

## PENERAPAN ALGORITMA C4.5 KLASIFIKASI TINGKAT KELULUSAN MAHASISWA

Taufiq<sup>1</sup>, Yulia Yudihartanti<sup>2</sup>

STMIK Banjarbaru

Jl. A. Yani Km 33,3 Loktabat, telp/fax 0511-7482881

pa\_tauw@yahoo.com, yuliydh@gmail.com

### **Abstract**

*STMIK Banjarbaru information system study program is an academic education program after secondary education, which has a burden of at least 144 (one hundred forty four) credits scheduled for 8 (eight) semesters of at most 14 (fourteen) semesters. The study period of students can be known earlier so the information system study program can take the necessary actions so that students can graduate in exactly 8 semesters. The research objective is to analyze and design a decision support system that can help decision making in classifying student graduation rates. The research method used is the C.45 Algorithm method to determine the decision tree pattern. The results of the analysis and design of this decision support system are expected to help in classifying student graduation rates. Based on the evaluation results by comparing the original decision with the decision of mining results C4.5, it can be calculated the percentage of accuracy so as to obtain a value of 94%. With this system, it is expected to overcome the classification problem.*

**Keywords:** c 4.5, graduation, classification

### **Abstrak**

*Program studi sistem informasi STMIK Banjarbaru adalah program pendidikan akademik setelah pendidikan menengah, yang memiliki beban sekurang-kurangnya 144 (seratus empat puluh empat) sks) yang dijadwalkan untuk 8 (delapan) semester paling lama 14 (empat belas) semester. Masa studi mahasiswa dapat diketahui lebih dini maka pihak program studi sistem informasi dapat melakukan tindakan-tindakan yang diperlukan supaya mahasiswa dapat lulus tepat 8 semester. Tujuan penelitian adalah menganalisis dan merancang suatu sistem pendukung keputusan yang dapat membantu pengambilan keputusan dalam mengklasifikasi tingkat kelulusan mahasiswa. Metode penelitian yang digunakan adalah metode Algoritma C.45 untuk menentukan pola pohon keputusan. Hasil dari analisis dan perancangan sistem pendukung keputusan ini diharapkan dapat membantu dalam mengklasifikasikan tingkat kelulusan mahasiswa. Berdasarkan hasil evaluasi dengan membandingkan keputusan asli dengan keputusan hasil mining C4.5, maka dapat dihitung persentase keakuratannya sehingga memperoleh nilai sebesar 94%. Dengan adanya sistem ini diharapkan dapat mengatasi permasalahan klasifikasi.*

**Kata kunci:** c 4.5, kelulusan, klasifikasi

## 1. PENDAHULUAN

Program Sarjana (S1) program studi sistem informasi STMIK Banjarbaru adalah program pendidikan akademik setelah pendidikan menengah, yang memiliki beban sekurang-kurangnya 144 (seratus empat puluh empat) sks) yang dijadwalkan untuk 8 (delapan) semester paling lama 14 (empat belas) semester.

Untuk memenuhi target kelulusan tepat waktu maka mahasiswa harus lebih giat mengikuti perkuliahan dibandingkan kegiatan lainnya. Ada hal lain tentunya yang menghambat mahasiswa tidak mendapat predikat kelulusan tepat waktu, diantaranya : sibuk bekerja, terlalu banyak mengulang matakuliah, punya masalah dengan kampus, dan yang paling penting adalah sibuk menjadi aktivis atau kegiatan kemahasiswaan.[1]

Satuhal yang belum banyak dilakukan perguruan tinggi adalah melakukan antisipasi terhadap mahasiswa yang berpotensi mengalami hambatan dalam belajarnya. Hal ini dikarenakan riwayat akademik pada saat disekolah yang tidak sepenuhnya mendukung dengan jurusan yang dijalani pada perguruan tinggi. Hal ini dianggap penting karena semakin awal mengetahui adanya potensi mahasiswa yang kemungkinan akan mengalami hambatan atau kesulitan dalam belajarnya, maka akan semakin cepat langkah langkah antisipatif yang akan dilakukan. [2]

Jika masa studi mahasiswa dapat diketahui lebih dini maka pihak program studi dapat melakukan tindakan-tindakan yang diperlukan supaya mahasiswa dapat lulus tepat 8 (Delapan) semester.

Decision tree (pohon keputusan) adalah pohon klasifikasi yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dikasuskan dalam decision tree, daerah pengambilan keputusan yang sebelumnya kompleks dapat diubah menjadi lebih sederhana. Banyak algoritma yang dapat dipakai dalam pembentuknya decision tree seperti ID3, CART, C4.5, dan lain-lain. Algoritma adalah urutan langkah-langkah yang logis untuk menyelesaikan suatu masalah. [3]

Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma yang lainnya. Kelebihan algoritma C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan numerik. Dalam mengkonstruksi pohon, algoritma C4.5 membaca seluruh sampel *data training* dari *storage* dan memuatnya ke memori. Hal inilah yang menjadi salah satu kelemahan algoritma C4.5 dalam kategori "skalabilitas" adalah algoritma ini hanya dapat digunakan jika *data training* dapat disimpan secara keseluruhan dan pada waktu yang bersamaan di memori.[4]

Data yang akan digunakan oleh peneliti adalah data lulusan mahasiswa program studi Sistem Informasi STMIK Banjarbaru. Dari kumpulan data tersebut, dapat diketahui informasi kelulusan yang dapat mempengaruhi beberapa keputusan program studi menggunakan data mining dengan algoritma C4.5.

Penelitian terdahulu mengenai metode Algoritma C4.5 telah dilakukan antara lain oleh Risnawati (2018) tentang Analisa kelulusan mahasiswa studi kasus STMIK Royal Kisaran. Perbedaan penelitian dengan penelitian sebelumnya yaitu terletak pada kriteria yang dijadikan acuan dalam penentuan kelulusan mahasiswa.

## 2. METODOLOGI PENELITIAN

Pada penelitian ini yang digunakan adalah data daftar lulusan mahasiswa STMIK Banjarbaru yang menjadi rujukan utama dalam membangun Aplikasi klasifikasi tingkat kelulusan mahasiswa menggunakan metode *Algoritma C4.5* seluruh populasi pada penelitian ini dipergunakan sebagai sampel yang diperoleh dari bagian program studi sistem informasi pada STMIK Banjarbaru.

Metode-metode yang digunakan untuk pengumpulan data dalam penelitian adalah sebagai berikut :

1. Wawancara  
Pengumpulan data dengan melakukan tanya jawab terhadap sumber data atau pihak-pihak yang berkepentingan dalam penelitian, dalam hal ini yaitu STMIK Banjarbaru.
2. Observasi  
Pengumpulan data dengan cara melakukan pengamatan dan pencatatan pada STMIK Banjarbaru sebagai objek penelitian.
3. Studi Kepustakaan  
Pengumpulan data dengan cara mengambil bahan dari dokumentasi, literature, jurnal, buku-buku dan internet yang berhubungan dengan penelitian.

### 2.1 Data Mining

Menurut (Turban, E, Dkk, 2005) Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistis, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakumulasi dari berbagai database besar. (Kusrini, & EmhaTaufiq Luthfi, 2009). [5]

Menurut (Larose, D.T, 2005) Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. [6]

### 2.2 Pengelompokan Data Mining

Menurut Larose (2005) *data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu [6] :

#### 2.2.1 Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Contoh prediksi dalam bisnis dan penelitian adalah :

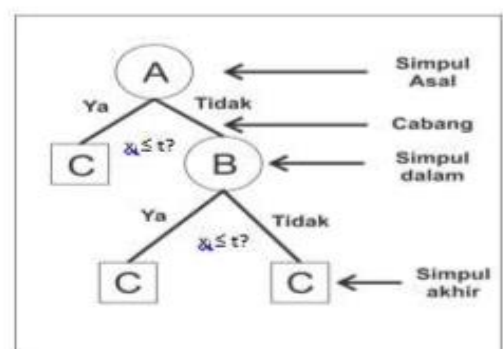
- a. Prediksi harga beras dalam tiga bulan yang akan datang.

- b. Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk memprediksi.

#### 2.2.2 Klasifikasi

Proses pembentukan pohon klasifikasi terdiri atas tiga tahapan, yaitu (Lewis and Roger, 2000) :

- a. Pemilihan (Classifier)
- b. Untuk membentuk pohon klasifikasi digunakan sampel data learning (L) yang masih bersifat heterogen. Setiap pemilihan hanya bergantung pada nilai yang berasal dari suatu variabel independen. Struktur pohon klasifikasi dapat dilihat pada Gambar 3. Penentuan Node Terminal Suatu node t akan menjadi node terminal atau tidak, akan dipilih kembali bila pada node t tidak terdapat penurunan keheterogenan secara berarti atau adanya batasan minimum n seperti hanya terdapat satu pengamatan pada tiap node anak. Umumnya jumlah kasus minimum dalam suatu terminal akhir adalah 5. Dan apabila hal itu terpenuhi maka pengembangan pohon dihentikan.



Gambar 1. Data mining sebagai salah satu tahapan dalam proses Knowledge Disc1ver

c. Penandaan Label Kelas Penandaan label node terminal berdasar aturan jumlah anggota kelas terbanyak. Proses pembentukan pohon klasifikasi berhenti saat terdapat hanya satu pengamatan dalam tiap simpul anak atau adanya batasan minimum n, semua pengamatan dalam tiap node anak identik, dan adanya batasan jumlah kedalaman pohon maksimal. Setelah terbentuk pohon maksimal tahap selanjutnya adalah pemangkasan pohon untuk mencegahnya terbentuk pohon klasifikasi yang berukuran sangat besar dan kompleks, sehingga diperoleh ukuran pohon yang layak. [7]

### 2.3 Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami, juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, pohon keputusan sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain. (Kusrini, & EmhaTaufiq Luthfi, 2009). [5]

### 2.4 Algoritma C4.5

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut [5] :

- Pilih atribut sebagai akar
- Buat cabang untuk masing-masing nilai
- Bagi kasus dalam cabang

- Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S<sub>i</sub>| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sebelum mendapatkan nilai Gain adalah dengan mencari nilai Entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar dari Entropi adalah sebagai berikut:

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i$$

Keterangan:

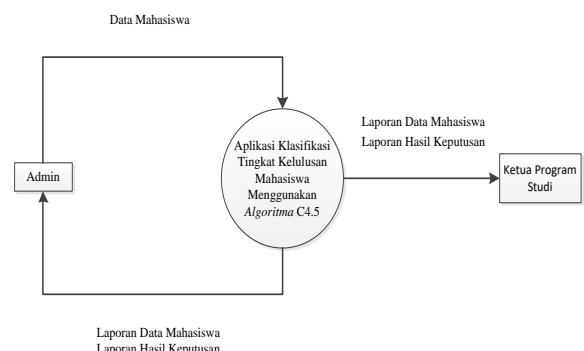
S : Himpunan Kasus

n : Jumlah partisi S

p<sub>i</sub> : Proporsi dari S<sub>i</sub> terhadap S

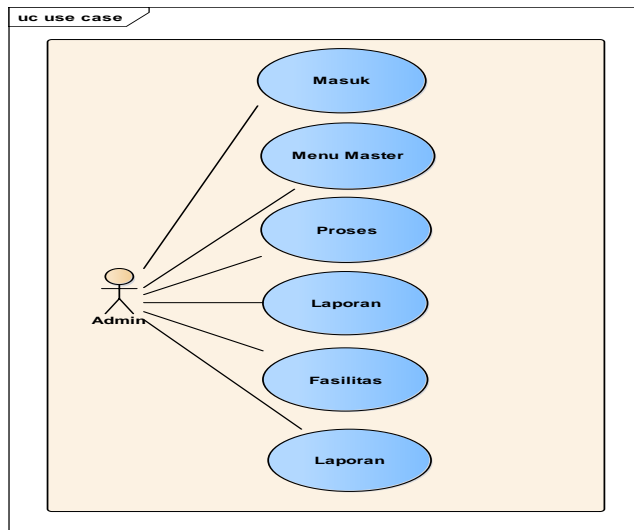
### 2.5 Model Data

Diagram konteks memberikan gambaran admin menginput data Mahasiswa, serta menerima Laporan Data Mahasiswa dan Laporan Hasil Keputusan.



Gambar 2. Diagram Konteks

Gambar *Use Case Diagram* ini menjelaskan interaksi antara admin dan sistem pengguna.



Gambar 3. Use Case Diagram

Dalam diagram ini digambarkan bahwa kegiatan yang terjadi dalam aplikasi ini adalah admin diharuskan *login* untuk dapat menjalankan berbagai fasilitas dalam aplikasi ini yaitu mengelola data master, mengelola penilaian, proses perhitungan serta membuat laporan.

## 2.6 Analisa Data

Data dari sistem adalah berupa atribut yang dimiliki oleh mahasiswa, nilai atributnya, dan nilai kemungkinannya. Data yang dimaksud adalah data yang mempunyai minimal dua kolom atribut. Satu kolom sebagai kolom atribut masukkan dan satu kolom sebagai kolom atribut target. Dari setiap kolom terdapat nilai-nilai yang akan dipergunakan untuk kalkulasi, dan nilai dari setiap atribut harus bersifat diskret. Beberapa komponen variabel yang digunakan yaitu:

### 1. Jenis Kelamin

Variabel ini berisi data jenis kelamin dari mahasiswa yang digunakan untuk pembentukan keputusan.

Pengelompokkan yang ada berdasarkan ketentuan yang dibuat program yaitu L (laki-laki) dan P (Perempuan).

### 2. IPK

Variabel ini berisi data ipk dari mahasiswa yang digunakan untuk pembentukan keputusan. Pengelompokkan yang ada berdasarkan ketentuan yang dibuat program yaitu : CUKUP BAIK, BAIK, MEMUASKAN, SANGAT MEMUASKAN dan CUMLAUDE

### 3. Lama TA

Variabel ini berisi data lama TA dari mahasiswa yang digunakan untuk pembentukan keputusan. Pengelompokkan yang ada berdasarkan ketentuan yang dibuat program yaitu : CEPAT, SEDANG dan LAMA

### 4. Tahun Lulus

Variabel ini berisi data tahun lulus dari mahasiswa yang digunakan untuk pembentukan keputusan. Pengelompokkan yang ada berdasarkan ketentuan yang dibuat program yaitu : 2015 dan 2016

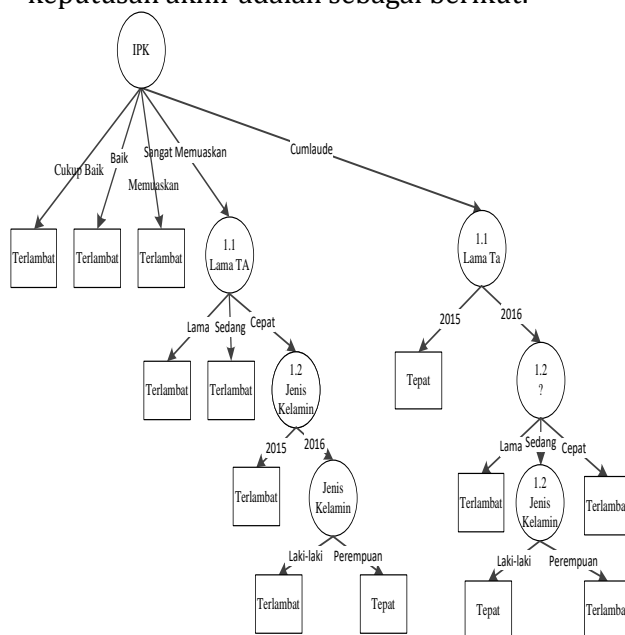
Berikut ini adalah penjelasan lebih terperinci mengenai tiap-tiap langkah dalam pembentukan pohon keputusan dengan menggunakan algoritma C4.5 untuk menyelesaikan permasalahan yang ada pada skripsi. Rumus dalam mencari nilai entropy dan gain untuk menyelesaikan data kasus dari penelitian adalah sebagai berikut :

Tabel 1. Perhitungan Total Kasus

1			Jml Kasus	Terlambat	T e p a t	Ent rop y	Gain
	Total		34	29	5	0.6	
	Jenis Kelamin						0
		Laki-laki	15	13	2	0.57	
		Perempuan	19	16	3	0.63	
	IPK						0.23
		Cukup Baik	4	4	0	0	
		Baik	6	6	0	0	
		Memuaskan	7	7	0	0	
		Sangat Memuaskan	12	10	2	0.65	

		Cumlaude	5	2	3	0.9 7	
	Lama TA	0.11					
	Cepat	15	11	4	0.8 6		
	Sedang	11	10	1	0.5		
	Lama	8	8	0	0		
	Tahun Lulus	0.01					
	2012	10	8	2	0.7 2		
	2013	24	21	3	0.5 3		

Dari hasil tabel data kasus yang telah dilakukan perhitungan menggunakan teknik algoritma C4.5 didapatkan nilai Gain tertinggi yaitu Nilai, sehingga Nilai dijadikan sebagai akar dari pohon keputusan selanjutnya. Setelah dilakukan perhitungan akhir dari data diatas didapat pohon keputusan akhir adalah sebagai berikut:



Gambar 4. Hasil Akhir Pohon Keputusan

Dari hasil perhitungan Level 1 maka ditemukan pohon keputusan Rute utama adalah IPK, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan lima 5 kriteria perhitungan node 1. Dari rute IPK memiliki lima cabang yaitu :

1. Cukup Baik
2. Baik
3. Memuaskan

4. Sangat Memuaskan

5. Cumlaude

Hasil dari perhitungan kriteria IPK maka ditemukan pola, apabila IPK Cukup Baik maka Terlambat, apabila IPK Baik maka Terlambat, apabila IPK Memuaskan maka Terlambat, apabila IPK Sangat Memuaskan dan IPK Cumlaude masih bertanya, maka akan dilanjutkan kembali perhitungan untuk menentukan gain tertinggi.

Dari hasil perhitungan Level 2 maka ditemukan pohon keputusan Rute utama adalah Lama TA, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan tiga (3) kriteria ditabel perhitungan node

1.1. Dari rute Lama TA memiliki tiga (3) cabang yaitu :

1. Cepat
2. Sedang
3. Lama

Hasil dari perhitungan kriteria Lama TA maka ditemukan pola, apabila Lama TA Lama maka Terlambat, apabila lama TA Sedang maka terlambat dan apabila lama TA Cepat masih bertanya, maka akan dilanjutkan kembali perhitungan untuk menentukan gain tertinggi.

Dari hasil perhitungan Level 2 maka ditemukan pohon keputusan Rute utama adalah Tahun Lulus, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan dua (2) kriteria perhitungan node 1.1. Dari rute Tahun Lulus memiliki dua (2) cabang yaitu :

1. 2015
2. 2016

Hasil dari perhitungan kriteria Tahun Lulus maka ditemukan pola, apabila Tahun lulus 2012 lama maka Terlambat, dan apabila Tahun lulus 2013 masih bertanya, maka akan dilanjutkan kembali perhitungan untuk menentukan gain tertinggi.

Dari hasil perhitungan Level 3 maka ditemukan pohon keputusan Rute utama adalah Tahun Lulus, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan dua (2) kriteria perhitungan node 1.2. Dari rute Tahun Lulus memiliki dua (2) cabang yaitu :

1. 2015

2. 2016

Hasil dari perhitungan kriteria Tahun Lulus maka ditemukan pola, apabila Tahun Lulus 2015 maka Terlambat, dan apabila Tahun Lulus 2016 masih bertanya, maka akan dilanjutkan kembali perhitungan untuk menentukan gain tertinggi.

Dari hasil perhitungan Level 3 maka ditemukan pohon keputusan Rute utama adalah Jenis Kelamin, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan dua (2) kriteria perhitungan node 1.3. Dari rute Jenis Kelamin memiliki dua (2) cabang yaitu :

1. Laki-laki
2. Perempuan

Hasil dari perhitungan kriteria Jenis Kelamin maka ditemukan pola, apabila Jenis Kelamin Laki-laki maka Terlambat, dan apabila Jenis Kelamin Perempuan maka Tepat.

Dari hasil perhitungan Level 3 maka ditemukan pohon keputusan Rute utama adalah Lama TA, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan tiga (3) kriteria ditabel 3.19 perhitungan node 1.2. Dari rute Jenis Kelamin memiliki tiga (3) cabang yaitu :

1. Cepat
2. Sedang
3. Lama

Hasil dari perhitungan kriteria Lama TA maka ditemukan pola, apabila Lama TA lama maka Terlambat, apabila Lama TA Cepat maka terlambat dan apabila Lama TA Sedang masih bertanya maka akan dilanjutkan kembali perhitungan untuk menentukan gain tertinggi.

Dari hasil perhitungan Level 3 maka ditemukan pohon keputusan Rute utama adalah Jenis Kelamin, dimana rute utama dilihat dari hasil gain tertinggi, dari perhitungan tiga (2) kriteria ditabel 3.20 perhitungan node 1.3. Dari rute Jenis Kelamin memiliki tiga (2) cabang yaitu :

1. Laki-laki
2. Perempuan

Hasil dari perhitungan kriteria Jenis Kelamin maka ditemukan pola, apabila Jenis Kelamin Perempuan maka Terlambat dan apabila Jenis Kelamin Laki-laki maka Tepat.

Jadi dilakukan perhitungan algoritma C4.5 itu untuk menemukan pola, dimana pola tersebut menghasilkan keputusan penentuan Tingkat Kelulusan Mahasiswa STMIK Banjarbaru, sehingga dengan melihat hasil pola keputusan penentuan Tepat Terlambatnya Mahasiswa STMIK Banjarbaru dengan algoritma C.45.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Hasil



Gambar 5. Menu Utama

Menu utama adalah tampilan depan program yang memuat *link-link* menuju ke *form* yang lain. Ini merupakan prosedur dasar menuju ke menu lainnya. Jika pengguna login sebagai admin, maka pada *form* menu utama memunculkan seluruh menu yang tersedia pada aplikasi seperti, menu master data, proses, laporan, fasilitas dan keluar

Gambar 6. Data Mahasiswa

*Form* ini digunakan untuk memasukkan data mahasiswa pada *Form* ini



digunakan untuk memasukkan nim mahasiswa, nama mahasiswa, tempat lahir, tanggal lahir, jenis kelamin, alamat, no, rt, rw, kelurahan, kecamatan, kota, no telepon dan nama orang tua. Untuk melakukan proses pengelolaan data pada *form* ini ada beberapa tombol yang disediakan yaitu tombol simpan, tambah, ubah, hapus, batal, dan keluar.

Gambar 7. Form Proses Data Mahasiswa

*Form* ini menampilkan hasil pemrosesan data mahasiswa dengan Metode *C4.5*. *Form* ini digunakan untuk memproses hasil dari Metode *C4.5*.

Gambar 8. Form Proses *C4.5* Pengelompokan Data

*Form* ini digunakan untuk menghitung jumlah kasus *entropy* dan gain dari tiap-tiap atribut. Gain yang dihasilkan dari tiap-tiap atribut akan dihitung kembali untuk menentukan gain atribut mana yang tertinggi, gain tertinggi yang didapat akan

digunakan untuk menentukan hasil keputusan.

### 3.2 Pengujian Sistem

*Pretest* dan *Posttest* adalah untuk membandingkan antara sebelum dan sesudah dibangunnya sistem. Berikut tabel perbandingan pengujian *pretest* dan *posttest* yaitu, sebagai berikut: Adapun perhitungan Reabilitas Sebagai Berikut :

Tabel 2. Perbandingan Pengujian *Pretest* dan *Posttest*

No	NIM	Kesesuaian		Ket
		Sebelum	Sesudah	
1	Mahasiswa 1	Terlambat	Terlambat	Sesuai
2	Mahasiswa 2	Terlambat	Terlambat	Sesuai
3	Mahasiswa 3	Terlambat	Terlambat	Sesuai
4	Mahasiswa 4	Terlambat	Terlambat	Sesuai
5	Mahasiswa 5	Terlambat	Tepat	Tidak Sesuai
6	Mahasiswa 6	Tepat	Tepat	Sesuai
7	Mahasiswa 7	Terlambat	Tepat	Tidak Sesuai
8	Mahasiswa 8	Terlambat	Terlambat	Sesuai
9	Mahasiswa 9	Tepat	Tepat	Sesuai
10	Mahasiswa 10	Terlambat	Terlambat	Sesuai
11	Mahasiswa 11	Terlambat	Terlambat	Sesuai
12	Mahasiswa 12	Terlambat	Terlambat	Sesuai
13	Mahasiswa 13	Terlambat	Terlambat	Sesuai
14	Mahasiswa 14	Terlambat	Terlambat	Sesuai
15	Mahasiswa 15	Terlambat	Terlambat	Sesuai
16	Mahasiswa 16	Terlambat	Terlambat	Sesuai
17	Mahasiswa 17	Terlambat	Terlambat	Sesuai
18	Mahasiswa 18	Terlambat	Terlambat	Sesuai
19	Mahasiswa 19	Terlambat	Terlambat	Sesuai
20	Mahasiswa 20	Terlambat	Terlambat	Sesuai



21	Mahasiswa 21	Terlambat	Terlambat	Sesuai
22	Mahasiswa 22	Terlambat	Terlambat	Sesuai
23	Mahasiswa 23	Terlambat	Terlambat	Sesuai
24	Mahasiswa 24	Terlambat	Terlambat	Sesuai
25	Mahasiswa 25	Tepat	Tepat	Sesuai
26	Mahasiswa 26	Terlambat	Terlambat	Sesuai
27	Mahasiswa 27	Terlambat	Terlambat	Sesuai
28	Mahasiswa 28	Terlambat	Terlambat	Sesuai
29	Mahasiswa 29	Terlambat	Terlambat	Sesuai
30	Mahasiswa 30	Terlambat	Terlambat	Sesuai
31	Mahasiswa 31	Terlambat	Terlambat	Sesuai
32	Mahasiswa 32	Terlambat	Terlambat	Sesuai
33	Mahasiswa 33	Tepat	Tepat	Sesuai
34	Mahasiswa 34	Tepat	Tepat	Sesuai

Keterangan:

: Urutan data yang sesuai sebelum dilakukan proses C4.5.

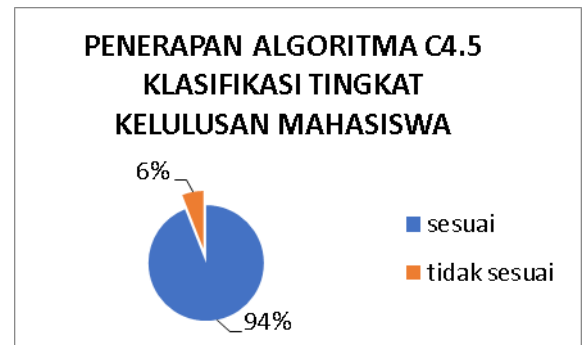
: Urutan data setelah dilakukan proses C4.5.

Maka dengan membandingkan antara data *pretest* dan data hasil *posttest* dengan sampel data yang berjumlah 34 data, maka didapat hasil 2 data yang tidak sesuai setelah dilakukan proses perhitungan C4.5. Dengan demikian didapat hasil untuk menghitung hasil perbandingan sebagai berikut:

$$\frac{\text{Jumlah Keputusan Yang Sesuai}}{\text{Jumlah Mahasiswa Keseluruhan}} \times 100 \%$$

Nilai persentase kesesuaian sebelum dan sesudah adanya aplikasi =  $\frac{32}{34} \times 100 \% = 94 \%$

Dari hasil perbandingan data maka diperoleh grafik perbandingan sebagai berikut:



Gambar 9. Grafik Perbandingan Pretest dan Posttest

#### 4. SIMPULAN

Aplikasi data mining dengan metode klasifikasi algoritma C4.5 ini dapat memproses data mahasiswa menjadi sebuah aturan-aturan yang berguna sebagai masukan penentu keputusan walaupun jumlah data dapat mempengaruhi persentase keakuratannya.

Berdasarkan hasil evaluasi dengan membandingkan keputusan asli dengan keputusan hasil mining C4.5, maka dapat dihitung persentase keakuratan kesesuaian sehingga memperoleh nilai sebesar 94 %.

#### DAFTAR PUSTAKA

- [1] Risnawati., "Analisis Kelulusan Mahasiswa Menggunakan Algoritma C.45", Jurnal Mantik Penusa Vol. 2, No. 1, pp.71-76, Juny 2018.
- [2] Muhammad Ardiansyah Sembiring, "Penerapan Metode Decission Tree Algoritma C45 Untuk Memprediksi Hasil Belajar Mahasiswa Berdasarkan Riwayat Akademik", Jurteksi, vol. 3. No. 1. Pp. 60-65, dec 2016.
- [3] Prasetyo, E, "Data Mining: Konsep dan Aplikasi Menggunakan Matlab", Andi Offset, Yogyakarta, 2012.
- [4] Fayyad, U, *Advances in Knowledge Discovery and Data Mining*. MIT Press, Yogyakarta, 2016.

- [5] Kusrini, dan E. T. Luthfy, **Algoritma Data Mining**, Andi Offset, Yogyakarta, 2009.
- [6] Larose, **Discovering Knowledge in Data: An Introduction to Data Mining**, New Jersey, John Willey and Sons, 2005.
- [7] Lewis, R.J, **An Introduction to Classification and regression tree Analysis**, Annual Meeting of the Society for Academic Emergency Medicine in San Fransisco, California: Departement of Emergency Medicine, 2000.