

## Klasifikasi Penjurusan pada Sekolah Menengah Atas (SMA) dengan Metode Algoritma C4.5

Yofhanda Septi Eirlangga<sup>1✉</sup>, Aldo Eko Syaputra<sup>2</sup>

<sup>1,2</sup>Universitas Adzkia

[fandaeir1langga@gmail.com](mailto:fandaeir1langga@gmail.com)

### Abstract

The Indonesian Minister of Education has stipulated 12 years of compulsory education, but only at high school (SMA) students can only determine the major they are interested in. In every new academic year the PPDB (New Student Admissions) committee has difficulty in determining majors because the determination of majors is still done manually so that there is a mismatch with the interests and talents of students caused by human error. So it is necessary to conduct a study with the aim of assisting the PPDB committee in determining the majors accurately and effectively so that the interests and talents of students can be channeled appropriately. Therefore, the C4.5 algorithm method is used which is one of the data mining algorithms that produces a decision tree that is suitable for classifying large amounts of data. The data used in this study were taken from the grade IX student report cards in 2022 as many as 44 report cards which have 11 subject value variables, namely Science, Social Sciences, Indonesian Language, PKN, Islamic Religious Education, MTK, English, BAM Arts and Culture, ICT and Physical Education. The results of this study obtained a decision tree (decision tree) with 17 rules (knowledge) that has been matched with real data with a very high level of accuracy. So that this research can help and become a benchmark for the PPDB committee in determining the majors of new students every time a new teaching begins at SMA Pertiwi 2 Padang.

Keywords: C4.5 Algorithm, High School Majors Classification, Course Determination, High Schools.

### Abstrak

Menteri Pendidikan Indonesia telah menetapkan wajib belajar 12 tahun namun hanya pada saat Sekolah Menengah Atas (SMA) saja peserta didik baru bisa menentukan jurusan yang diminatinya. Disetiap tahun ajaran baru panitia PPDB (Penerimaan Peserta Didik Baru) mengalami kesulitan dalam menentukan jurusan dikarenakan penentuan jurusan masih dilakukan manual sehingga adanya ketidakcocokan dengan minat bakat peserta didik yang diakibatkan oleh human error. Maka perlu dilakukan sebuah penelitian dengan tujuan untuk membantu panitia PPDB dalam menentukan jurusan secara akurat dan efektif agar minat dan bakat peserta didik dapat tersalurkan dengan tepat. Oleh karena itu digunakanlah metode algoritma C4.5 yang merupakan salah satu algoritma data mining yang menghasilkan pohon keputusan (*decision tree*) yang cocok dalam pengklasifikasian data yang banyak. Data yang dipergunakan dalam penelitian ini diambil dari nilai rapor peserta didik kelas IX tahun 2022 sebanyak 44 data rapor yang memiliki 11 variabel nilai mata pelajaran yaitu IPA, IPS, Bahasa Indonesia, PKN, Pendidikan Agama Islam, MTK, bahasa Inggris, BAM Seni Budaya, TIK dan Penjas. Hasil dari penelitian ini didapatkan pohon keputusan (*decision tree*) dengan 17 *rule (knowledge)* yang telah di cocokkan dengan data *real* dengan tingkat akurasi yang sangat tinggi. Sehingga menjadikan penelitian ini dapat membantu serta menjadi patokan bagi panitia PPDB dalam menentukan jurusan peserta didik baru setiap dimulainya ajaran baru di SMA Pertiwi 2 Padang.

Kata kunci : Algoritma C4.5, Klasifikasi Penjurusan, Penentuan Jurusan, Sekolah Menengah Atas.

*JIDT is licensed under a Creative Commons 4.0 International License.*



### 1. Pendahuluan

Teknologi informasi berkembang begitu pesat dan menyentuh pada semua sisi kehidupan [1]. Agar kegiatan manusia lebih mudah dan efisien, banyak peneliti melakukan kolaborasi antara disiplin ilmu komputer dan disiplin lain [2]. Termasuk salah satu kolaborasi berlangsung antara ilmu komputer dan ilmu Pendidikan [3].

Pendidikan adalah bagian terpenting dalam kehidupan manusia karena pendidikan dapat mengubah sikap dan pola pikir manusia itu sendiri [4]. Pemerintah Indonesia telah menerapkan wajib belajar 12 tahun [5]. yang terdiri dari SD 6 tahun, SMP 3 tahun dan SMA 3 tahun.

Dari SD hingga SMP, setiap siswa akan mengambil mata pelajaran yang sama. Namun, saat memasuki Sekolah Menengah Atas (SMA), setiap siswa akan perlu memilih jurusan [6].

Pemilihan jurusan oleh siswa sesuai minat ketika mendaftar di sekolah menengah atas memiliki tujuan untuk memberikan kesempatan bagi siswa untuk mengembangkan sikap kompetensi dan keterampilan siswa sesuai dengan minat, bakat, dan keterampilan akademik mereka dalam kelompok mata pelajaran dalam ilmu pengetahuan. Penentuan jurusan akan mempengaruhi tingkat akademik berikutnya dan akan mempengaruhi bidang ilmu pengetahuan atau studi untuk siswa yang akan melanjutkan ke tingkat

universitas sehingga pemilihan jurusan yang salah dapat membahayakan siswa dan masa depan mereka [7].

Jika melihat dari masalah yang ada, maka kita perlu sebuah metode yang dapat mengklasifikasikan data, klasifikasi sendiri mempunyai arti yaitu mengkotak-kotakan ataupun memisahkan *class* dari *dataset* yang digunakan [8]. Didalam data mining terdapat banyak metode untuk mengklasifikasikan data [9]. Salah satu metode untuk mengklasifikasikan data dalam data mining adalah algoritma C4.5 [10],[11]. Metode ini menghasilkan pohon keputusan dan rule [12].

Pohon keputusan ini nantinya yang akan menjadi acuan dalam membuat keputusan dalam menentukan sebuah hasil [13]. Pohon keputusan ini didapatkan dari data mentah yang kemudian dibersihkan dari atribut-atribut yang tidak dipakai ataupun tidak relevan [14] setelah data dibersihkan makadata tersebut siap diproses dan disebut dengan *dataset* [15], setelahnya data diolah menggunakan persamaan yang telah ditentukan dan didapatkanlah sebuah pohon keputusan (*decision tree*) dan rule atau aturan [16].

Dalam kasus ini metode algoritma C4.5 digunakan dalam membantu panitia Penerimaan Peserta Didik Baru (PPDB) dalam mengklasifikasikan penjurusan bagi siswa yang baru masuk di jenjang Sekolah Menengah Atas (SMA), algoritma C4.5 sendiri merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Decision Tree). Pohon keputusan merupakan metode klasifikasi dan prediksi yang terkenal [17], [18]. Pohon keputusan berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target [19].

Diharapkan dengan menggunakan metode ini dapat menyelesaikan penelitian ini, dan dapat membantu panitia PPDB dalam melakukan tugasnya dalam menentukan jurusan bagi siswa baru dan juga dapat meminimalisir kesalahan penentuan jurusan bagi siswa, agar proses pengembangan sikap kompetensi dan keterampilan kompetensi siswa sesuai dengan minat, bakat, dan keterampilan akademik mereka dalam kelompok mata pelajaran ilmu pengetahuan memiliki hasil yang baik untuk dapat dilanjutkan ke perguruan tinggi.

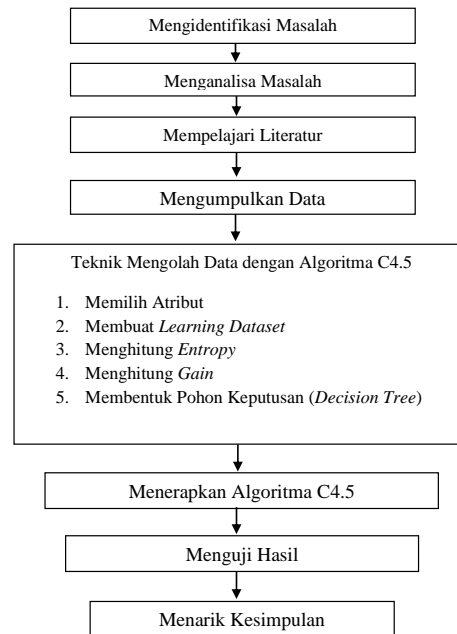
## 2. Metodologi Penelitian

### 2.1 Pendahuluan

Metodologi penelitian adalah tata cara melaksanakan penelitian ini, seperti mendefinisikan ruang lingkup masalah yang ada, menganalisa masalah, menentukan tujuan dari penelitian, mempelajari literatur, bagaimana teknik pengumpulan data, bagaimana teknik mengolah data dan implementasi Algoritma C4.5 serta pengujian hasil.

### 2.2 Penjelasan Kerangka Kerja Penelitian

Penjelasan kerangka kerja adalah rincian rencana kerja yang sudah disusun guna memandu tahap-tahap pekerjaan yang akan dilakukan pada penelitian ini agar dapat mencapai hasil yang diharapkan. Adapun gambaran dari kerangka kerja penelitian ini bisa dilihat pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

### 2.3 Uraian Kerangka Kerja Penelitian

Uraian kerangka kerja merupakan penjelasan secara rinci terhadap masing-masing sub dari kerangka kerja pada gambar 1, berikut penjelasan dari kerangka kerja

#### a. Mengidentifikasi Masalah

Tahap ini merupakan kerja awal dalam menentukan tujuan akhir dan melakukan pemecahan serta memberi solusi atas masalah yang ditemukan pada SMA Pertiwi 2 Padang dengan menggunakan algoritma C4.5

#### b. Menganalisa Masalah

Menganalisis masalah merupakan langkah selanjutnya agar masalah yang sudah teridentifikasi pada langkah sebelumnya dapat terpecahkan. Pada Langkah ini digambarkan proses untuk mendapatkan hasil rule.

#### c. Mempelajari Literatur

Mempelajari literatur ini bermaksud untuk mencari sumber literatur yang pas dengan penelitian guna mendapat gambaran dan contoh-contoh kongkrit yang sudah di lakukan sebelumnya.

#### d. Mengumpulkan Data

Langkah ini bermaksud untuk mempersiapkan data yang digunakan untuk diteliti yang didapat langsung dari SMA Pertiwi 2 Padang. Data yang di peroleh kemudian diolah secara manual menggunakan

algoritma C4.5 dan menggunakan Microsoft Excel 2013. Pada data nilai rapor siswa terdapat 11 variabel antara lain IPA, IPS, Bahasa Indonesia, PKN, Pendidikan Agama Islam, MTK, bahasa inggris, BAM Seni Budaya, TIK dan Penjas.

e. Teknik Mengolah Data Dengan Metode Algoritma C4.5

Tahapan selanjutnya yaitu mengolah data dengan Algoritma C4.5 ada beberapa proses yang akan dilakukan untuk mendapatkan hasil rule dan pohon keputusan adapun langkah-langkah yang dilakukan ialah sebagai berikut:

a. Memilih Atribut

Pemilihan atribut disini ialah sebagai eliminasi data awal agar atribut yang diperlukan saja yang akan dimasukan dalam pengolahan algoritma C4.5 agar hasil menjadi lebih akurat dan sesuai dengan tujuan yang telah di rencanakan diawal.

b. Membuat Learning Dataset

Learning dataset adalah sekelompok data table yang terdiri dari atribut-atribut yang telah dipilih dan siap untuk diolah.

c. Menghitung Entropy

Dari tabel Learning Dataset selanjutnya mencari Entropy masing-masing atribut dengan menggunakan formula yang telah di tentukan, jika telah terkumpul

entropy dari seluruh atribut, maka siap untuk masuk ke Langkah selanjutnya.

d. Menghitung Gain

Langkah ini sama hal nya dengan Entropy, Gain masing-masing atribut juga harus dicari dengan menggunakan formula yang telah ditentukan, setelah semua atribut dicari gain nya, maka gain yang tertinggi akan menjadi node akar (root node).

e. Membentuk Pohon Keputusan (Decision Tree).

Setelah didapat gain tertinggi, maka selanjutnya atribut gain tersebut akan menjadi node akar (root node), dan atribut yang tersisa akan di ulang mencari entropy dan gain nya, sehingga terbentuk daun (leaf) dari node akar tadi, langkah tersebut diulang sampai semua atribut habis, dan terbentuk pohon keputusan akhir [20].

3. Hasil dan Pembahasan

Tahap awal dalam penelitian ini adalah mempersiapkan data yang digunakan untuk diteliti yang didapat langsung dari SMA Pertiwi 2 Padang. Data yang di peroleh kemudian diolah secara manual menggunakan algoritma C4.5 dan menggunakan Microsoft Excel 2013. Pada data awal nilai rapor siswa terdapat 11 variabel dan hanya 9 yang akan digunakan sebagai data awal antara lain IPA, IPS, Bahasa Indonesia, PKN, Pendidikan Agama Islam, MTK, bahasa inggris, Seni Budaya, dan Penjas. Berikut data nilai siswa pada Tabel 1.

Tabel 1. Data Nilai Rapor Semester Genap Siswa Kelas IX

Nama	PAI	PKN	B.I	B.ING	MTK	IPA	IPS	SBUD	PJOK
Nur U.L	87	85	84	83	79	95	83	88	88
Nesti Violin	85	88	85	85	88	90	95	83	85
Nahdatull F	87	82	72	92	81	81	65	77	86
M. Yaser A	86	81	87	80	83	76	87	88	95
M. Ferino	79	77	77	76	74	75	80	79	80
Marshanda F	92	87	97	85	84	89	88	89	92
Kasturi	83	80	80	80	80	80	80	80	80
Faiza Zahra	80	80	84	75	76	77	78	80	83
Diyana S	82	93	91	86	80	80	87	84	86
Dinda P D	85	90	85	85	86	85	89	91	86
Adinda M R	83	83	89	81	82	84	85	90	93
Berlian N A	85	88	94	78	81	78	79	87	92
Ayu Aprilia	89	87	87	86	86	85	87	95	88
Artha Ria D	95	90	90	90	83	90	87	90	89
Yuni Astuti	85	87	88	83	82	82	90	90	90
Sultan Thoriq	80	85	88	81	82	85	95	86	85
Sevilla S P	88	93	94	80	82	90	90	93	92
Sentia W	84	83	87	82	84	79	85	80	85
Rizka H	89	89	92	86	81	93	85	92	90
Ricky R	90	88	85	85	95	85	90	85	90

Setelah mendapatkan data diatas, mak dilanjutkan ke proses Algoritma C4.5. Langkah pemrosesan Algoritma C4.5 sebagai berikut:

a. Pemilihan Atribut

Dari Tabel 1 diatas dilakukan proses Data Selection yang mana nilai SBUD dan PENJAS tidak dimasukan dalam data penelitian karena tidak semua anak memiliki nilai tersebut, jadi hanya nilai pengetahuan dari 7 mata pelajaran yang akan menjadi sampel data yang digunakan. Setelah proses Data Selection selesai,

kemudian masuk ke dalam proses data cleaning yang mana proses penghapusan record dan variable pada table 1 yang tidak komplit atau valid, karena data ini tidak akan berpengaruh dan tidak akan digunakan dalam prediksi penentuan jurusan siswa. Setelah proses diatas selesai, berikutnya masuk ke proses data transformastion yaitu pengelompokkan masing-masing nilai berdasarkan keterangan dari SMA Pertiwi 2 Padang, dimana pada atribut dilakukan pengkategorian seperti pada Tabel 2.

Tabel 2. Transformasi Masing-Masing Nilai

Nilai	Transformasi
90-100	A
80-89	B
70-79	C
50-69	D
0-50	E

b. Membuat *Learning Dataset*

Pembuatan *Learning Dataset* pada Tabel 3 dibawah mengacu pada tranformasi nilai pada Tabel 2 diatas.

Tabel 3. *Data Transformation* Nilai Rapor Siswa

PKN	B.IND	B.ING	MTK	IPA	IPS	KEPUTUSAN
B	B	B	C	A	B	IPA
B	B	B	B	A	A	IPS
B	C	A	B	B	C	IPS
B	B	B	B	C	B	IPA
C	C	C	C	C	B	IPA
B	A	B	B	B	B	IPS
B	B	B	B	B	B	IPS
B	B	C	C	C	C	IPS
A	A	B	B	B	B	IPA
A	B	B	B	B	B	IPA
B	B	B	B	B	B	IPS
B	A	C	B	C	C	IPA
B	B	B	B	B	B	IPS
A	A	A	B	A	B	IPA
B	B	B	B	B	A	IPA
B	B	B	B	B	A	IPA
A	A	B	B	A	A	IPA
B	B	B	B	C	B	IPA
B	A	B	B	A	B	IPA
B	B	B	A	B	A	IPA

c. Menghitung *Entropy*

Untuk mendapatkan atribut mana yang akan menjadi Node 1, maka harus dilakukan perhitungan terhadap entropy total. Jumlah data yang digunakan sebagai sampel penelitian berjumlah 44, Dimana terbagi menjadi 21 sampel IPA dan 23 sampel IPS, Untuk menghitung nilai entropy total dapat dilihat dengan Persamaan (1).

$$Entropy(S) = \sum_{j=1}^k - p_j \log_2 p_j \tag{1}$$

Dimana *S* ialah himpunan (dataset) kasus. *k* merupakan banyaknya partisi *S*. Dan *p<sub>j</sub>* merupakan probabilitas dari Sum (Ya) / total kasus.

Langkah selanjutnya adalah menghitung nilai entropy untuk setiap atribut berdasarkan klasifikasi. Untuk menghitung nilai entropy dapat menggunakan persamaan (1) untuk setiap atribut. Berikut bentuk perhitungannya:

Atribut Nilai PKN

a. Entropy A

$$= \left(-\frac{5}{6} * \log_2 \left(\frac{5}{6}\right)\right) + \left(-\frac{1}{6} * \log_2 \left(\frac{1}{6}\right)\right) = 0,65002242$$

b. Entropy B

$$= \left(-\frac{15}{35} * \log_2 \left(\frac{15}{35}\right)\right) + \left(-\frac{15}{35} * \log_2 \left(\frac{15}{35}\right)\right) = 0,98522814$$

c. Entropy C

$$= \left(-\frac{1}{3} * \log_2 \left(\frac{1}{3}\right)\right) + \left(-\frac{2}{3} * \log_2 \left(\frac{2}{3}\right)\right) = 0,918295834$$

Dan seterusnya sampai semua perhitungan atribut selesai dilakukan

d. Menghitung *Gain*

Perhitungan nilai *Gain* setiap atribut berdasarkan nilai *Entropy*. Nilai *Gain* tertinggi akan menjadi node 1 dengan Persamaan (2).

$$Gain(A) = Entropy(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} \times Entropy(S_i) \tag{2}$$

*S* merupakan data sampel yang digunakan untuk trening data. *A* ialah data atribut. *|S<sub>i</sub>|* merupakan keseluruhan untuk sampel nilai *V*. *|S|* adalah total semua data sampel. Dan *Entropy(S<sub>i</sub>)* adalah sampel yang mempunyai nilai *i*.

A. *Gain* (Total, PKN)

$$= Entropy(S) - \sum_{i=1}^n \frac{|PKN_i|}{|Total|} * Entropy(PKN_i) = 0,9985091 - \left(\left(\frac{6}{44} * 0,65002242\right) + \left(\frac{35}{44} * 0,98522814\right) + \left(\frac{3}{44} * 0,389975\right)\right) = 0,998509099$$

Dan seterusnya sampai semua perhitungan atribut selesai dilakukan. Setelah menghitung entropy dan gain, maka selanjutnya menentukan gain tertinggi untuk membentuk pohon keputusan.

Kunci pencarian entropy:

a. Jika pada kolom “IPA” ada yang bernilai nol (0) dan juga pada kolom “IPS” ada yang bernilai nol (0), sudah bisa dipastikan entropy nya juga bernilai nol (0).

b. Jika pada kolom “IPA” ada yang bernilai satu (1) dan juga pada kolom “IPS” ada yang bernilai satu (1), sudah bisa dipastikan entropy nya juga bernilai satu (1).

Kunci pemberian Keputusan

a. Jika pada kolom “IPS” bernilai 0 dan kolom “IPA” mempunyai nilai contohnya 1 maka keputusan adalah “IPA”.

b. Jika pada kolom “IPA” bernilai 0 dan kolom “IPS” mempunyai nilai contohnya 1 maka keputusan adalah “IPS.”

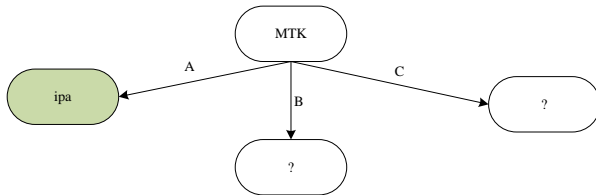
c. Jika kedua kolom “IPS” dan “IPA” sama-sama bernilai 0 berarti tidak mempengaruhi keputusan. Hasil nilai gain terbesar bisa dilihat pada Tabel 4.

Tabel 4. Penentuan Gain Terbesar

Node		Jumlah	IPA	IPS	Entropy	Gain
1	Total	44	21	23	0,998509099	
	PKN					0,0635544
	A	6	5	1	0,650022422	
	B	35	15	20	0,985228136	
	C	3	1	2	0,918295834	
	B.IND					0,0632882
	A	13	9	4	0,89049164	
	B	27	11	16	0,975119065	
	C	4	1	3	0,811278124	
	B.ING					0,0256371
	A	5	2	3	0,970950594	
	B	32	17	15	0,997180399	
	C	7	2	5	0,863120569	
	MTK					0,1156718
	A	4	4	0	0	
	B	33	15	18	0,994030211	
	C	7	2	5	0,863120569	
	IPA					0,032838
	A	7	5	2	0,873499722	
	B	29	12	17	0,978449329	
	C	8	4	4	1	
	IPS					0,0593744
	A	10	6	4	0,942179356	
	B	28	14	14	1	
	C	6	1	5	0,650022422	

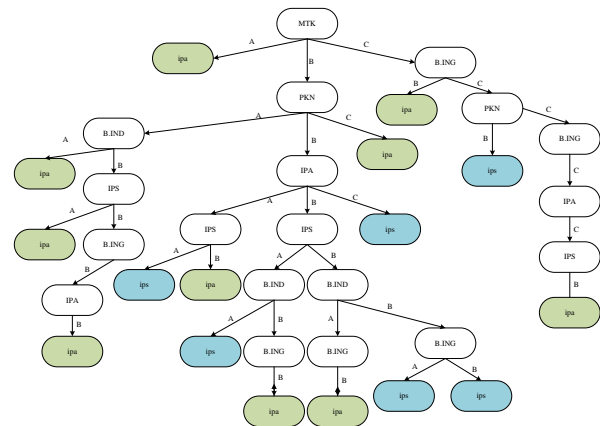
e. Membentuk Pohon Keputusan (*Decision Tree*)

Pada pencarian pohon keputusan awal dapat dilihat pada Tabel 4 diatas dari nilai gain terbesar yang berada pada atribut “MTK” dengan angka sebesar 0,1156718. Dengan begitu “MTK” dinobatkan menjadi akar pohon node 1 yang disajikan pada Gambar 2.



Gambar 2. Hasil Pohon Keputusan Dari Node 1

Setelah didapatkan atribut node 1, selanjutnya mencari atribut yang akan dijadikan cabang node 1.2. Pada cabang Node 1 terdapat 3 perhitungan klasifikasi yang berbeda yaitu atribut “MTK” klasifikasi “A”, “B” dan “C”. Kemudian yang akan dilakukan adalah perhitungan untuk mencari pohon keputusan cabang node 1.2 dan seterusnya dengan menggunakan langkah-langkah seperti sebelumnya sampai didapatkan pohon keputusan akhir pada Gambar 3.



Gambar 3. Pohon Keputusan Akhir

Rule (*Knowledge*) yang didapatkan dari perhitungan pohon keputusan akhir berdasarkan ketentuan persamaan algoritma ini mendapatkan 17 aturan (*rule*) dalam mengklasifikasi jurusan siswa.

Algoritma C4.5 telah berhasil dalam menganalisa data untuk penentuan jurusan siswa dan didapatkan rule-rule yang bisa membantu memudahkan panitia PPDB SMA Pertiwi 2 Padang dalam mengambil keputusan penetapan jurusan siswa baru.

4. Kesimpulan

Penelitian yang telah dilakukan pada SMA Pertiwi 2 Padang dalam mengklasifikasi data penjurusan siswa baru dengan menggunakan metode algoritma C4.5 dan melakukan uji data menggunakan aplikasi RapidMiner, dapat memberikan pengetahuan untuk mengambil



keputusan penentuan jurusan karena menghasilkan rule yang sangat mudah untuk dipahami sehingga dapat meningkatkan kinerja panitia PPDB SMA Pertiwi 2 Padang.

### Ucapan Terimakasih

Kami mengucapkan terimakasih sebesar-besarnya kepada Universitas Adzka yang telah mendanai penelitian ini dan menyediakan fasilitas dengan Kontrak Nomor. 19/KONTRAK-PENELITIAN/UNIV-ADZ/2022.

### Daftar Referensi

- [1] Ifongki, I. (2020). Penerapan Data Mining Menggunakan Algoritma C4.5 Terhadap Pengaruh Penjualan Kopi Pada Pt. Jpw Indonesia. *Jurnal Sistem Informasi Dan Informatika (Simika)*, 3(1), 40–54. <https://doi.org/10.47080/simika.v3i1.836>
- [2] Azwanti, N. (2018). Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 13(1), 33. <https://doi.org/10.30872/jim.v13i1.629>
- [3] Gustian, D., & Hundayani, R. D. (n.d.). Combination of AHP Method With C4 . 5 in The Level Classification Level Out Students. 4–9. <https://doi.org/10.1109/CED.2017.8308098>
- [4] Yudi Hidayat, E., Sabiq Taufiqurrahman, A., Luthfiarta, A., Zeniarja, J., Agus Santoso, H., Nugraha, A., & Rismiyati. (2018). Implementation of Weighted Naive Bayes Algorithm for Major Determination in Indonesian High School. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISEmantic 2018*, 580–584. <https://doi.org/10.1109/ISEMANTIC.2018.8549761>
- [5] Priya, A., & Meenakshi, E. (2017). Detection of phishing websites using C4.5 data mining algorithm. *RTEICT 2017 - 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, Proceedings*, 2018-Janua, 1468–1472. <https://doi.org/10.1109/RTEICT.2017.8256841>
- [6] Gustian, D., Yulitasari, S., Hundayani, D., Muslih, & Nunik. (2018). Comparison of C4.5 method based optimization algorithm to determine eligibility of beneficiaries of direct community assistance (Case study: Kelurahan Cicurug). *3rd International Conference on Computing, Engineering, and Design, ICCED 2017*, 2018-March, 1–6. <https://doi.org/10.1109/CED.2017.8308109>
- [7] Hajrahnur, S., Nasrun, M., Setianingsih, C., & Murti, M. A. (2018). Classification of posts Twitter traffic jam the city of Jakarta using algorithm C4.5. *2018 International Conference on Signals and Systems, ICSigSys 2018 - Proceedings*, 294–300. <https://doi.org/10.1109/ICSIGSYS.2018.8372776>
- [8] Harman, R. (2018). Penerapan Algoritma C.45 Untuk Memprediksi Indikator Website Yang Baik. *Jurnal Ilmiah Informatika*, 06(02), 54–59. <https://doi.org/10.33884/jif.v6i02.545>
- [9] Jia, D., Xie, W., Chen, Z., & Qiang, B. (2018). Publication recommendations of manuscripts using improved C4.5 decision tree algorithm. *Advances in Intelligent Systems and Computing*, 686, 710–714. [https://doi.org/10.1007/978-3-319-69096-4\\_100](https://doi.org/10.1007/978-3-319-69096-4_100)
- [10] Budiman, E., Haviluddin, Dengan, N., Kridalaksana, A. H., Wati, M., & Purnawansyah. (2018). Performance of Decision Tree C4.5 Algorithm in Student Academic Evaluation. *Lecture Notes in Electrical Engineering*, 488, 380–389. [https://doi.org/10.1007/978-981-10-8276-4\\_36](https://doi.org/10.1007/978-981-10-8276-4_36)
- [11] Elisa, E. (2017). Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. *Jurnal Online Informatika*, 2(1), 36. <https://doi.org/10.15575/join.v2i1.71>
- [12] Fiandra, Y. A., Defit, S., & Yuhandri, Y. (2017). Penerapan Algoritma C4.5 untuk Klasifikasi Data Rekam Medis berdasarkan International Classification Diseases (ICD-10). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(2), 82–89. <https://doi.org/10.29207/resti.v1i2.48>
- [13] Sulastris, H., & Gufroni, A. I. (2017). Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 3(2), 299–305. <https://doi.org/10.25077/teknosi.v3i2.2017.299-305>
- [14] Suryani, D., Labellapansa, A., & Marsela, E. (2018). Proceedings of the Second International Conference on the Future of ASEAN (ICoFA) 2017 – Volume 2. In *Proceedings of the Second International Conference on the Future of ASEAN (ICoFA) 2017 – Volume 2 (Vol. 2)*. Springer Singapore. <https://doi.org/10.1007/978-981-10-8471-3>
- [15] Suyatno, J. A., Nhita, F., & Rohmawati, A. A. (2018). Rainfall forecasting in Bandung regency using C4.5 algorithm. *2018 6th International Conference on Information and Communication Technology, ICOICT 2018*, 0(c), 324–328. <https://doi.org/10.1109/ICOICT.2018.8528725>
- [16] Muslim, M. A., Nurzahputra, A., & Prasetyo, B. (2018). Improving accuracy of C4.5 algorithm using split feature reduction model and bagging ensemble for credit card risk prediction. *2018 International Conference on Information and Communications Technology, ICOIACT 2018*, 2018-Janua(1996),141–145. <https://doi.org/10.1109/ICOIACT.2018.8350753>
- [17] Tamin, F., & Iswari, N. M. S. (2017). Implementation of C4.5 algorithm to determine hospital readmission rate of diabetes patient. *Proceedings of 2017 4th International Conference on New Media Studies, CONMEDIA 2017*, 2018-Janua, 15–18. <https://doi.org/10.1109/CONMEDIA.2017.8266024>
- [18] Asroni, A., Masajeng Respati, B., & Riyadi, S. (2018). Penerapan Algoritma C4.5 untuk Klasifikasi Jenis Pekerjaan Alumni di Universitas Muhammadiyah Yogyakarta. *Semesta Teknika*, 21(2), 158–165. <https://doi.org/10.18196/st.212222>
- [19] Pertiwi, A. G., Widyaningtyas, T., & Pujiyanto, U. (2018). Classification of province based on dropout rate using C4.5 algorithm. *Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology, SIET 2017*, 2018-Janua, 410–413. <https://doi.org/10.1109/SIET.2017.8304173>
- [20] Nugraha, A., Arista Harum Perdana, M., Agus Santoso, H., Zeniarja, J., Luthfiarta, A., & Pertiwi, A. (2018). Determining the Senior High School Major Using Agglomerative Hierarchical Clustering Algorithm. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISEmantic 2018*, 225–228. <https://doi.org/10.1109/ISEMANTIC.2018.8549834>