

Penerapan Data Mining dalam Menganalisa Pola Kelayakan Siswa Pada Kelas Unggulan Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) pada SMP N. 2 Rantau Selatan

Masyuni Hutasuhut , Dina Octavina, Jufri Halim
Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received May 31th, 2019

Revised June 12th, 2019

Accepted Augs 02th, 2019

Keyword:

Data Mining

Iterative Dichtomizer 3

ID3

Kelas Unggulan

ABSTRACT

Kelas unggulan Pada SMP Negeri 2 Rantau Selatan merupakan program yang dilakukan oleh pihak sekolah untuk meningkatkan kualitas siswanya. Kelas unggulan Pada SMP Negeri 2 Rantau selama ini dibentuk dengan tes ujian masuk yang dilakukan pihak Sekolah SMP Negeri 2 Rantau Selatan kepada siswa baru. Kelas unggulan yang telah dibentuk dirasa masih kurang efektif karena ada beberapa siswa yang memang tidak layak ditempatkan dan ada juga siswa yang memang layak ditempatkan pada kelas unggulan. Untuk itu perlu dilakukan analisa menganalisa pola pembentukan kelas unggulan yang telah dibentuk oleh pihak sekolah yang nantinya dapat digunakan sebagai acuan yang lebih baik lagi untuk masa yang akan datang. Untuk mendapatkan pola tersebut digunakan data mining dengan menggunakan metode Iterative Dichtomizer 3 (ID3) dalam menganalisa pola pada kelas unggulan diharapkan bisa memberikan hasil yang lebih baik bagi pihak sekolah dalam menempatkan siswa berprestasi di kelas unggulan nantinya. Hasil yang diperoleh dengan menggunakan metode ini yaitu berupa pola dalam bentuk pohon keputusan yang diterjemahkan menjadi aturan-aturan (rule) yang berjumlah 28 rule. Berdasarkan rule tersebut dapat dijadikan acuan dalam pembentukan kelas unggulan di masa mendatang.

Copyright © 2019 STMIK Triguna Dharma.
All rights reserved.

First Author

Nama :Masyuni Hutasuhut
Kantor :STMIK Triguna Dharma
Program Studi :Sistem Informasi
E-Mail :yunihutasuhut@gmail.com

1. PENDAHULUAN

Adanya kebebasan setiap sekolah mengelola potensi yang ada di sekolah, maka dimungkinkan adanya pelayanan yang optimal terhadap siswa yang memiliki kecerdasan tinggi dan siswa yang memiliki kecerdasan rendah sesuai dengan potensi yang ia miliki. Oleh karena itu, pihak sekolah merespon potensi siswa yang memiliki kecerdasan tinggi dan bakat khusus dengan menyelenggarakan kelas unggulan. Penyelenggaraan kelas unggulan merujuk pada amanat Undang-Undang Republik Indonesia Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional Bab IV bagian kesatu Pasal 5 ayat 4.

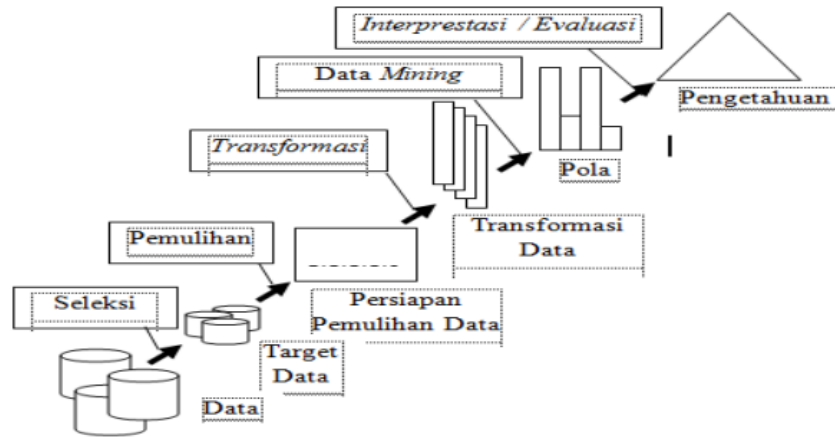
Dalam pembentukan kelas unggulan pada SMP Negeri 2 Rantau Selatan selama ini yang dilakukan oleh pihak sekolah dinilai masih kurang objektif dan terdapat beberapa masalah diantaranya ada beberapa siswa yang tidak layak dan ditempatkan pada kelas unggulan dan ada pula siswa yang berpotensi dan layak tidak ditempatkan pada kelas unggulan dikarenakan seleksi yang dilakukan oleh pihak sekolah.

Adapun data yang dikelola dalam penelitian ini akan dikelola dengan menggunakan ID3. Dengan menggunakan metode ID3 dapat diterapkan pada permasalahan yang ada yaitu menganalisa data siswa baru berprestasi agar lebih selektif dalam membentuk kelas unggulan.

2. METODE PENELITIAN

2.1 Knowledge Discovery in Databases (KDD)

Data Mining sering disebut sebagai Knowledge Discovery in Database (KDD) yang bertugas untuk mengekstrak pola atau model dari data dengan menggunakan suatu algoritma yang spesifik [1]. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat. Data Mining merupakan salah satu langkah dari serangkaian proses interative KDD. Terhadap proses KDD dapat dilihat pada gambar [2].



Gambar 1. Tahapan dalam Proses Knowledge Discovery in Database

2.2 Data Mining

Data Mining merupakan proses ekstraksi data menjadi informasi yang sebelumnya belum tersampaikan, dengan teknik yang tepat proses Data Mining akan memberikan hasil yang optimal [3].

Data Mining merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang statistik, pembelajaran mesin dan sistem manajemen basis data [4].

Dan Secara garis besar metode pelatihan yang digunakan dalam teknik-teknik *Data Mining* dibedakan ke dalam dua pendekatan, yaitu

- 2.1 *Unsupervised Learning*, metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*). Guru di sini adalah label dari data.
- 2.2 *Supervised Learning*, yaitu metode belajar dengan adanya latihan dan pelatih. Dalam pendekatan ini, untuk menemukan fungsi keputusan, fungsi pemisah atau fungsi regresi, digunakan beberapa contoh data yang mempunyai output atau label selama proses [5].

2.3 Algoritma Iterative Dichotomiser 3 (ID3)

Algoritma ID3 atau *Iterative Dichotomiser 3* (ID3) merupakan sebuah metode yang digunakan untuk membangkitkan pohon keputusan. Algoritma pada metode ini menggunakan konsep dari *entropi* informasi. Secara ringkas, cara kerja Algoritma ID3 dapat digambarkan sebagai berikut [6]. Pemilihan atribut dengan menggunakan *Information Gain*

1. Pilih *atribut* dimana nilai *information gain*-nya terbesar,
2. Buat simpul yang berisi atribut tersebut,
3. Proses perhitungan *information gain* akan terus dilaksanakan sampai semua data masuk dalam kelas yang sama. Atribut yang telah dipilih tidak diikuti lagi dalam perhitungan nilai *information gain*.

Pada algoritma ID3 pengurangan *entropy* disebut dengan *informasi gain*. Pembagian sample S terhadap atribut A dapat dihitung *information gain* dengan rumus :

$$Gain(S,A) = Entropy(S) - \sum_{vs \text{ nilai } A} \left| \frac{S_v}{S} \right| entropy(S_v)$$

Entropi digunakan untuk mengukur jumlah informasi yang ada pada atribut dengan rumus. :

$$Entropy(S) = -(P_+ \times \log_2 P_+ + P_- \times \log_2 P_-)$$

3. ANALISA DAN HASIL

3.1. Analisis Data

Analisis kebutuhan data yang dilakukan untuk menerapkan algoritma ID3 pada sistem yang akan dirancang untuk menentukan pola yang ada pada SMP Negeri 2 Rantau Selatan ini berdasarkan hasil tes pada

calon siswa dalam bentuk nilai akan dirubah terlebih dahulu menjadi batasan atau *range* agar memudahkan proses perhitungan pada algoritma nantinya. Adapun jumlah siswa yang akan dijadikan contoh untuk pembentukan pola ini berjumlah 88 orang siswa. Berikut ini merupakan penerapan algoritma ID3 dalam menentukan pola pembentukan kelas unggulan.

Tabel 1. Hasil Pemilihan Siswa Kelas Unggulan

NO	Nama Siswa	Nil. Seleksi	Nil. Psikotes	Nil. Wawancara	Kelas Unggulan
1	SAHAT MARTUA SINAGA	80	65	75	tidak
2	BRYAN HARRY PRAYOGI OMPUSUNGGU	78	60	90	tidak
3	RAHMAD ZUANDI	88	58	90	Ya
4	TONI SANJAYA	79	79	90	Ya
5	SABRINA EMILIA MALIK RAMBE	87	60	90	Ya
6	DZAITUN JAHARA	89	60	90	Ya
7	ARMANDA LAILA	67	87	79	tidak
8	SITI JAHARA	60	90	87	Ya
9	ANISA SUSANTI	58	90	89	Ya
10	RATNA EKA SARI	79	79	67	tidak
11	ISHAK IMMANUEL SIMANJUNTAK	60	60	87	tidak
12	ANDIKA PERMANA SANDI	60	58	90	tidak
13	NIZAR IBRAHAMKA SIAHAAN	63	79	90	tidak
14	SUSANA TRI IRWANA	90	60	90	Ya
15	ALFI SAHRIN RAMBE	79	87	90	Ya
16	RAMADHANI RITONGA	87	90	65	Ya
17	JONATAN NABABAN	89	90	60	Ya
18	BAKTIAR SAMSURI SIREGAR	67	90	58	tidak
19	RISNA DAYANTI GULO	87	90	79	Ya
20	NUR YURIDA KURNIA	79	79	60	tidak
...					
88	RAJA MULIA DALIMUNTHE	87	87	87	Ya

Dari tabel diatas maka tiap-tiap kumpulan nilai akan dijadikan atribut yang akan diberikan *range* dengan goal = hasil agar memudahkan proses perhitungan dengan menggunakan ID3. Atribut yang akan digunakan adalah:

1. Atribut nilai seleksi dengan *range*:
 - a. Sangat bagus (SB) : 80 - 100
 - b. Bagus (B) : 66 - 79
 - c. Kurang bagus (KB) : 0 - 65
2. Atribut nilai psikotes dengan *range*:
 - a. Sangat Tinggi (ST) : 80 - 100
 - b. Tinggi (T) : 66 - 79
 - c. Kurang tinggi (KT) : 0 - 65
3. Atribut nilai wawancara dengan *range*:

- a. Sangat memuaskan (SM) : 80 - 100
- b. Memuaskan (M) : 66 - 79
- c. Kurang memuaskan (KM) : 0 – 65

Konversi nilai siswa agar memudahkan proses perhitungan, tabel hasil pemilihan siswa kelas unggulan menjadi bentuk range untuk tiap-tiap atribut dapat dilihat pada tabel dibawah ini.

Tabel 2 Konversi Nilai Pemilihan Siswa Kelas Unggulan

NO	Nama Siswa	Nil. Seleksi	Nil. Psikotes	Nil. Wawancara	Kelas Unggulan
1	SAHAT MARTUA SINAGA	sangat bagus	kurang tinggi	memuaskan	tidak
2	BRYAN HARRY PRAYOGI OMPUSUNGGU	Bagus	kurang tinggi	sangat memuaskan	tidak
3	RAHMAD ZUANDI	sangat bagus	kurang tinggi	sangat memuaskan	Ya
4	TONI SANJAYA	Bagus	tinggi	sangat memuaskan	Ya
5	SABRINA EMILIA MALIK RAMBE	sangat bagus	kurang tinggi	sangat memuaskan	Ya
6	DZAITUN JAHARA	sangat bagus	kurang tinggi	sangat memuaskan	Ya
7	ARMANDA LAILA	Bagus	sangat tinggi	memuaskan	tidak
8	SITI JAHARA	kurang bagus	sangat tinggi	sangat memuaskan	Ya
9	ANISA SUSANTI	kurang bagus	sangat tinggi	sangat memuaskan	Ya
10	RATNA EKA SARI	Bagus	tinggi	memuaskan	tidak
11	ISHAK IMMANUEL SIMANJUNTAK	kurang bagus	kurang tinggi	sangat memuaskan	tidak
12	ANDIKA PERMANA SANDI	kurang bagus	kurang tinggi	sangat memuaskan	tidak
13	NIZAR IBRAHAMKA SIAHAAN	kurang bagus	tinggi	sangat memuaskan	tidak
14	SUSANA TRI IRWANA	sangat bagus	kurang tinggi	sangat memuaskan	Ya
15	ALFI SAHRIN RAMBE	Bagus	sangat tinggi	sangat memuaskan	Ya
16	RAMADHANI RITONGA	sangat bagus	sangat tinggi	kurang memuaskan	Ya
17	JONATAN NABABAN	sangat bagus	sangat tinggi	kurang memuaskan	Ya
18	BAKTIAR SAMSURI SIREGAR	Bagus	sangat tinggi	kurang memuaskan	tidak
19	RISNA DAYANTI GULO	sangat bagus	sangat tinggi	memuaskan	Ya
20	NUR YURIDA KURNIA	Bagus	tinggi	kurang memuaskan	tidak
...					
88	RAJA MULIA DALIMUNTHE	sangat bagus	sangat tinggi	sangat memuaskan	Ya

Setelah data dikonversi menjadi *range* maka selanjutnya dilakukan perhitungan dengan menggunakan algoritma ID3 sebagai berikut:

Langkah awal dimulai dari node (atribut) akar,

Hitung semua entropy untuk node akar (semua data) terhadap komposisi kelas dimana atribut kelas unggulan = “ya” bernilai (+) dan atribut kelas unggulan = “tidak” bernilai (-).

Kelas unggulan [+53, -35] = 88 siswa

$$E(\text{semua}) = - \left[\left(\frac{53}{88} \right) \times \log_2 \left(\frac{53}{88} \right) + \left(\frac{35}{88} \right) \times \log_2 \left(\frac{35}{88} \right) \right]$$

$$= 0.9696055$$

Selanjutnya dihitung entropy untuk setiap fitur terhadap kelas. Untuk entropy nilai seleksi, nilai psikotes dan nilai wawancara dengan rumus yang sama. Setelah nilai entropy diperoleh maka selanjutnya mencari nilai gain.

Gain Nilai Wawancara

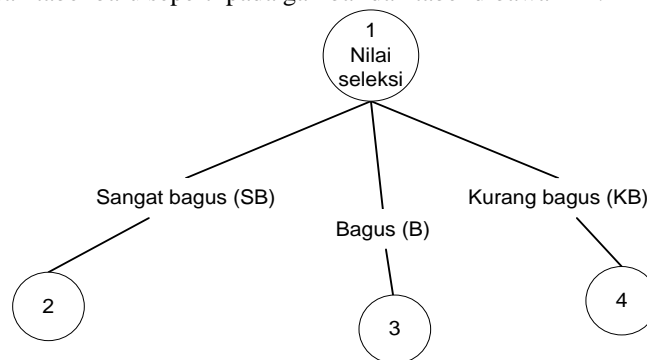
$$= 0.969605599 - \left(0.818955502 \times \frac{51}{88} + 0.998363673 \times \frac{21}{88} + 0.818955502 \times \frac{16}{88} \right)$$

$$= \mathbf{0.093822}$$

Tabel 3 Hasil perhitungan Entropy dan Gain untuk Node akar

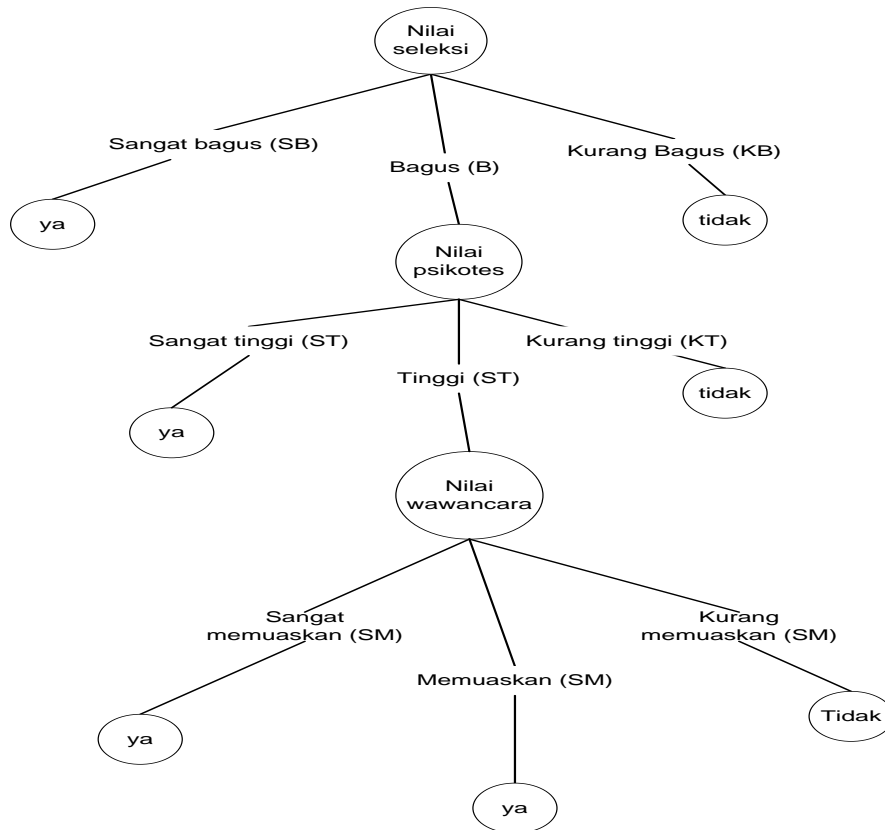
Node			Jumlah	Ya	Tidak	Entropy	Gain
1	Total		88	53	35	0.9696055	
Nilai seleksi	SB		46	40	6	0.615254	0.221181
	B		28	11	17	0.966618633	
	KB		14	3	11	0.749595257	
Nilai psikotes	ST		39	29	10	0.821280942	0.049167
	T		26	13	13	1	
	KT		23	3	11	0.998636	
Nilai wawancara	SM		51	38	13	0.818955502	0.093822
	M		21	10	11	0.998363673	
	KM		16	5	11	0.896038233	

Hasil yang didapat dari tabel 3.3 menunjukkan bahwa gain tertinggi ada di fitur “nilai seleksi” sehingga “nilai seleksi” dijadikan node akar. Untuk cabangnya, digunakan 3 nilai di dalam node akar. Selanjutnya data akan terpecah menjadi 3 kelompok yaitu “sangat bagus”, “bagus” dan “kurang bagus”. Sehingga menghasilkan pohon dan tabel baru seperti pada gambar dan tabel dibawah ini.



Gambar 2. Hasil Pohon Node Ke 1

Perhitungan lanjutan dilakukan untuk mencari node selanjutnya dengan langkah yang sama hingga diperoleh hasil gambar pembentukan pola induksi *decision tree* pada gambar berikut.



Gambar 3. Ranting Decision Tree

3.2. Rule Decision Tree

Pohon yang dihasilkan dari ranting decision tree dapat dinyatakan dalam bentuk IF THEN sebagai berikut. Dari ranting decision tree didapat rule pola baru yaitu :

1. Rule 1 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya
2. Rule 2 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=Ya
3. Rule 3 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=Ya
4. Rule 4 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya
5. Rule 5 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=Ya
6. Rule 6 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya
7. Rule 7 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
8. Rule 8 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
9. Rule 9 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak
10. Rule 10 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
11. Rule 11 IF Nilai Seleksi=sangat bagus and Nilai Psikotes=tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
12. Rule 12 IF Nilai Seleksi=bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=Ya
13. Rule 13 IF Nilai Seleksi=bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya

14. Rule 14 IF Nilai Seleksi=bagus and Nilai Psikotes=tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya
15. Rule 15 IF Nilai Seleksi=bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
16. Rule 16 IF Nilai Seleksi=bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak
17. Rule 17 IF Nilai Seleksi=bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
18. Rule 18 IF Nilai Seleksi=bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
19. Rule 19 IF Nilai Seleksi=bagus and Nilai Psikotes=tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
20. Rule 20 IF Nilai Seleksi=bagus and Nilai Psikotes=tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
21. Rule 21 IF Nilai Seleksi=bagus and Nilai Psikotes=tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak
22. Rule 22 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=Ya
23. Rule 23 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=kurang tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak
24. Rule 24 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=kurang memuaskan THEN Kelas Unggulan=tidak
25. Rule 25 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
26. Rule 26 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=sangat tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak
27. Rule 27 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=tinggi and Nilai Wawancara=memuaskan THEN Kelas Unggulan=tidak
28. Rule 28 IF Nilai Seleksi=kurang bagus and Nilai Psikotes=tinggi and Nilai Wawancara=sangat memuaskan THEN Kelas Unggulan=tidak

4. KESIMPULAN

Dari hasil penelitian ini diperoleh kesimpulan penerapan data Mining dengan algoritma ID3 dapat memberikan rule atau aturan untuk kelayakan siswa untuk masuk kelas unggulan atau tidak. Hasil yang diperoleh dengan menggunakan metode ini yaitu berupa pola dalam bentuk pohon keputusan yang diterjemahkan menjadi aturan-aturan (rule) yang berjumlah 28 rule. Dan dari analisis dari pohon keputusan yang terbentuk jika nilai seleksi sangat bagus sudah dapat dipastikan siswa akan masuk kelas unggulan.

REFERENSI

- [1] M. Hasan, "Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Menggunakan Algoritma Naive Bayes Berbasis Forward Selection," Vol. 9, Pp. 317–324, 2017.
- [2] K. Erwansyah, "Implementasi Data Mining Untuk Menganalisa Hubungan Data Penjualan Produk Bahan Kimia Terhadap Persediaan Stok Barang Menggunakan Algoritma FP (Frequent Pattern) Growth Pada PT . Grand Multi Chemicals," *J-SISKO TECH J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 2, no. 2, pp. 30–40, 2019.
- [3] G. Abdillah *et al.*, "Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di PDAM Tirta Raharja Menggunakan Algoritma K-Means," vol. 2016, no. Sentika, pp. 18–19, 2016.
- [4] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Citec J.*, vol. 2, no. 3, pp. 207–217, 2015.
- [5] W. Muslehatin and M. Ibnu, "Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau," pp. 18–19, 2017.
- [6] Y. Elmande and P. P. Widodo, "Pemilihan Criteria Splitting Dalam Algoritma Iterative Dichotomiser 3 (ID3) Untuk Penentuan Kualitas Beras : Studi Kasus Pada Perum Bulog Divre Lampung," vol. 4, no. 1, 2012.