol. 3, no. 1, 2022, doi: [10.1017/dce.2022.21](https://doi.org/10.1017/dce.2022.21).
- [2] Yusuf Maulana, Riki Winanjaya, and Fitri Rizki, "Penerapan Data Mining dengan Algoritma C.45 Dalam Memprediksi Penjualan Tempe," *Bull. Comput. Sci. Res.*, vol. 2, no. 2, pp. 53–58, 2022, doi: [10.47065/bulletincsr.v2i2.163](https://doi.org/10.47065/bulletincsr.v2i2.163).
- [3] N. P. Setyadi, "Penerapan Data Mining Untuk Prediksi Hasil Produksi Karet Menggunakan Algoritma Decision Tree C4 . 5," vol. 2, no. 7, pp. 1–11, 2022. <http://dx.doi.org/10.12928/jstie.v2i1.2613>
- [4] T. Salsabilla, "Implementasi algoritma c4.5 untuk klasifikasi produk laris sepeda motor honda pada cv cendana motor cepiring 1) 1,2)," vol. 7, no. 2, pp. 164–171, 2022. <https://doi.org/10.36341/rabit.v7i2.2489>
- [5] M. Meilana, Y. Astuti, I. R. Wulandari, I. Sulistyowati, and B. A. Mimartiningtyas, "Algoritma Naive Bayes untuk Mengklasifikasikan Kepribadian Siswa SMP Berdasarkan Tipologi Hippocrates-Galenus," *Sistemasi*, vol. 10, no. 2, p. 480, 2021, doi: [10.32520/stmsi.v10i2.1339](https://doi.org/10.32520/stmsi.v10i2.1339).
- [6] M. Adibulasyhar, "Klasifikasi Karakteristik Kepribadian Manusia Berdasarkan Tipologi Hippocrates-Galenus Menggunakan Metode Fk-Nn," *Univ. Muhammadiyah Gresik*, pp. 6–21, 2019.
- [7] L. S. Sani and E. Munastiwi, "Karakteristik Siswa Program Tahfidz dalam Proses Menghafal di SMA Al-Muhajirin Purwakarta Ditinjau dari Tipologi Kepribadian Hippocrates-Galenus," *IQ (Ilmu Al-qur'an) J. Pendidik. Islam*, vol. 5, no. 01, pp. 87–107, 2022, doi: [10.37542/iq.v5i01.622](https://doi.org/10.37542/iq.v5i01.622).
- [8] R. Efkelin, R. A. Utami, and Y. Mailintina, "Hubungan Kepemimpinan Dengan Kinerja Perawat di Ruang Anggrek dan Gladiola Rumah Sakit Husada Jakarta," *J. Kesehat. Holist.*, vol. 7, no. 1, pp. 1–7, 2023, doi: [10.33377/jkh.v7i1.131](https://doi.org/10.33377/jkh.v7i1.131).
- [9] L. L. Yee and M. Ma Ma, "Identification of Poison using C4.5 Algorithm," *Int. J. Sci. Res. Sci. Eng. Technol.*, vol. 7, no. 2, pp. 218–222, 2020, doi: [10.32628/ijsrset207247](https://doi.org/10.32628/ijsrset207247).
- [10] A. A. Elacio, L. L. Lacatan, A. A. Vinluan, and F. G. Balazon, "Machine Learning Integration of Herzberg ' s Theory using C4 . 5 Algorithm," vol. 9, no. 1, pp. 57–63, 2020. <https://doi.org/10.30534/ijatcse/2020/1191.12020>.
- [11] F. M. J. M. Shamrat *et al.*, "Performance Evaluation among ID3 , C4 . 5 , and CART Decision Tree Algorithms," no. March, pp. 19–20, 2021. DOI: [10.1007/978-981-16-5640-8_11](https://doi.org/10.1007/978-981-16-5640-8_11)
- [12] S. Xuanyuan, S. Xuanyuan, and Y. Yue, "Application of C4.5 Algorithm in Insurance and Financial Services Using Data Mining Methods," *Mob. Inf. Syst.*, vol. 2022, 2022, doi: [10.1155/2022/5670784](https://doi.org/10.1155/2022/5670784).
- [13] D. Fahrudy *et al.*, "INTELLIGENT SYSTEM FOR CLASSIFICATION OF STUDENT PERSONALITY," vol. 5, no. 1, pp. 1–9, 2022. <https://doi.org/10.31598/sintechjournal.v5i1.969>
- [14] G. S. Reddy and S. Chittineni, "Entropy based C4.5-SHO algorithm with information gain optimization in data

- mining,” *PeerJ Comput. Sci.*, vol. 7, pp. 1–22, 2021, doi: 10.7717/PEERJ-CS.424. doi: [10.7717/PEERJ-CS.424](https://doi.org/10.7717/PEERJ-CS.424).
- [15] S. Ucha Putri, E. Irawan, F. Rizky, S. Tunas Bangsa, P. A. -Indonesia Jln Sudirman Blok No, and S. Utara, “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5,” *Januari*, vol. 2, no. 1, pp. 39–46, 2021. <https://doi.org/10.30645/kesatria.v2i1.56>
- [16] P. B. N. Setio, D. R. S. Saputro, and B. Winarno, “PRISMA, Prosiding Seminar Nasional Matematika Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5,” *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [17] A. Nazir, A. Akhyar, Y. Yusra, and E. Budianita, “Toddler Nutritional Status Classification Using C4.5 and Particle Swarm Optimization,” *Sci. J. Informatics*, vol. 9, no. 1, pp. 32–41, 2022, doi: [10.15294/sji.v9i1.33158](https://doi.org/10.15294/sji.v9i1.33158).
- [18] A. Tumanggor and P. S. Hasugian, “Penerapan Data Mining Untuk Memprediksi Tingkat Kemampuan Anak Dalam Mengikuti Mata Pelajaran Dengan Metode C4.5 Pada SDN 105351 Bakaran Batu,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 4, no. 1, pp. 57–63, 2021, doi: [10.32672/jnkti.v4i1.2718](https://doi.org/10.32672/jnkti.v4i1.2718).
- [19] D. Y. Hakim Tanjung, “Optimalisasi Algoritma C4.5 untuk Prediksi Kerusakan Mesin ATM,” *Infosys (Information Syst. J.)*, vol. 6, no. 1, p. 12, 2021, doi: 10.22303/infosys.6.1.2021.12-21. <http://dx.doi.org/10.22303/infosys.6.1.2021.12-21>