

Penerapan Algoritma CART Dalam Menentukan Jurusan Siswa di MAN 1 Inhil

Siti Monalisa ^{[1]*}, Fakhri Hadi ^[2]

Jurusan Sistem Informasi^{[1], [2]}

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

siti.monalisa@uin-suska.ac.id^[1], fakhri.hadi@student.uin-suska.ac.id^[2]

Abstract— MAN 1 Inhil is a school that applies ministerial regulations to determine the direction of student majors at the beginning of entry, namely in class X. Determination of majors is done by considering several indicators, namely the results of academic tests, interviews, and student interest. The calculation in determining this course is very simple, namely by adding up the values of each indicator and dividing them together so that an average value is obtained. If the value is fulfilled then the student is grouped based on their interests. This can lead to errors in decision making by the school due to subjective factors by prioritizing student interests and not based on the value of the academic test results. Therefore we need methods and algorithms to help make decisions well, the decision tree method. One algorithm that can be used is CART algorithm to classify majors with three indicators, namely Natural Sciences, Social Sciences and Religion. The results of this study indicate that the CART algorithm is able to predict correctly, from 360 data classified using the CART algorithm, it can be concluded that 71 data majoring in religion and correctly classified by CART. 144 data majoring in Natural Sciences, 119 data correctly classified and 24 data classified as IPS, and 1 data classified as religion. Of 146 data majoring in social studies, 129 were classified correctly, 16 data were classified as natural sciences. Therefore it can be concluded that CART algorithm has an 80% accuracy so that it can be used in decision making

Keywords— CART algorithm, Classification, Decision Tree, MAN 1 Inhil

Abstrak— MAN 1 Inhil merupakan sekolah yang menerapkan peraturan menteri untuk melakukan penentuan jurusan siswa di awal masuk yaitu pada kelas X. Penentuan jurusan dilakukan dengan mempertimbangkan beberapa indikator yaitu hasil tes akademik, wawancara, dan minat siswa. Perhitungan dalam penentuan jurusan ini sangat sederhana yaitu dengan menjumlahkan nilai setiap indikator dan dibagi keseluruhannya sehingga didapatkan nilai rata-rata. Jika nilai tersebut terpenuhi maka siswa tersebut dikelompokkan berdasarkan minatnya. Hal ini bisa menimbulkan kesalahan dalam pengambilan keputusan oleh pihak sekolah karena bersifat subjektif dengan mengutamakan minat siswa dan bukan berdasarkan nilai dari hasil tes akademik. Oleh karena itu butuh metode dan algoritma dalam membantu mengambil keputusan dengan baik yaitu metode decision tree. Salah satu algoritma yang bisa digunakan yaitu algoritma CART untuk melakukan pengklasifikasian jurusan dengan tiga indikator yaitu IPA, IPS dan Agama. Hasil penelitian ini menunjukkan bahwa algoritma CART mampu melakukan prediksi dengan benar.

Berdasarkan 360 data yang diklasifikasikan disimpulkan bahwa jurusan agama terdiri dari 71 yang diklasifikasikan secara benar. Jurusan IPA berjumlah 144 meliputi 119 diklasifikasikan secara benar dan 24 diklasifikasikan sebagai IPS dan 1 diklasifikasikan sebagai agama. Jurusan IPS berjumlah 146 meliputi 129 diklasifikasikan secara benar dan 16 diklasifikasikan sebagai IPA. Oleh karena itu dapat disimpulkan bahwa algoritma CART memiliki tingkat akurasi 80% sehingga bisa digunakan dalam pengambilan keputusan.

Kata Kunci— Algoritma CART, Decision Tree, Klasifikasi, MAN 1 Inhil

I. PENDAHULUAN

Saat ini penentuan jurusan pada tingkat SMA dan setingkatnya dimulai sejak siswa masuk yaitu mulai dari kelas X. Hal ini bertujuan agar siswa lebih mendalami bidang keilmuan yang diminatinya sejak awal masuk sekolah menenahi atas sehingga siswa bisa memfokuskan dirinya dalam mengembangkan kemampuan dan minat yang diminatinya. Pemilihan jurusan yang tidak tepat bisa saja merugikan siswa tersebut dan juga karirnya di masa yang akan datang [1].

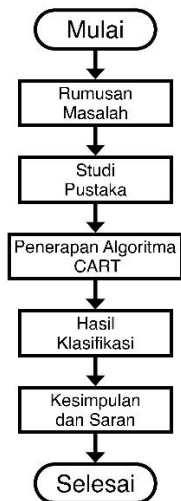
Salah satu sekolah yang menerapkan sistem penjurusan tersebut adalah sekolah MAN 1 di Indragiri Hilir (Inhil) dengan jurusan yang tersedia yaitu IPA, IPS, dan MAK (Agama). Saat ini, pemilihan jurusan di MAN 1 Inhil dilakukan pada saat calon siswa dinyatakan lulus masuk di sekolah ini. Sistem seperti ini sangat efektif dilakukan, jika sebelumnya siswa tersebut telah memiliki persiapan dan pengetahuan mengenai jurusan yang akan dipilih. Namun sebaliknya, jika siswa tersebut belum mengetahui arah kemampuan mengenai jurusan yang akan dipilih maka siswa tersebut akan kebingungan dalam memilih jurusan dan bisa berakibat pada salah ambil jurusan.

Indikator yang digunakan oleh pihak sekolah dalam menentukan jurusan siswa adalah hasil tes akademik, wawancara, dan minat siswa. Indikator tersebut dijumlahkan dan dibagi sehingga didapat nilai rata-rata. Jika nilai tersebut terpenuhi maka siswa tersebut dikelompokkan berdasarkan minatnya. Perhitungan seperti ini sangat sederhana. Hal ini bisa menimbulkan kesalahan dalam pengambilan keputusan oleh pihak sekolah karena bersifat subjektif yaitu dengan mengutamakan minat siswa bukan berdasarkan nilai dari hasil tes akademik.

Seringkali dijumpai dalam memilih jurusan, siswa hanya ikut-ikutan teman atau hanya memiliki informasi sedikit dari teman maupun oranglain mengenai jurusan yang dipilih. Oleh karena tidak memperhatikan nilai dan peminatan kurikulumnya sehingga menyebabkan salah jurusan dan berakibat putus sekolah di tengah jalan [2]. Terdapat beberapa hambatan yang dihadapi oleh siswa dalam memilih jurusan yaitu hambatan yang berasal dari dalam berupa kemampuan diri dan hambatan yang berasal dari luar yaitu permintaan atau paksaan orang tua dalam memilih jurusan yang mana dilatarbelakangi oleh pekerjaan di masa depan yang diharapkan oleh orang tua untuk anaknya dimasa yang akan datang [3].

Berdasarkan hasil angket yang disebarakan kepada siswa MAN 1 Inhil bahwa mayoritas siswa memperoleh jurusan berdasarkan hasil pilihan yang telah diisi pada formulir pemilihan jurusan. Selain itu, berdasarkan angket bahwa tingkat kepuasan siswa serta nilai akhir yang dihasilkanpun bervariasi, mulai dari hasil yang memuaskan hingga hasil yang kurang memuaskan. Oleh karena itu perlu adanya algoritma dalam menentukan atau memutuskan jurusan yang tepat sehingga diharapkan nilai akhir siswa mayoritas memuaskan. Metode yang bisa digunakan dalam pengambilan keputusan adalah decision tree. Salah satu algoritma yang sederhana dan kuat dalam pengambilan keputusan adalah algoritma CART [4]. Algoritma CART menghasilkan 87% tingkat akurasi data uji dan data latih [5] sehingga dapat digunakan dalam pengambilan keputusan oleh pihak sekolah dan memberikan informasi dalam menentukan jurusan yang benar [6].

II. METODE PENELITIAN



Gambar 1. Metodologi Penelitian

Metode Penelitian memberikan penjelasan tentang langkah-langkah, data, lokasi penelitian, metode evaluasi yang digunakan serta penjelasan terstruktur tentang algoritma atau metode dari penelitian yang dibahas.

Metode penelitian pada penelitian ini adalah decision tree, CART dan software WEKA yang digunakan untuk menguji hasil perhitungan manual. Metode tersebut akan dijelaskan dibawah ini.

A. Pohon Keputusan (Decission Tree)

Decission tree adalah metode pengambilan keputusan teknik klasifikasi dengan struktur flowchart yang mirip dengan pohon. Model yang dibentuk oleh metode ini ini sangat mudah dipahami sehingga menjadikan metode ini sangat umum dan populer [7]. Terdiri dari beberapa algoritma dalam membangun tree yaitu CART, ID3 dan C4.5. Penelitian ini menggunakan algoritma CART.

B. Algoritma CART (Classification And Regression Tree)

Merupakan algoritma yang umum dan banyak digunakan yang mengintegrasikan berbagai faktor dari sumber yang berbeda [8] untuk masalah klasifikasi dan regresi berdasarkan biner rekursif [9]. Algoritma ini tergolong dalam model nonparametric [10] yang tidak memerlukan bentuk fungsional dan telah terbukti menjadi alat ampuh untuk masalah prediksi dan klasifikasi [9]. CART mengandalkan partisi biner rekursif [11] dari data dasar pembangunan pohon regresi. Pohon didirikan dari dataset yang dikumpulkan di simpul pohon akar dan setiap node dibagi menjadi dua node turun menggunakan variabel pemisahan (salah satu variabel prediktor) . Pemilihan variabel pemisahan mencari penurunan variabilitas dalam-node dan untuk peningkatan variabilitas antara-node, dengan setiap partisi yang diperoleh menghasilkan pohon dengan variabilitas yang lebih sedikit daripada pohon sebelumnya.

Berdasarkan penelitian [6] menunjukkan bahwa berdasarkan nilai precision, nilai precision metode naive bayes lebih baik dibandingkan dengan metode yang lain dengan nilai 77,51%. Sedangkan berdasarkan nilai recall dan accuracy, decision tree lebih baik dibandingkan dengan metode yang lain dengan nilai recall sebesar 90,80% dan nilai accuracy sebesar 79,14% [1].

Tahapan-tahapan yang dilakukan pada algoritma CART ini sebagai berikut [12]–[14]: Pertama, lakukan penyusunan calon cabang (candidate split) terhadap seluruh variabel prediktor dengan lengkap. Daftar yang berisi calon cabang tersebut dinamakan calon cabang mutakhir. Kedua, lakukan penilaian seluruh calon cabang mutakhir dengan melakukan perhitungan nilai besaran kesesuaian $\Phi(s|t)$. Lakukan tahapan kedua sampai tidak adalagi noktah keputusan dan akan berhenti jika tidak ada lagi noktah keputusan. Untuk mencari kesesuaian $\Phi(s|t)$ dari calon cabang s pada noktah keputusan t menggunakan persamaan sebagai berikut:

$$\Phi(s|t) = 2P_L P_R Q(s|t) \dots \dots \dots (1)$$

$$\text{Dimana } Q(s|t) = \sum_{i=1}^{\text{Jumlah}} |P(j|tL) - P(j|tR)| \dots \dots \dots (2)$$

t_L = cabang kiri dari noktah keputusan t

t_R = cabang kanan dari noktah keputusan t

$$\text{dimana } P_L = \frac{\text{jumlah catatan pada calon kiri } t_L}{\text{jumlah catatan pada data latihan}} \dots \dots \dots (3)$$

$$P(j|t_L) = \frac{j \text{ calon cabang kiri } t_L}{\text{noktah keputusan } t} \dots \dots \dots (4)$$

$$P_L = \frac{\text{jumlah catatan pada calon kiri } t_R}{\text{jumlah catatan pada data latihan}} \dots \dots \dots (5)$$

$$P(j|t_R) = \frac{j \text{ calon cabang kiri } t_R}{\text{noktah keputusan } t} \dots \dots \dots (6)$$

C. Weka Tools

WEKA merupakan salah satu tools yang mampu melakukan perbandingan beberapa algoritma *machine learning* yang digunakan dalam pengaplikasian pada permasalahan data mining. Tools ini bisa dikembangkan oleh siapa saja karena bersifat open source dan dikembangkan pertamakali oleh University of Waikato di Selandia Baru [15] sebagai penambangan data. Pada tools ini, data-data di uji prosedur-prosedurnya untuk melakukan eksplorasi dan permodelan guna menghasilkan hubungan yang tersembunyi dari data tersebut [16].

III. HASIL DAN PEMBAHASAN

Data inputan berjumlah 439 record yang berasal dari data rekapitulasi pendaftaran dan seleksi siswa kelas X. Adapun target dan klasifikasi dalam penentuan jurusan menggunakan atribut seperti Jenis Kelamin, Nilai Tes Siswa berupa Nilai Matematika, IPA, IPS dan Agama, ditunjukkan pada TABEL I.

TABLE I. DATA REKAPUTULASI (POTONGAN DATA)

No.	Nama	JK	MTK	IPS	...	Jurusan
1.	Muh Ariansyah	LK	85	75	...	Agama
2.	Robiansyah	LK	70	85	...	Agama
3.	M.Hidayat	PR	90	70	...	Agama
...
439	Muhammad Ridho	LK	0	0	0	Tidak Lulus

A. Preprocessing Data

Tahapan ini bertujuan untuk memilih data dan atribut yang sesuai dan lengkap. Awalnya atribut berjumlah 11, namun setelah dilakukan penyeleksian dan pembersihan data menjadi 8 atribut yang dapat digunakan. Atribut seperti peringkat, nilai rata-rata, dan keterangan lulus/tidak lulus dihapus dan tidak digunakan. Selain itu, data yang digunakan berjumlah 360 record dari yang sebelumnya berjumlah 439 record data. Data hasil pada tahap ini di tunjukkan pada TABEL II.

TABLE II. DATA HASIL PREPROCESSING (POTONGAN DATA)

No.	Nama	JK	MTK	IPS	...	Jurusan
1.	Muh Ariansyah	LK	85	75	...	Agama
2.	Robiansyah	LK	70	85	...	Agama
3.	M.Hidayat	PR	90	70	...	Agama
...
360	Selfi Nurdiati	PR	25	25	...	IPS

B. Data Latih Dan Data Uji

Berdasarkan data yang dihasilkan pada TABEL II selanjutnya digunakan 15 data latih pada TABEL III. Selanjutnya akan dilakukan pengujian data dengan mencari label atau jurusannya berdasarkan perhitungan algoritma CART. Data uji pada penelitian ini di tunjukkan pada TABEL IV.

TABLE III. DATA LATIH

No.	Nama	JK	MTK	IPA	...	JUR
1	Muh Weldi	LK	70	85	...	AGAMA
2	Cici Alia	PR	60	90	...	AGAMA
3	Syaifira A	PR	65	90	...	IPA
4	Nia Zufiani	PR	65	85	...	IPA
5	Abdul Rahman	LK	60	80	...	IPA
6	Muh Rizki	LK	60	61,43	...	AGAMA
7	Muh Arya	LK	55	70	...	AGAMA
8	Meilani S	PR	45	52,86	...	AGAMA
9	Widia Putri	PR	50	44,29	...	IPA
10	Ihsan D	LK	60	47,14	...	IPA
11	Meiry M	PR	45	52,86	...	IPS
12	Windi A	PR	35	52,86	...	IPS
13	Puspa Y	PR	30	47,14	...	IPS
14	Suandi	PR	35	52,86	...	IPS
15	Tirta	PR	45	47,14	...	IPS

TABLE IV. DATA UJI

No.	Na ma	JK	MTK	IP A	IPS	AGAMA	JUR
1	Muh Reza	LK	75	75	58,71	82,98	?

C. Penerapan Metode Cart

Tahapan pertama pada Algoritma CART ini adalah menentukan calon noktah keputusan dengan cara menyusun calon cabang (candidate split) yang dilakukan terhadap seluruh atribut prediktor yaitu Jenis Kelamin, Rata-rata nilai IPA, Rata-rata nilai IPS, Rata-rata nilai Agama, Rata-rata nilai Matematika, sehingga terbentuk daftar calon cabang mutakhir seperti yang ditunjukkan pada TABEL V.

TABLE V. DAFTAR CALON MUTAKHIR

Nomor calon cabang	Calon cabang kiri (t_L)	Calon cabang kanan (t_R)
1	Jenis Kelamin = Laki-laki	Jenis Kelamin = Perempuan
2	Rata-rata nilai IPA ≤ 50	Rata-rata nilai IPA > 50
3	Rata-rata nilai IPA ≤ 80	Rata-rata nilai IPA > 80
4	Rata-rata nilai IPS ≤ 50	Rata-rata nilai IPS > 50
5	Rata-rata nilai IPS ≤ 60	Rata-rata nilai IPS > 60
6	Rata-rata nilai Agama ≤ 60	Rata-rata nilai Agama > 60
7	Rata-rata nilai Agama ≤ 80	Rata-rata nilai Agama > 80
8	Rata-rata nilai Matematika ≤ 30	Rata-rata nilai Matematika > 30
9	Rata-rata nilai Matematika ≤ 50	Rata-rata nilai Matematika > 50

Data yang telah diolah pada Tabel 5 selanjutnya di seleksi berdasarkan data latih sehingga menghasilkan data calon mutakhir pada TABEL VI.

TABLE VI. DATA CALON MUTAKHIR

Nomor Calon Cabang	Calon Cabang Kiri (tL)	Calon Cabang Kanan (tR)
1	8	7
2	7	8
3	11	4
4	4	11
5	9	6
6	5	10
7	9	6
8	4	11
9	7	8

Setelah mendapatkan calon cabang kiri dan kanan, selanjutnya menghitung *candidate split purity left* (P_L) dan *purity right* (P_R) dengan menggunakan persamaan (3) dan (5) sehingga menghasilkan nilai yang ada pada TABEL VII.

TABLE VII. NILAI P_L DAN P_R

No	P_L (Purity Left)	P_R (Purity Right)
1	8:15 = 0,5333	7:15 = 0,4667
2	7:15 = 0,4667	8:15 = 0,5333
3	11:15 = 0,7333	4:15 = 0,2667
4	4:15 = 0,2667	11:15 = 0,7333
5	9:15 = 0,6000	6:15 = 0,4000
6	5:15 = 0,3333	10:15 = 0,6667
7	9:15 = 0,6000	6:15 = 0,4000
8	4:15 = 0,2667	11:15 = 0,7333
9	7:15 = 0,4667	8:15 = 0,5333

Selanjutnya melakukan penandaan pada label kelas menggunakan persamaan (4) dan (6) untuk jurusan IPA, IPS dan AGAMA. Sehingga menghasilkan perhitungan pada TABEL VIII.

TABLE VIII. NILAI (j/T_L) dan $P(j/T_R)$

No	Jurusan	$P(j/T_L)$	$P(j/T_R)$
1	IPA	3/8 = 0,3750	2/7 = 0,2857
	IPS	2/8 = 0,2500	3/7 = 0,4286
	AGAMA	3/8 = 0,3750	2/7 = 0,2857
2	IPA	1/7 = 0,1429	4/8 = 0,5000
	IPS	5/7 = 0,7143	0/8 = 0,0000
	AGAMA	1/7 = 0,1429	4/8 = 0,5000
3	IPA	3/11 = 0,2727	2/4 = 0,5000
	IPS	5/11 = 0,4545	0/4 = 0,0000
	AGAMA	3/11 = 0,2727	2/4 = 0,5000
4	IPA	2/4 = 0,5000	3/11 = 0,2727
	IPS	2/4 = 0,5000	3/11 = 0,2727
	AGAMA	0/4 = 0,0000	5/11 = 0,4545
5	IPA	3/9 = 0,3333	2/6 = 0,3333
	IPS	5/9 = 0,5556	5/9 = 0,5556
	AGAMA	1/9 = 0,1111	4/6 = 0,6667
6	IPA	0/5 = 0,0000	5/10 = 0,5000
	IPS	5/5 = 1,0000	0/10 = 0,0000

No	Jurusan	$P(j/T_L)$	$P(j/T_R)$
7	AGAMA	0/5 = 0,0000	5/10 = 0,5000
	IPA	4/9 = 0,4444	1/6 = 0,1667
	IPS	5/9 = 0,5556	0/6 = 0,0000
8	AGAMA	0/9 = 0,0000	5/6 = 0,8333
	IPA	0/4 = 0,0000	5/11 = 0,4545
	IPS	4/4 = 1,0000	1/11 = 0,0909
9	AGAMA	0/4 = 0,0000	5/11 = 0,4545
	IPA	1/7 = 0,1429	4/8 = 0,5000
	IPS	5/7 = 0,7143	0/8 = 0,0000
	AGAMA	1/7 = 0,1429	4/8 = 0,5000

Selanjutnya persamaan 2 digunakan untuk mencari nilai $Q(s/t)$ dengan cara mencari selisih nilai $P(j/T_L)$ dan $P(j/T_R)$ pada setiap label atau jurusan dan nilai yang dihasilkan bersifat *absolute* lalu selanjutnya hasil selisihnya dijumlahkan. Sehingga menghasilkan nilai $Q(s/t)$ seperti ditunjukkan pada TABEL IX.

Berdasarkan TABEL IX, selanjutnya dihitung nilai kesesuaian (*goodness*) calon cabang $\phi(s/t)$ menggunakan Persamaan 2 dengan mengalikan nilai $2P_L P_R$ dan $Q(s/t)$. Nilai $2P_L P_R$ didapat dari $(2P_L \times P_R)$ kalikan 2. Hasil perhitungannya ditunjukkan pada TABEL X.

Calon cabang tertinggi berada pada nomor calon cabang ke 6 yaitu 0,8889 berdasarkan TABEL X diatas. Nilai yang dihasilkan pada calon cabang kiri Rata-rata nilai Agama ≤ 60 dan calon cabang kanan Rata-rata nilai Agama >60 sehingga calon cabang inilah yang dipilih sebagai root node pada tahap ini. Namun, tetap dilakukan perhitungan kembali sampai memiliki noktah keputusan dengan iterasi selanjutnya, setelah menghilangkan no 6 yang terpilih sebelumnya. Hal ini akan menghasilkan jumlah kesesuaian hasil prediksi jurusan yang memilih jurusan IPA dan IPS dan Agama. Perhitungan iterasi terakhir ditunjukkan pada Gambar I. yang menghasilkan pohon keputusan.

TABLE IX. DATA PERHITUNGAN $Q(s/t)$

No	Jurusan	$P(j/T_L)$	$P(j/T_R)$	Selisih (Absolute)	$Q(s/t)$
1	IPA	0,375	0,2857	0,0893	0,3571
	IPS	0,25	0,4286	0,1786	
	AGAMA	0,375	0,2857	0,0893	
2	IPA	0,1429	0,5	0,3571	1,4286
	IPS	0,7143	0	0,7143	
	AGAMA	0,1429	0,5	0,3571	
3	IPA	0,2727	0,5	0,2273	0,9091
	IPS	0,4545	0	0,4545	
	AGAMA	0,2727	0,5	0,2273	
4	IPA	0,5	0,2727	0,2273	0,9091
	IPS	0,5	0,2727	0,2273	
	AGAMA	0	0,4545	0,4545	
5	IPA	0,3333	0,3333	0	1,1111
	IPS	0,5556	0	0,5556	
	AGAMA	0,1111	0,6667	0,5556	
6	IPA	0	0,5	0,5	2
	IPS	1	0	1	
	AGAMA	0	0,5	0,5	

No	Jurusan	P(j T _L)	P(j T _R)	Selisih (Absolute)	Q(s/t)
7	IPA	0,4444	0,1667	0,2778	1,6667
	IPS	0,5556	0	0,5556	
	AGAMA	0	0,8333	0,8333	
8	IPA	0	0,4545	0,4545	1,8182
	IPS	1	0,0909	0,9091	
	AGAMA	0	0,4545	0,4545	
9	IPA	0,1429	0,5	0,3571	1,4286
	IPS	0,7143	0	0,7143	
	AGAMA	0,1429	0,5	0,3571	

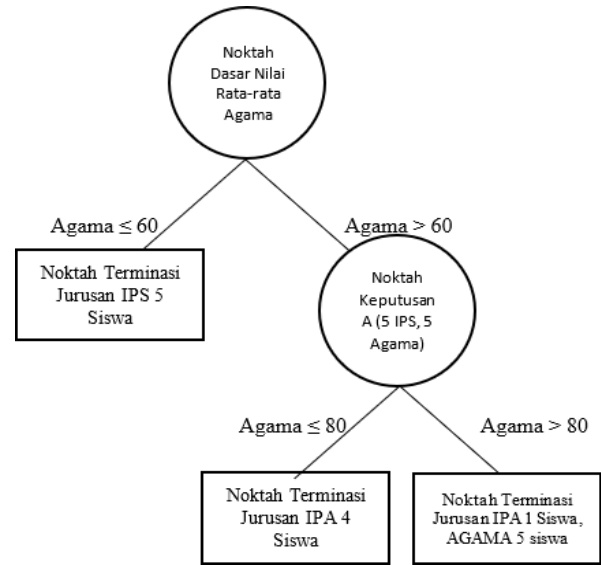
TABLE X. DATA PERHITUNGAN Q (s/t)

No	2P _L P _R	Q(s/t)	φ(s/t)
1	0,4978	0,3571	0,1778
2	0,4978	1,4286	0,7111
3	0,3911	0,9091	0,3556
4	0,3911	0,9091	0,3556
5	0,48	1,1111	0,5333
6	0,4444	2	0,8889
7	0,48	1,6667	0,8
8	0,3911	1,8182	0,7111
9	0,4978	1,4286	0,7111

Berdasarkan TABEL IX, selanjutnya dihitung nilai kesesuaian (goodness) calon cabang φ(s/t) menggunakan Persamaan 2 dengan mengalikan nilai 2P_LP_R dan Q(s/t). Nilai 2P_LP_R didapat dari (2P_L x P_R) kalikan 2. Hasil perhitungannya ditunjukkan pada TABEL X.

Calon cabang tertinggi berada pada nomor calon cabang ke 6 yaitu 0,8889 berdasarkan TABEL X diatas. Nilai yang dihasilkan pada calon cabang kiri Rata-rata nilai Agama ≤ 60 dan calon cabang kanan Rata-rata nilai Agama >60 sehingga calon cabang inilah yang dipilih sebagai root node pada tahap ini. Namun, tetap dilakukan perhitungan kembali sampai memiliki noktah keputusan dengan iterasi selanjutnya, setelah menghilangkan no 6 yang terpilih sebelumnya. Hal ini akan menghasilkan jumlah kesesuaian hasil prediksi jurusan yang memilih jurusan IPA dan IPS dan Agama. Perhitungan iterasi terakhir ditunjukkan pada Gambar I. yang menghasilkan pohon keputusan.

Iterasi selanjutnya adalah Calon cabang nomor 7 yang memiliki nilai tertinggi. Nilai yang dihasilkan adalah calon cabang kiri rata-rata nilai agama ≤ 80 dan calon cabang kanan > 80 sehingga dijadikan cabang kedua. Masing-masing cabang memiliki keterangan yang berbeda. Pada cabang dengan nilai rata-rata agama ≤ 80 memiliki keterangan yang homogen maka proses pencabangan dihentikan, untuk cabang dengan nilai rata-rata aga- ma > 80 memiliki keterangan belum homogen tetapi sudah memenuhi dari dari seluruh data latih yaitu 15 data, maka proses pencabangan juga dihentikan. Selanjutnya dilakukan pengklasifikasian data uji pada TABEL XI.



Gambar 2. Hasil pohon keputusan akhir

TABLE XI. DATA UJI

Nama	MTK	IPA	IPS	Agama	Jurusan	Prediksi
Muh Reza	75	75	58,71	82,98	Agama	?

Berdasarkan kesimpulan rule, rata-rata nilai agama merupakan nilai yang diprioritaskan terlebih dahulu untuk menentukan jurusan. Siswa yang bernama Muhammad Muh Reza tidak memenuhi syarat dikarenakan rata-rata nilai agama ≤ 60 sehingga tidak memenuhi masuk jurusan IPS, rata-rata nilai agama ≤ 80 sehingga tidak memenuhi masuk jurusan IPA, rata-rata nilai agama > 80 sehingga memenuhi masuk jurusan Agama. Sehingga disimpulkan bahwa jurusan yang cocok untuk Muhammad Muh Reza adalah Agama. Hasilnya ditunjukkan pada TABEL XII.

TABLE XII. DATA UJI PREDIKSI

Nama	MTK	IPA	IPS	Agama	Jurusan	Prediksi
Muh Reza	75	75	58,71	82,98	Agama	Agama

D. Pengujian Tools Weka

Menggunakan 15 data maka dilakukan pengujian algoritma Cart pada tools Weka. Hasilnya sebagai berikut :

AGAMA <59.72: IPS (5/0) AGAMA ≥ 59.72

— AGAMA <78.38: IPA (4/0)

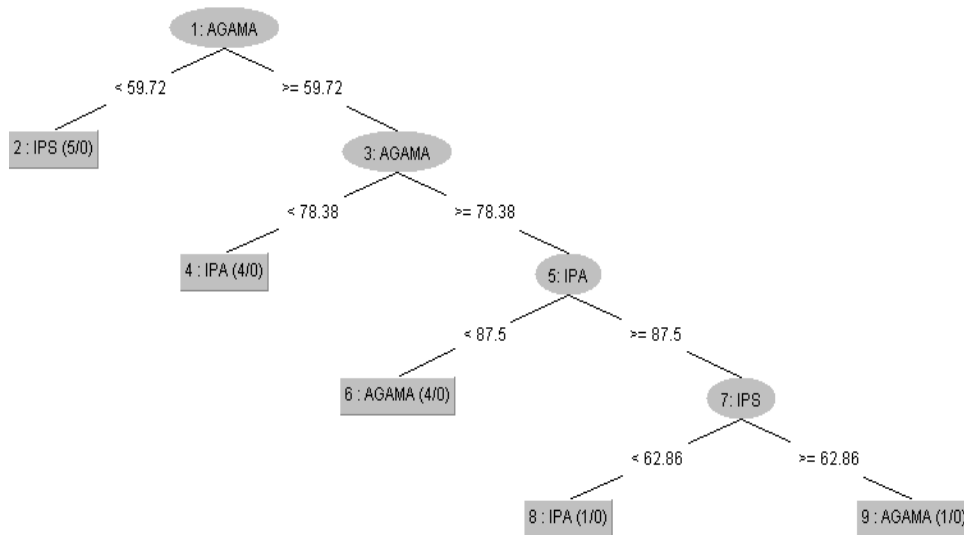
— AGAMA ≥ 78.38

— — IPA <87.5: AGAMA (4/0)

— — — IPA ≥ 87.5

— — — — IPS <62.86: IPA (1/0)

— — — — IPS ≥ 62.86: AGAMA (1/0)



Gambar 3. Pohon Visualisasi dengan Weka

Pohon visualisasi Algoritma CART dapat dilihat pada Gambar II. dan hasil prediksi dari kesimpulan rule ditunjukkan pada TABEL XIII. Sedangkan Uji coba tingkat akurasi

ditunjukkan pada TABEL XIV dengan akurasi yang tergolong tinggi yaitu 80%.

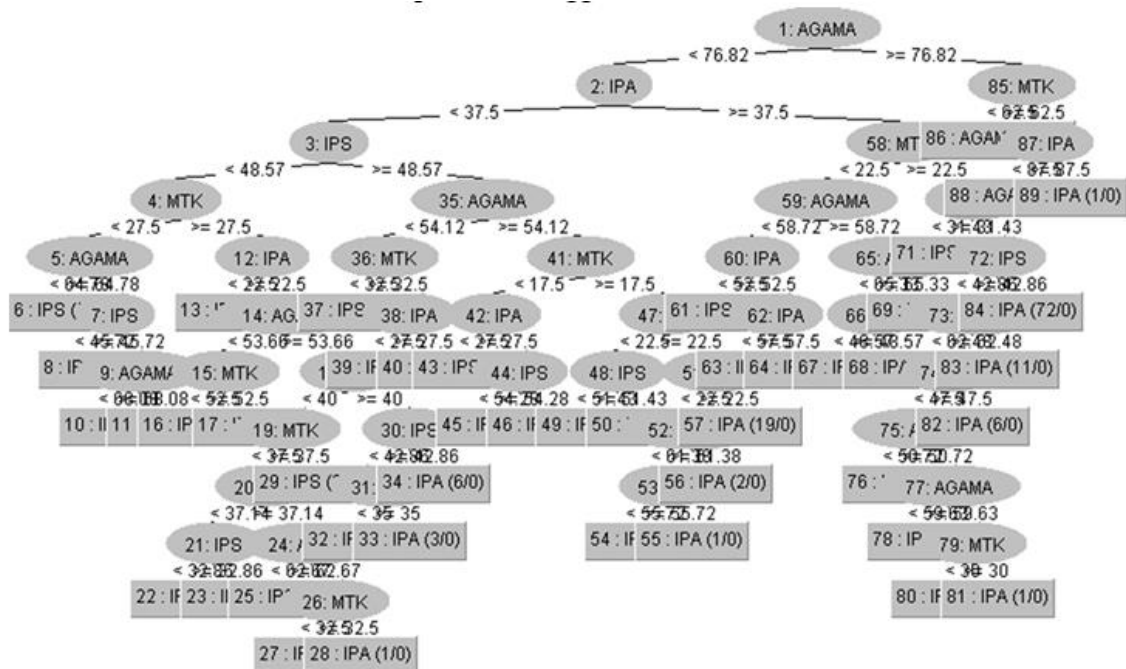
TABLE XIII. HASIL PREDIKSI DATA UJI

No	NAMA	IPA	IPS	Agama	MTK	Prediksi	Jurusan
1	Syaifra	90.0	58.57	84.26	65.0	AGAMA	IPA
2	Meiry M	45.0	52.86	51.18	15.0	IPS	IPS
3	Cici A	90.0	67.14	94.74	60.0	AGAMA	AGAMA
4	Nia Z	85.0	75.71	65.15	65.0	IPA	IPA
5	Suandi	35.0	52.86	48.05	25.0	IPS	IPS
6	Muhammad Rizqi A	60.0	61.43	97.87	60.0	AGAMA	AGAMA
7	Wilia Permata A	50.0	44.29	72.5	80.0	IPA	IPA
8	Windi A	35.0	52.86	54.49	20.0	IPS	IPS
9	Meileny S	45.0	52.86	89.04	65.0	IPS	AGAMA
10	Ihsan D	60.0	47.14	64.96	45.0	IPA	IPA
11	Tirta A Pratama	45.0	47.14	46.76	20.0	IPS	IPS
12	M Arya Y	55.0	70.0	84.08	45.0	IPS	AGAMA
13	Abdul R Sidiq	80.0	70.0	72.68	60.0	IPA	IPA
14	Puspa Y	30.0	47.14	45.11	40.0	IPS	IPS

TABLE XIV. NILAI AKURASI ALGORITMA CART)

No.	Jenis	Persentase
1	Akurasi	80%
2	Precision	0,794
3	Recall	0,8

Berdasarkan node akar yang telah diperoleh terjadi perbedaan antara perhitungan manual dan tools Weka. Pada perhitungan manual diperoleh Nilai Agama ≤ 60 dan Nilai Agama > 60 . Perhitungan *Tools Weka* diperoleh Nilai Agama > 59.72 dan Nilai Agama ≥ 59.72 . Perbedaan ini terjadi karena ketika pembentukan Calon cabang mutakhir pada perhitungan manual yang terdapat pada TABEL V berbeda dengan cara perhitungan Tools Weka, karena pembentukan Calon cabang mutakhir ini yang bersifat subjektif.



Gambar 4. Pohon keputusan menggunakan Tools Weka

Tahapan selanjutnya melakukan Uji coba seluruh data dengan *Tools Weka*. Sehingga terbentuklah Node dari pohon keputusan sebagai berikut : Jumlah Node Akar yaitu 1 node, jumlah Node Internal yaitu 43 node, jumlah Lengan yaitu 89 lengan dan jumlah Node Daun (Internal) yaitu 49 node. Untuk melihat Pohon Visualisasi keseluruhan data ditunjukkan pada Gambar III. dan untuk melihat hasil klasifikasi dari kesimpulan rule dan pohon yang terbentuk dapat dilihat pada TABEL XV.

TABLE XV. HASIL KESIMPULAN RULE

No	NAMA	IPA	IPS	Agama	MTK	Prediksi	Jurusan
1	Himantul	45.0	45.0	78.57	58.71	IPA	IPA
2	Muh Andri	30.0	35.0	47.14	60.0	IPS	IPA
3	Murniyanti	50.0	50.0	38.57	58.9	IPA	IPA
4	Nursiffa Z	40.0	30.0	41.43	62.76	IPS	IPA
5	Muh Yusof	25.0	35.0	47.14	64.96	IPS	IPA
6	Aura Putri	20.0	55.0	70.0	67.35	IPA	IPA
7	Nabiala P	30.0	40.0	41.43	64.78	IPS	IPA
8	Muh. Farid G	25.0	40.0	61.43	71.03	IPA	IPA
9	Wahyudi Saputra	20.0	45.0	58.57	67.9	IPA	IPA
10	Feni Ahmad	30.0	45.0	44.29	59.26	IPA	IPA
11	Tresna P	35.0	40.0	47.14	70.66	IPA	IPA
12	Sarmina	35.0	40.0	41.43	68.09	IPS	IPA
13	Tanzani	45.0	70.0	58.57	75.26	IPA	IPA
...
360	Muh Akbar H	20.0	20.0	52.86	82.98	AGAMA	AGAMA

IV. KESIMPULAN

Penelitian ini menghasilkan akurasi sebesar 88,61% dan nilai *precision* dan *recall* masing-masing adalah 0,903 dan 0,905, angka tersebut jika dilihat dari tingkat diagnose kurva ROC. Kedua algoritma ini termasuk kedalam kategori *Excellent Classification* dalam penentuan jurusan menggunakan algoritma CART. Sehingga algoritma CART dan C4.5 sangat direkomendasikan untuk mengklasifikasikan penjurusan, dengan catatan jika atribut yang dimiliki dalam penentuan jurusan mayoritas kedalam bentuk numerik maka algoritma CART lebih dianjurkan diterapkan dalam melakukan klasifikasi penjurusan siswa, karena berdasarkan perhitungannya, algoritma CART memang lebih baik untuk menghadapi data-data yang bersifat numerik.

REFERENSI

- [1] Y. S. Nugroho, "Klasifikasi dan Klustering Penjurusan Siswa SMA Negeri 3 Boyolali," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 1, no. 1, p. 1, 2015.
- [2] F. Rini, N. Kahar, and Juliana, "Penerapan Algoritma K-Means Pada Pengelompokan Data Siswa Baru Berdasarkan Jurusan Di Smk Negeri 1 Kota Jambi Berbasis Web ", *Semin. Nas. APTIKOM*, pp. 94–99, 2016.
- [3] I. M. Prabowo and Subiyanto, "Sistem Rekomendasi Penjurusan Sekolah Menengah Kejuruan Dengan Algoritma C4.5," *J. Kependidikan*, vol. 1, no. 1, pp. 139–149, 2017.
- [4] F. E. Pratiwi, F. E. Pratiwi, and I. Zain, "Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara," *J. Sains dan Seni ITS*, vol. 3, no. 1, pp. D54–D59, 2014.
- [5] D. Mutiah, M. Bettiza, and M. R. Rathomi, "Penerapan Algoritma Classification And Regression Tree (CART) Untuk Pemilihan Jurusan SMA (STUDI KASUS : SMA Negeri 3 Bintang Kabupaten Bintang)," pp. 1–13, 2015.
- [6] Y. S. Nugroho and N. S. Haryati, "Klasifikasi dan Klustering Penjurusan Siswa SMA Negeri 3 Boyolali," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 1, no. 1, pp. 1–6, 2015.

- [7] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 1, pp. 213–219, 2014.
- [8] J. Han, K. Mao, T. Xu, J. Guo, Z. Zuo, and C. Gao, "A soil moisture estimation framework based on the CART algorithm and its application in China," *J. Hydrol.*, vol. 563, pp. 65–75, 2018.
- [9] S. M. Hamze-Ziabari and T. Bakhshpoori, "Improving the prediction of ground motion parameters based on an efficient bagging ensemble model of M5' and CART algorithms," *Appl. Soft Comput. J.*, vol. 68, pp. 147–161, 2018.
- [10] M. M. Ghiasi, S. Zendejboudi, and A. A. Mohsenipour, "Decision tree-based diagnosis of coronary artery disease: CART model," *Comput. Methods Programs Biomed.*, vol. 192, p. 105400, 2020.
- [11] L. L. Doove, S. Van Buuren, and E. Dusseldorp, "Recursive partitioning for missing data imputation in the presence of interaction effects," *Comput. Stat. Data Anal.*, vol. 72, pp. 92–104, 2014.
- [12] Mardiani, "Penerapan Klasifikasi Dengan Algoritma Cart Untuk Prediksi Kuliah Bagi Mahasiswa Baru," *Semin. Nas. Apl. Teknol. Inf.*, vol. 2012, no. Snati, pp. 15–16, 2012.
- [13] F. E. dan I. Z. Pratiwi, "Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara," *J. Sains Dan Seni Pomits*, vol. 3, no. 1, pp. 2337–3520, 2014.
- [14] Hariati, M. Wati, and B. Cahyono, "Penerapan Algoritma C4.5 Decision Tree pada Penentuan Penerima Program Bantuan Pemerintah Daerah Kabupaten Kutai Kartanegara," *Jurti*, vol. 2, no. 1, pp. 27–36, 2018.
- [15] A. Naik and L. Samant, "Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime," *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 662–668, 2016.
- [16] emadwiandr, "Sistem Rekomendasi Penjurusan Sekolah Menengah Kejuruan Dengan Algoritma C4.5," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.