

# Implementasi Data Mining Untuk Pengembangan Sistem Rekomendasi Pemilihan SMK Dengan Menggunakan Algoritma Cart

Ade Septi Binjori

Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: adeseptibinjori1234@gmail.com

## Abstrak

Rekomendasi pemilihan Sekolah Menengah Kejuruan (SMK) adalah prosedur yang dilakukan untuk memberikan informasi pendukung kepada siswa yang tamatan SMP atau MTS dalam pemilihan Sekolah Menengah Kejuruan (SMK) untuk melanjutkan jenjang pendidikan SMP atau MTS ke jenjang pendidikan Sekolah Menengah Kejuruan (SMK). Sebagian besar negara memiliki sistem pendidikan formal yang umumnya bersifat wajib. Pada setiap tahun, siswa akan mencari sekolah lanjutan yang cocok untuk dirinya. Persoalan muncul ketika terdapat banyak pilihan sekolah yang memberikan beragam tawaran dan pilihan kepada calon siswanya. Oleh karena itu diperlukan sebuah sistem yang tidak hanya memberikan informasi tentang sekolah saja, melainkan juga mampu memberikan rekomendasi sekolah berdasarkan kriteria yang diinginkan dari masing-masing calon siswa. Pohon keputusan yang dihasilkan CART merupakan pohon biner yang memiliki nilai atribut dengan memilih percabangan yang paling optimal dalam menghitung setiap variabel-variabel. Prinsip dari pohon klafikasi adalah memilah seluruh amatan menjadi dua gugus amatan menjadi dua gugus amatan berikutnya sehingga diperoleh jumlah amatan minimum pada tiap gugus amatan.

**Kata Kunci:** Rekomendasi, Data Mining, Metode CART.

## Abstract

Recommendation for choosing a Vocational High School (SMK) is a procedure that is carried out to provide supporting information for students graduating from SMP or MTS in the selection of Vocational High Schools (SMK) for advanced junior high school or MTS education to Vocational High School (SMK) education. Most countries have formal education systems that are usually needed. Every year, students will look for schools. Problems arise and contain many choices that provide a variety of offers and choices to prospective students. Therefore we need a system that not only provides information about schools, and also is not able to contribute to the appropriate school of each prospective student. The decision tree that produces CART is a binary tree that has attribute values by selecting the most optimal branching in each variable. The principle of the clafication tree is to separate all observations into two observation groups into the next two observation groups, that is, the minimum number of tantrums when the observation group.

**Keywords:** Recommendation, Data Mining, CART Method.

## 1. PENDAHULUAN

Dalam penelitian ini masalah yang dihadapi Dinas Pendidikan Tapanuli Utara mengambil keputusan dalam hasil rekomendasi pemilihan Sekolah Menengah Kejuruan (SMK) bagi tamatan jenjang pendidikan SMP atau MTS yaitu membuat pertimbangan kepada siswa untuk memilih salah satu Sekolah Menengah Kejuruan (SMK) di Kabupaten Tapanuli berdasarkan kriteria-kriteria yang dibutuhkan. Data mining merupakan suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Salah satu algoritma data mining yaitu algoritma CART (Classification and Regression Tree) [1].

CART (Classification and Regression Tree) adalah salah satu metode atau algoritma dari salah satu teknik eksplorasi data yaitu teknik pohon keputusan. CART (Classification and Regression Tree) dikembangkan untuk melakukan analisis klasifikasi pada peubah respon baik yang nominal, ordinal, maupun kontinu. CART (Classification and Regression Tree) menghasilkan suatu pohon klasifikasi jika peubah responnya kategorik dan menghasilkan pohon regresi jika peubah responnya kontinu. Sebelumnya Algoritma CART sudah pernah dilakukan penelitian dengan judul Klasifikasi Pengangguran Terbuka Menggunakan Algoritma CART (Classification and Regression Tree) Diprovinsi Sulawesi Utara[2]. Nilai tingkat kesalahan yang paling kecil pada pohon klasifikasi yang diasilkan akan cenderung membuat pohon ini digunakan untuk memperkirakan respon. Oleh karena itu penulis mencoba untuk menyelesaikan permasalahan yang ada di Dinas Pendidikan Dan Kebudayaan Tapanuli Utara dengan salah satu teknik yang digunakan dalam pengolahan data yaitu data mining dengan menerapkan algoritma CART (Classification and Regression Tree) dengan memanfaatkan aplikasi Weka sebagai alat bantu untuk menerapkan rules yang akan dibangun menggunakan data mining dengan menerapkan algoritma CART (Classification and Regression Tree).

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mini

Data Mining adalah Serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data [3].

## 2.2 Rekomendasi

Rekomendasi adalah bentuk komunikasi sekaligus promosi tidak langsung dilakukan oleh para konsumen yang sudah pernah membeli produk atau jasa yang kemudian menceritakan berbagai pengalaman yang berkaitan dengan produk atau jasa tersebut kepada orang lain[4].

## 2.3 Pemilihan Sekolah

Secara umum pemilihan sekolah merupakan suatu hal yang sangat penting dikarenakan pilihan sekolah akan mempengaruhi pendidikan masa depan. Di era globalisasi saat ini membuat keputusan untuk memilih sekolah yang tepat tidak mudah. Selain jumlah sekolah yang banyak, setiap sekolah juga memberi beragam tawaran dan pilihan kepada para calon siswanya [5].

## 2.4 Sekolah Menengah Kejuruan

Sekolah Menengah kejuruan (SMK) adalah bagian dari sistem pendidikan yang mempersiapkan seseorang agar lebih mampu bekerja pada suatu kelompok pekerjaan atau satu bidang pekerjaan daripada bidang-bidang pekerjaan lainnya. Dengan pengertian bahwa setiap bidang studi adalah pendidikan kejuruan sepanjang bidang studi tersebut dipelajari lebih mendalam dan kedalaman tersebut dimaksudkan sebagai bekal memasuki dunia kerja Evans dalam Djojonegoro (1999) [6].

## 2.5 Algoritma Cart

CART (Classification and Regression Trees) adalah salah satu metode atau algoritma dari salah satu teknik eksplorasi data decision tree. Metode ini dikembangkan oleh Leo Breiman, Jerome H. Friedman, Richard S. Olshen dan Charles J[2].

Langkah – langkah Metode Cart :

1. Susunlah calon cabang (candidate split) yang dilakukan terhadap seluruh variabel prediktor. Daftar yang berisi calon cabang disebut calon cabang mutakhir.
2. Berikan penilaian keseluruhan calon cabang mutakhir dengan menghitung besaran  $\phi(s|t)$
3. Tentukan cabang yang memiliki kesesuaian  $\phi(s|t)$ . Setelah noktah keputusan tidak ada lagi, Algoritma Cart dihentikan. Kesesuaian (*goodness*)  $\phi(s|t)$  dari calon cabang  $s$  pada noktah keputusan  $t$ , didefinisikan sebagai persamaan-persamaan berikut :

$$\phi(s|t) = 2P_L P_R Q(s|t) \tag{1}$$

$$Q(s|t) = \sum_{j \neq 1}^{\text{jumlah kategori}} |P(j|t_L) - P(j|t_R)| \tag{2}$$

$t_L$  = cabang kiri dari noktah keputusan  $t$   
 $t_R$  = cabang kanan dari noktah keputusan  $t$

$$P_L = \frac{\text{calon cabang kiri}(t_L)}{\text{data latihan}} \tag{3}$$

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

$$P_R = \frac{\text{calon cabang kanan}(t_R)}{\text{data latihan}} \tag{5}$$

$$P(j|t_R) = \frac{j \text{ pada calon cabang kanan}(t_R)}{\text{noktah keputusan } t} \tag{6}$$

## 3. HASIL DAN PEMBAHASAN

Setiap tahun pendidikan di Indonesia semakin maju dan berkembang seiring dengan semakin bertambahnya jumlah sekolah atau lembaga pendidikan. Sementara itu berdasarkan data dari Dinas Pendidikan Kabupaten Tapanuli Utara, bertambahnya jumlah sekolah yang berdiri juga diimbangi dengan jumlah siswa yang semakin meningkat pada tiap tahunnya. Sekolah-sekolah ini tidak jarang didirikan dalam jarak yang berdekatan dalam satu wilayah seperti halnya dengan sekolah menengah kejuruan di kota Tapanuli Utara. Dewasa ini, sekolah tersebut tidak hanya bersaing dalam hal akademik melalui prestasi sekolah maupun akreditasi namun juga bersaing dalam mutu pendidikan yang baik untuk menarik minat calon siswa. Namun, permasalahan yang sering dihadapi yaitu banyak faktor-faktor yang akhirnya melatar belakangi siswa untuk memilih sebuah calon sekolah baik variabel eksternal sekolah maupun internal. Sehingga siswa dituntut cerdas dalam menentukan calon sekolah terbaik dan sesuai dengan kriterianya. Penelitian ini dilakukan dalam rangka membantu siswa untuk memilih calon sekolah menengah kejuruan terbaik sesuai dengan kriteria-kriteria yang diinginkan. Penerapan metode CART digunakan dalam penelitian ini untuk merekomendasikan calon sekolah terbaik bagi siswa dengan memanfaatkan data sekolah menengah atas di Kota Tapanuli Utara. Metode Cart diterapkan untuk mencari probabilitas terbesar dari semua instance pada atribut target seperti Nama Sekolah, Uang Sekolah, tingkat kelulusan dan nilai rata-rata minimal UN. Hasil dari penelitian ini yaitu sebuah sistem rekomendasi yang dapat membantu menyajikan informasi sekolah-sekolah menengah atas yang sesuai

dengan kriteria. Sehingga dapat menjadi pertimbangan bagi siswa untuk memilih calon sekolah menengah kejuruan yang terbaik sesuai dengan kriteria yang diajukan. Salah satu teknik pengolahan data yang dapat digunakan untuk masalah tersebut adalah algoritma *cart* melalui aplikasi WEKA versi 3.8.

Dalam menyelesaikan masalah tersebut diperlukan data dapat dilakukan analisa pemilihan SMK terhadap data khusus pada pengembangan sistem rekomendasi pemilihan SMK sebelumnya. Dengan mengetahui SMK apa saja yang diminati oleh siswa baru. Dengan tujuan untuk menentukan calon cabang terhadap seluruh prediktor dalam pemilihan SMK

Berikut ini adalah data pengembangan sistem rekomendasi pemilihan SMK yang akan dijadikan sampel untuk analisa dan pengujian

**Table 1.** Data Training

N O	Nama Sekolah	Uang Sekolah	Tingkat Kelulusan	Nilai Rata –Rata UN	Jenis Sekolah
1	SMK Negeri 1 Siatas Barita	High ( $\leq 75.000$ )	High	25	Good
2	SMK Negeri 1 Siborong-borong	Medium ( $\leq 50.000$ )	High	75	Good
3	SMK Negeri 1 Muara	Low ( $\leq 25.000$ )	Low	25	Bad
4	SMKNegeri 1 Pagaran	Low ( $\leq 25.000$ )	Low	50	Bad
5	SMKNegeri 1 Losida Siatas Barita	High ( $\leq 75.000$ )	Medium	25	Bad
6	SMK Hki Tarutung	Medium ( $\leq 50.000$ )	Medium	50	Good
7	SMK Nusantara Siborong-borong	Medium ( $\leq 50.000$ )	Medium	75	Good
8	Smk Karya Tarutung	Low ( $\leq 25.000$ )	Medium	25	Good
9	SMK Swasta Dharma Bakti Siborong- Borong	Medium ( $\leq 50.000$ )	Medium	45	Good
10	SMK Hkbp Sarulla	Medium ( $\leq 50.000$ )	Medium	50	Good
11	SMK Hkbp 1 Saipaholon	Medium ( $\leq 50.000$ )	Low	60	Bad
12	SMK Hkbp 2 Sipaholon	High ( $\leq 75.000$ )	Low	45	Bad
13	SMK Swasta Parsaoran Parmonangan	High( $\leq 75.000$ )	Low	50	Bad
14	SMK Swasta Gunung Karmel	Low ( $\leq 25.000$ )	Low	70	Good
15	SMK N 1 Pahae Julu	Low ( $\leq 25.000$ )	High	60	Good
16	SMK Baginda Mulia	Medium ( $\leq 50.000$ )	High	30	Good

Keterangan :

1. Nilai Rata-Rata UN:
  - a. Low (0-30),
  - b. Medium (31-50),
  - c. High (51-75)
2. Uang Sekolah (Rp):
  - a. Low ( $\leq 25.000$ ),
  - b. Medium ( $\leq 50.000$ ),
  - c. High ( $\leq 75.000$ )
3. Tingkat Kelulusan
  - a. Low (0-29%)
  - b. Medium (30-59%)
  - c. High (60-100%)

Berikut adalah tahapan-tahapan penyelesaian pada Algoritma CART untuk membentuk sebuah pohon keputusan dari data *sekolah* yang akan dijadikan acuan untuk pengembangan sistem rekomendasi pemilihan SMK pada Dinas pendidikan dan kebudayaan Tapanuli Utara yaitu :

Langkah pertama algoritma CART adalah menyusun calon cabang (*candidate split*). Penyusunan ini dilakukan terhadap seluruh variabel prediktor secara lengkap (*exhaustive*). Pada kasus ini, yang menjadi variabel prediktor seluruhnya adalah Nama Sekolah ,Jumlah Siswa dan Akreditasi. Mari kita susun calon cabang yang lengkap untuk masing-masing variabel prediktor ini.Perhatikanlah bahwa pembuatan calon cabang akan selalu patuh pada ciri khas algoritma CART, yaitu adanya noktah keputusan yang selalu bercabang dua atau bercabang biner.

Calon cabang untuk variabel prediktor Uang Sekolah adalah sebagai berikut:

1. Uang Sekolah = High, Uang Sekolah = Low, Medium
2. Uang Sekolah = Low, Uang Sekolah = Medium, High
3. Uang Sekolah = Medium, Uang Sekolah = Low, High

Calon cabang untuk variabel prediktor Tingkat Kelulusan adalah sebagai berikut:

1. Tingkat Kelulusan = High, Tingkat Kelulusan = Low, Medium
2. Tingkat Kelulusan = Low, Tingkat Kelulusan = Medium, High
3. Tingkat Kelulusan = Medium, Tingkat Kelulusan = Low, High

Calon cabang untuk variable predictor Nilai Rata- Rata UN yang dapat diusulkan adalah sebagai berikut :

1. Nilai Rata-Rata UN = High, Tingkat Kelulusan = Low, Medium
2. Nilai Rata-Rata UN = Low, Tingkat Kelulusan = Medium, High
3. Nilai Rata-Rata UN = Medium, Tingkat Kelulusan = Low, High

Pada variabel prediktor Jenis Sekolah, calon cabang yang dapat diusulkan adalah sebagai berikut:

1. Jenis Sekolah = Good, Jenis Sekolah = Bad
2. Jenis Sekolah = Bad, Jenis Sekolah = Good

**Tabel 2.** Keterangan Data Jenis Sekolah

No.	Nama Kategori	Nilai
1.	Good	10
2.	Bad	6

Mengingat ciri khas algoritma CART yang setiap noktah keputusannya bercabang biner, maka calon cabang akan diberi nama calon cabang kiri dan calon cabang kanan.

**Tabel 3.** Daftar calon cabang mutakhir (iterasi-1)

Calon Cabang	Calon Cabang Kiri (tL)	Calon Cabang Kanan (tR)
1	Uang Sekolah = High	Uang Sekolah = Low, Medium
2	Uang Sekolah = Medium	Uang Sekolah = Low, High
3	Uang Sekolah = Low	Uang Sekolah = Medium, High
4	Tingkat Kelulusan = High	Tingkat Kelulusan = Low, Medium
5	Tingkat Kelulusan = Medium	Tingkat Kelulusan = Low, High
6	Tingkat Kelulusan = Low	Tingkat Kelulusan = Medium, High
7	Nilai Rata-Rata UN = High	Nilai Rata-Rata UN = Low, Medium
8	Nilai Rata-Rata UN = Medium	Nilai Rata-Rata UN = Low, High
9	Nilai Rata-Rata UN = Low	Nilai Rata-Rata UN = Medium, High
10	Jenis Sekolah = Good	Jenis Sekolah = Bad
11	Jenis Sekolah = Bad	Jenis Sekolah = Good

Langkah kedua algoritma ini akan menilai kinerja keseluruhan calon cabang yang ada pada daftar calon cabang mutakhir. Untuk saat ini, daftar calon cabang mutakhir. Lalu, bagaimana mengukur kinerja masing-masing calon cabang yang terdaftar pada daftar tersebut?

Hasil perhitungan nilai kesesuaian bagi tiap calon cabang yang masih terdapat dalam daftar calon cabang mutakhir berikut ini :

**Tabel 4.** Perhitungan nilai kesesuaian untuk calon cabang 1, 2, , 4, 5, 6, 7, 8, 9, 10,11 (iterasi-1)

Calon Cabang	PL	PR	Jenis Sekolah	P(j tL)	P(j tR)	2.PL.PR	Q(s t)	φ(s t)																																																																																																																		
1	1/16= 0,062	11/1= 0,687	Good	1/1= 1	7/7=1	0,851	1	0,343																																																																																																																		
			Bad	0	3/3=1				2	7/16= 0,437	9/16= 0,562	Good	7/16=0,437	4/4=1	0,491	-0,437	-0,241	Bad	0	5/5=1	3	5/16= 0,312	10/1= 0,625	Good	3/16=0,187	7/7=1	0,39	1,688	0,658	Bad	2/16=0,125	3/3=1	4	4/16=0,25	12/1=0,75	Good	1/6=0,1660	5/15=0,333	0,375	0,584	0,219	Bad	0	4/6=0,25	5	6/16=0,375	10/1=0,625	Good	4/4=1	4/4=1	0,468	-1	0	Bad	0	5/5=1	6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468	Bad	0	1/1=1	7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1
2	7/16= 0,437	9/16= 0,562	Good	7/16=0,437	4/4=1	0,491	-0,437	-0,241																																																																																																																		
			Bad	0	5/5=1				3	5/16= 0,312	10/1= 0,625	Good	3/16=0,187	7/7=1	0,39	1,688	0,658	Bad	2/16=0,125	3/3=1	4	4/16=0,25	12/1=0,75	Good	1/6=0,1660	5/15=0,333	0,375	0,584	0,219	Bad	0	4/6=0,25	5	6/16=0,375	10/1=0,625	Good	4/4=1	4/4=1	0,468	-1	0	Bad	0	5/5=1	6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468	Bad	0	1/1=1	7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1						
3	5/16= 0,312	10/1= 0,625	Good	3/16=0,187	7/7=1	0,39	1,688	0,658																																																																																																																		
			Bad	2/16=0,125	3/3=1				4	4/16=0,25	12/1=0,75	Good	1/6=0,1660	5/15=0,333	0,375	0,584	0,219	Bad	0	4/6=0,25	5	6/16=0,375	10/1=0,625	Good	4/4=1	4/4=1	0,468	-1	0	Bad	0	5/5=1	6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468	Bad	0	1/1=1	7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																		
4	4/16=0,25	12/1=0,75	Good	1/6=0,1660	5/15=0,333	0,375	0,584	0,219																																																																																																																		
			Bad	0	4/6=0,25				5	6/16=0,375	10/1=0,625	Good	4/4=1	4/4=1	0,468	-1	0	Bad	0	5/5=1	6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468	Bad	0	1/1=1	7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																														
5	6/16=0,375	10/1=0,625	Good	4/4=1	4/4=1	0,468	-1	0																																																																																																																		
			Bad	0	5/5=1				6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468	Bad	0	1/1=1	7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																										
6	6/16=0,375	10/1=0,625	Good	9/9=1	9/9=1	0,468	-1	-0,468																																																																																																																		
			Bad	0	1/1=1				7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428	Bad	0	3/3=1	8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																																						
7	5/16=0,312	11/1=0,687	Good	1/1=1	7/7=1	0,428	-1	-0,428																																																																																																																		
			Bad	0	3/3=1				8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468	Bad	0	4/4=1	9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																																																		
8	6/16=0,375	10/1=0,625	Good	3/3=1	4/4=1	0,468	-1	-0,468																																																																																																																		
			Bad	0	4/4=1				9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428	Bad	0	1/1=1	10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																																																														
9	5/16=0,312	11/1=0,687	Good	1/1=1	6/6=1	0,428	-1	-0,428																																																																																																																		
			Bad	0	1/1=1				10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,	Bad	6/6=1	6/6=1	11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																																																																										
10	10/16=0,625	6/16=0,375	Good	10/10=1	10/10=1	0,468	0	0,																																																																																																																		
			Bad	6/6=1	6/6=1				11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0	Bad	10/10=1	10/10=1																																																																																																						
11	6/16=0,375	10/1=0,625	Good	6/6=1	6/6=1	0,468	0	0																																																																																																																		
			Bad	10/10=1	10/10=1																																																																																																																					

Untuk mendapatkan gambaran mengenai cara mendapatkan Tabel 4.2, berikut adalah penjelasan mengenai perhitungan nilai kesesuaian, misalnya bagi calon cabang nomor 1, yaitu calon cabang yang terdiri atas calon cabang kiri, yaitu Uang Sekolah = High, dan calon cabang kanan, yaitu Uang Sekolah = (Low, Medium). Ada 1 buah catatan yang memenuhi Uang Sekolah = High yaitu catatan 2018. Adapun jumlah catatan pada data latihan yang dipertimbangkan saat ini 16 buah, yaitu Smk Negeri 1 Siatas Barita, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Pagaran, Smk Negeri 1 Losida Siatas Barita, Smk Hki Tarutung, Smk Nusantara Siborong-borong, Smk Karya Tarutung,

Smk Swasta Dharma Bakti Siborong-Borong, Smk Hkbp Sarulla, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Saipaholon, Smk Swasta Parsaoran Parmonangan, Smk Swasta Gunung Karmel, Smk N 1 Pahae Julu, Smk Baginda Mulia

$$PL = \frac{\text{jumlah catatan dengan Uang Sekolah} = \text{High pada calon cabang kiri ke} - 1}{\text{jumlah catatan pada data latihan pada noktah keputusan yang tengah dibahas}} = \frac{1}{16} = 0,062$$

Ada 15 buah catatan yang memenuhi Uang Sekolah = Uang Sekolah = (Low, Medium), yaitu catatan Smk N 1 Muara, Smk N 1 Pagaran, Smk Swasta Gunung Karmel, Smk N 1 Pahae Julu, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Losida Siatas Barita, Smk Nusantara Siborong-borong, Smk Swasta Dharma Bakti Siborong-Borong, Smk Hkbp Sarulla, Smk Hkbp 1 Sipaholon, Smk Baginda Mulia sehingga dari persamaan didapatkan:

$$PR = \frac{\text{jumlah catatan dengan Uang Sekolah} = \text{Low, Medium pada calon cabang kanan ke} - 1}{\text{jumlah catatan pada data latihan pada noktah keputusan yang tengah dibahas}} = \frac{15}{16} = 0,937$$

Dari 11 buah catatan yang memenuhi syarat calon cabang kiri, yaitu Uang Sekolah = Medium adalah Catatan Jenis Sekolah Good

1. Enam catatan diantaranya, yaitu Catatan Jenis Sekolah Good Memenuhi Uang Sekolah = Medium sehingga dari persamaan didapatkan:

$$P(j|tL) = \frac{\text{jumlah catatan dengan Uang Sekolah} = \text{Medium pada calon cabang kiri ke} - 2}{\text{jumlah catatan pada noktah keputusan } t \text{ yang tengah dibahas}} = \frac{2}{16} = 0,125$$

2. Catatan memenuhi Uang Sekolah = Low sehingga dari persamaan didapatkan:

$$P(j|tL) = \frac{\text{jumlah catatan dengan Tingkat Kelulusan} = \text{rendah pada calon cabang kiri ke} - 3}{\text{jumlah catatan pada noktah keputusan } t \text{ yang tengah dibahas}} = \frac{3}{5} = 0,6$$

Sepuluh buah catatan yang memenuhi syarat calon cabang kanan, yaitu Tingkat Kelulusan = Tingkat Kelulusan = Low, Medium adalah catatan Smk Negeri 1 Muara, Smk N 1 Pagaran, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Sipaholon, Smk Swasta Parsaoran Parmonangan, Smk Swasta Gunung Karmel, Smk Negeri 1 Losida Siatas Barita, SMK HKI Tarutung, SMK Nusantara Siborong-borong, SMK Karya Tarutung.

1. Sembilan catatan diantaranya, yaitu catatan Smk Negeri 1 Muara, Smk N 1 Pagaran, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Sipaholon, Smk Swasta Parsaoran Parmonangan, Smk Swasta Gunung Karmel, Smk Negeri 1 Siatas Barita, Smk Negeri 1 Siborong-borong, Smk Baginda Mulia Tingkat Kelulusan = Low, High sehingga dari persamaan didapatkan:

$$P(j|tR) = \frac{\text{jumlah catatan dengan Tingkat Kelulusan} = \text{Low, High pada calon cabang kanan ke} - 5}{\text{jumlah catatan pada noktah keputusan } t \text{ yang tengah dibahas}} = \frac{5}{9} = 0,312$$

2. Sepuluh catatan diantaranya, yaitu Smk Negeri 1 Losida Siatas Barita, Smk Hki Tarutung, Smk Nusantara Siborong-borong, Smk Karya Tarutung, Smk Swasta Dharma Bakti Siborong-borong, Smk Hkbp Sarulla, Smk Negeri 1 siatas Barita, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Pahae Julu, Smk Baginda Mulia = Medium, High sehingga dari persamaan didapatkan:

$$P(j|tR) = \frac{\text{jumlah catatan Tingkat Kelulusan} = \text{Medium, High pada calon cabang kanan ke} - 6}{\text{jumlah catatan pada noktah keputusan } t \text{ yang tengah dibahas}} = \frac{6}{10} = 0,6$$

Selanjutnya didapatkan:

1. Nilai dari besaran  
2. PL.PR = 2 (0,312) (0,625) = 0,39
2. Nilai dari besaran adalah

$$Q(s|t) = \sum_{j=1}^{\text{jlh kategori}} |P(j|tL) - P(j|tR)| = |1 - 0,187| + |1 - 0,125| = 1,688$$

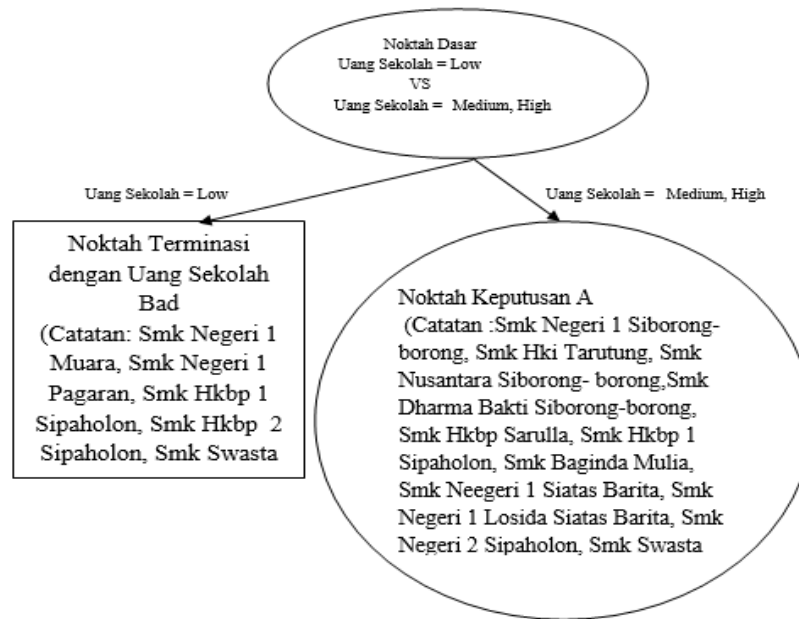
Dan akhirnya didapatkan pula :

$$\Phi(s|t) = 2 PL PR \Phi(s|t) = 2(0,312)(0,625)(1,688) = 0,658$$

Perhitungan nilai kesesuaian terhadap calon cabang nomor 1 dapat diambil analoginya sehingga pembaca dapat memeriksa perhitungan sejenis terhadap calon cabang yang lain. Selengkapnya, hasil perhitungan itu tersaji pada Tabel 4.4.

Langkah ketiga algoritma ini adalah menentukan calon cabang manakah yang akan benar-benar dijadikan cabang. Hal ini dilakukan dengan memilih calon cabang yang memiliki nilai kesesuaian  $\Phi(s|t)$  terbesar. Setelah itu, gambarkanlah percabangan sesuai hasil menjalankan algoritma. Jika tidak ada lagi noktah keputusan, pelaksanaan algoritma CART akan dihentikan. Namun, jika masih terdapat noktah keputusan, pelaksanaan algoritma dilanjutkan dengan kembali ke langkah kedua, dengan terlebih dahulu membuang calon cabang yang telah berhasil menjadi cabang.

Dari Tabel 4. tampak bahwa calon cabang nomor 3 adalah calon cabang dengan nilai besaran kesesuaian terbesar dari pada calon cabang yang lain, maka calon cabang inilah yang akan kita pilih sebagai cabang pada tahap ini sehingga kita memperoleh Gambar 1 dari iterasi-1 ini.



Gambar 1. Pohon keputusan bagi masalah Pemilihan SMK (iterasi-1)

1. Pada noktah dasar, kita masih berhadapan dengan seluruh catatan, yaitu catatan Smk Negeri 1 Siatas Barita, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Siborong-borong, Smk Negeri 1 Pagaran, Smk Negeri 1 Losida Siatas Barita, Smk Hki Tarutung, Smk Nusantara Siborong-borong, Smk Karya Tarutung, Smk Swasta Dharma Bakti Siborong-Borong, Smk Hkbp Sarulla, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Saipaholon, Smk Swasta Parsaoran Parmonangan, Smk Swasta Gunung Karmel, Smk N 1 Pahae Julu, Smk Baginda Mulia
2. Calon cabang nomor 3 kini telah benar-benar menjadi cabang. Adapun cabang kirinya, yaitu cabang yang memenuhi syarat Uang Sekolah = Low dipenuhi oleh catatan Smk Negeri 1 Muara, Smk Negeri 1 Pagaran, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Sipaholon, Smk Swasta Parsaoran Parmonangan dan karena nilai variabel prediktor bagi kedua catatan ini seluruhnya adalah Jenis Sekolah = Bad, maka noktah terminasi akan dihasilkan.
3. Cabang kanannya, yaitu cabang yang memenuhi syarat Uang Sekolah = Low, Medium dipenuhi oleh catatan Smk Negeri 1 Muara, Smk Negeri 1 Pagaran, Smk Karya Tarutung, Smk Swasta Gunung Karmel, Smk Negeri 1 Pahae Julu, Smk Negeri 1 Siborong-borong, Smk Hki Tarutung, Smk Nusantara Siborong-borong, Smk Swasta Dharma Bakti Siborong-borong, Smk Hkbp Sarulla, Smk Hkbp 1 Sipaholon, Smk Hkbp 2 Sipaholon, Smk Baginda Mulia, dan karena nilai variabel prediktor bagi catatan-catatan ini ada yang memiliki Jenis Sekolah = Bad, maka dihasilkanlah noktah keputusan. Karena merupakan noktah keputusan pertama, kita dapat menyebutnya noktah keputusan A. Noktah ini akan bercabang lebih lanjut karena bukan termasuk noktah terminasi.

#### 4. KESIMPULAN

Berdasarkan implementasi dan pengujian Pemnfaatan Metode CART Dalam Pengembangan Sistem Rekomendasi Pemilihan SMK maka dapat ditarik beberapa kesimpulan Metode Cart pohon klafikasi bisa diterapkan untuk memilah seluruh amatan menjadi dua gugus amatan dan memilah kembali gugus amatan tersebut menjadi dua gugus amatan berikutnya, sehingga diperoleh jumlah amatan minimum pada gugus amatan berikutnya. Sehingga dapat membentuk pohon keputusan dengan data – data yang diperoleh Dinas Pendidikan Tapanuli Utara Berdasarkan data pada Dinas Pendidikan Tapanuli Utara yang di jadikan data training, metode Cart berhasil mengklasifikasikan data yang diuji dengan menggunakan aplikasi Weka versi 3.8 dengan persentase keakuratan sebesar 100%..

#### REFERENCES

- [1] S. B. Alan, B. Efori and G. Perमानan , "PEMANFAATAN ALGORITMA CLASSIFICATION AND REGRESSION TRESS (CART) UNTUK MEMPREDIKSI OMSET SPANDUK PADA CV. MOEHA ADVERTISING," KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer), vol. 1, no. 1, pp. 345-354, 2017.
- [2] Budi Santosa, Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis, 1st ed. Yogyakarta: Graha Ilmu, 2007.
- [3] Kusriani and Emha Taufiq Luthfi, Algoritma Data Mining. Yogyakarta: Andi, 2009.
- [4] Beta Noranita and Nurdin Bahtiar, "Implementasi Data Mining Untuk Menentukan Pola Hubungan Tingkat Kelulusan Mahasiswa dengan Induk Mahasiswa," pp. 156-163, Oktober 2010.

- [5] Dana Sulisty Kusumo, "Data Mining dengan Algoritma Apriori pada RDBMS Oracle," Penelitian dan Pengembangan Telekomunikasi, vol. 8, 2003.
- [6] Finn Lee and Juan Santana, Data Mining: Meramalkan Bisnis Perusahaan, 1st ed. Jakarta: Elex Media Komputindo, 2010.
- [7] Berry Michael and Gordon Linoff, Data Mining: Techniques For Marketing, Sales, Customer Relationship Management, 2nd ed. Canada: Wiley Publishing Inc, 2004.