

PENERAPAN METODE C4.5 UNTUK KLASIFIKASI MAHASISWA BERPOTENSI DROP OUT

Asmaul Husnah Nasrullah

asmaulhusnel@gmail.com

UNIVERSITAS ICHSAN GORONTALO

Abstrak

Kualitas pendidikan di perguruan tinggi dapat dilihat dari tingginya tingkat keberhasilan mahasiswa dan rendahnya kegagalan mahasiswa. Salah satu indikator kegagalan mahasiswa adalah kasus Drop Out (berhenti studi). Permasalahan Drop Out menjadi sesuatu yang menarik untuk diteliti, karena hal ini dapat mempengaruhi kualitas pendidikan tersebut. Fakultas Ekonomi UNISAN Gorontalo merupakan Fakultas favorit di UNISAN Gorontalo sehingga memiliki jumlah mahasiswa kurang lebih 1000 mahasiswa sampai tahun 2017. Namun rasio jumlah mahasiswa lulus dan tidak lulus tidak seimbang. Sehingga menghasilkan jumlah mahasiswa Drop Out kurang lebih 200 mahasiswa pertahunnya. Untuk mengatasi masalah tersebut diusulkan model baru dengan memanfaatkan sebuah metode komputasi C4.5, agar menghasilkan pola berdasarkan hasil klasifikasi yang tepat dalam menentukan mahasiswa yang berpotensi Drop Out. Adapun hasil yang diperoleh dari penerapan metode C4.5 pada penelitian ini yaitu ditemukannya 17 rule yang dapat dijadikan sebagai pola untuk menentukan mahasiswa yang berpotensi Drop Out.

Kata kunci: Drop Out, Klasifikasi, Metode C4.5

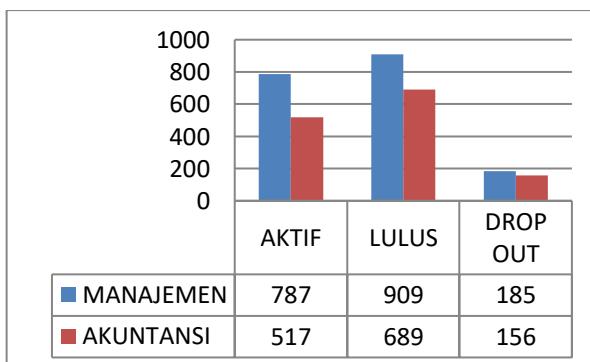
Abstract

The quality of education in universities can be seen from the high level of student success and the low failure of students. One indicator of student failure is the case of Drop Out (stop study). The problem of Drop Out becomes something interesting to study, because this can affect the quality of education. Faculty of Economics UNISAN Gorontalo is a favorite Faculty in UNISAN Gorontalo so it has a number of students of approximately 1000 students until 2017. But the ratio of the number of graduate students and not pass unbalanced. So as to produce the number of students Drop Out approximately 200 students per year. To solve the problem, we proposed a new model by utilizing a C4.5 computation method, in order to produce a pattern based on the results of the correct classification in determining the potential Drop Out students. The results obtained from the application of method C4.5 in this research is the discovery of 17 rules that can be used as a pattern to determine the potential students Drop Out.

Keywords: Drop Out, Classification, C4.5 Method

1. Pendahuluan

Universitas Ichsan Gorontalo (UNISAN) berdiri sejak tahun 2001, merupakan salah satu Perguruan Tinggi Swasta terbaik yang ada di Kota Gorontalo dan memiliki 5 Fakultas. UNISAN bertujuan menghasilkan lulusan berkualitas, inovatif, kreatif serta memiliki semangat kerakyatan yang mampu mendukung pembangunan nasional, memajukan masyarakat serta mempertinggi mutu kebudayaan. Fakultas Ekonomi adalah salah satu fakultas yang memiliki mahasiswa terbanyak diantara fakultas yang ada. Fakultas ini terdiri dari 2 program studi yaitu Manajemen dan Akuntansi. Berikut jumlah mahasiswa yang berstatus aktif, lulus dan Drop Out sampai dengan tahun akademik 2016/2017 [1].



Gambar 1. Rekap Data Mahasiswa

Tingginya tingkat keberhasilan mahasiswa dan rendahnya tingkat kegagalan mahasiswa merupakan cermin kualitas dari suatu Perguruan Tinggi. Kegagalan mahasiswa dan faktor-faktor penyebabnya merupakan topik yang menarik untuk diteliti. Saat ini, Institusi Perguruan Tinggi di Gorontalo berada dalam lingkungan yang sangat kompetitif, setiap perguruan tinggi berusaha untuk terus memperbaiki manajemennya demi meningkatkan mutu pendidikan dan meningkatkan akreditasi. Berdasarkan buku panduan akademik UNISAN alas an mahasiswa Drop Out terbagi tiga poin yaitu; 1).

Drop Out karena nonaktif selama 3 tahun akademik, 2) Drop Out karena Prestasi Akademik (tidak menyelesaikan 15 sks di 2 semester pertama dan IPK < 2.00 dan 3) Drop Out karena melewati masa Maksimal Studi. Keberadaan mahasiswa drop out di UNISAN dimulai dari ketidak aktifan mahasiswa sendiri di masa perkuliahan dan ini tentunya berpengaruh terhadap mahasiswa lulus tidak tepat waktu. Semakin banyak mahasiswa non-aktif maka semakin banyak juga mahasiswa lulus tidak tepat waktu yang berakibat besarnya peluang untuk drop out.

>>Semakin banyak mahasiswa non-aktif maka semakin banyak juga mahasiswa lulus tidak tepat waktu yang berakibat besarnya peluang untuk drop out.<<

<< lulus tepat waktu berarti lulus dalam kurun waktu 4 thn (8 smstr) krn status mhswa nonaktif (tdk aktif selama 6 smst) maka lulusnya tidak tepat waktu, dan jika lulus tidak tepat waktu maka berpotensi drop out krn lama masa studi akan bertambah.

>> Apa yang dimaksud mahasiswa non aktif di sini? Apakah mahasiswa lulus tidak tepat waktu = mahasiswa drop out?

<< mahasiswa yg tidak mengurus KRS selama 6 smstr, mhs lulus tidak tepat waktu tidak sama dengan mhs drop out. Tapi mahasiswa yg lulus tdk tepat waktu berpotensi menjadi DO.

Berdasarkan uraian diatas, mahasiswa yang berpotensi drop out dapat menimbulkan masalah yang dapat menurunkan mutu pendidikan dan akreditasi perguruan tinggi. oleh karena itu, diperlukan analisa terhadap data-data mahasiswa Drop out yang telah lalu sehingga dapat diketahui karakteristik mahasiswa drop out sejak dulu. Berdasarkan data mahasiswa yang dimiliki fakultas ekonomi yang cukup besar sehingga bisa dimanfaatkan dalam pengolahan data mining. Dimana data mining digunakan untuk menggali dan mendapatkan informasi dari data dengan jumlah besar [3]. Salah satu metode data mining adalah pengklasifikasian data. Klasifikasi merupakan kegiatan dalam mengekstrak dan kemudian memprediksi label kategori untuk masing-masing data. Dengan kata lain, klasifikasi adalah proses pencarian model yang dapat membedakan kelas data dengan tujuan agar model tersebut dapat digunakan untuk memprediksi kelas dari suatu obyek yang belum diketahui kelasnya. Salah satu algoritma klasifikasi yang bisa digunakan adalah C4.5 [4].

Permasalahan diatas dapat diatasi dengan melakukan penelitian lebih lanjut dengan membuat suatu aplikasi data mining yang akan digunakan untuk memprediksi mahasiswa yang memasuki semester V ke atas apakah berpotensi Drop Out atau tidak dengan menggunakan atribut yaitu Jenis Kelamin, Umur, Agama, Kelas, Asal Daerah, Dan Nilai IPS Semester I – Semester IV maka penulis menggunakan model pohon keputusan algoritma C4.5. Model ini dipilih karena proses learning dan klasifikasi pada algoritma C4.5 sederhana dan cepat [4][5]. Secara umum, model algoritma C4.5 mempunyai tingkat akurasi yang tinggi.

Berdasarkan latar belakang, maka permasalahan dapat dirumuskan yaitu sebagai berikut “Bagaimana hasil penerapan algoritma C4.5 untuk klasifikasi mahasiswa yang berpotensi mengalami drop out ? ”. Penelitian ini bertujuan agar dapat diketahui hasil penerapan algoritma C4.5 dalam melakukan klasifikasi terhadap mahasiswa yang berpotensi drop out.

Penelitian ini diharapkan dapat memberikan hasil berupa sistem / aplikasi yang dapat digunakan oleh Fakultas Ekonomi (khususnya) dan seluruh Fakultas di Universitas Ichsan Gorontalo (umumnya) dalam mengelompokkan (meng-klasifikasikan) mahasiswa yang berpotensi drop out. Serta dapat memberikan acuan bagi pihak pengambil kebijakan dalam menentukan strategi terkait permasalahan drop out tersebut.

2. Metode

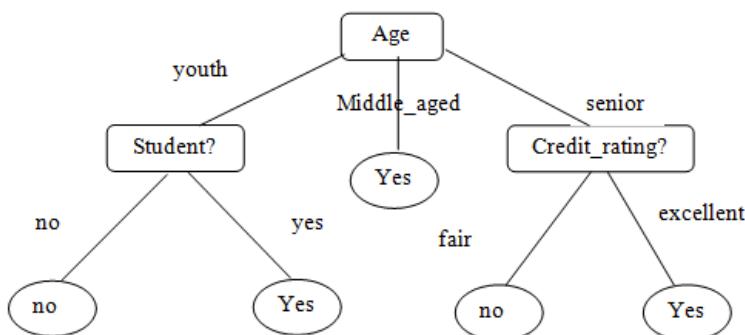
2.1 Algoritma C4.5

Pada akhir tahun 1970 sampai awal tahun 1980 J. Ross Quinlan, seorang peneliti di bidang machine learning, membuat sebuah algoritma decision tree yang dikenal dengan ID3 (Iterative Dichotomiser). Quinlan kemudian membuat algoritma C4.5 (sering disebut dengan pohon keputusan) yang merupakan pengembangan dari algoritma ID3 [10]. Algoritma ini memiliki kelebihan, yaitu mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan



dalam bentuk gambar (pohon keputusan) [11]. Algoritma C4.5 merupakan struktur pohon dimana terdapat simpul yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas. Algoritma C4.5 secara rekursif mengunjungi setiap simpul keputusan, memilih pembagian yang optimal, sampai tidak bisa dibagi lagi. Algoritma C4.5 menggunakan konsep information gain atau entropy reduction untuk memilih pembagian yang optimal [10].

Pohon keputusan mirip sebuah struktur pohon dimana terdapat node internal (bukan daun) yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas. Pohon keputusan bekerja mulai dari akar paling atas, jika diberikan sejumlah data uji, misalnya X dimana kelas dari data X belum diketahui, maka pohon keputusan akan menelusuri mulai dari akar sampai node dan setiap nilai dari atribut sesuai data X diuji apakah sesuai dengan aturan pohon keputusan, kemudian pohon keputusan akan memprediksi kelas dari tupel X.



Gambar 2. Contoh Konsep Pohon Keputusan untuk memprediksi kelas dari tupel X.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [4], yaitu :

1. Menyiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
2. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus:

$$Entropi(S) = \sum_{i=1}^k -p_i \log_2 p_i \quad \dots \dots \dots \quad (1)$$

Keterangan:

S = himpunan kasus

k = jumlah partisi S

p_i = probabilitas yang didapat dari jumlah (ya/tidak) dibagi total kasus

3. Kemudian hitung nilai gain menggunakan rumus:

$$Gain(S, A) = Entropi(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropi(S_i) \quad \dots \dots \dots \quad (2)$$

Keterangan:

S = himpunan kasus

A = fitur

n = jumlah partisi atribut A

$$\begin{cases} |S_i| &= \text{proporsi } S_i \text{ terhadap } S \\ |S'| &= \text{jumlah kasus dalam } S' \end{cases}$$

4. Ulangi langkah ke-2 hingga semua record terpartisi.
5. Proses partisi pohon keputusan akan berhenti saat :
 - a. Semua record dalam simpul N mendapat kelas yang sama.
 - b. Tidak ada atribut di dalam record yang dipartisi lagi.
 - c. Tidak ada record di dalam cabang yang kosong

2.2 Penelitian Terkait

Berikut ini adalah beberapa penelitian terdahulu yang terkait dengan algoritma C4.5, secara garis besar tinjauan studi untuk penelitian ini meliputi: Penelitian yang dilakukan oleh Siti Nurhayati, Kusrini, Emha Taufiq Luthfi, 2016, Prediksi Mahasiswa Drop Out Menggunakan Metode Support Vector Machine. Tingginya tingkat keberhasilan mahasiswa dan rendahnya tingkat kegagalan mahasiswa dapat mencerminkan kualitas dari suatu perguruan tinggi. Salah satu indikator kegagalan mahasiswa adalah kasus drop out. Untuk mengatasi permasalahan, dilakukan prediksi menggunakan metode support vector machine. Support Vector Machine berusaha mencari hyperplane yang optimal dimana dua kelas pola dapat dipisahkan dengan maksimal, parameter yang digunakan pada Support Vector Machine hanya parameter kernel dalam satu parameter C yang memberikan pinalti pada titik data yang diklasifikasikan secara acak. Dalam Support Vector Machine bobot (w) dan bias (b) merupakan solusi global optium dari quadratic programming sehingga cukup dengan sekali running akan menghasilkan solusi yang akan selalu sama untuk pilihan kernel dan parameter yang sama. Melalui penerapan support vector machine diharapkan untuk mendapatkan parameter Support Vector Machine yang digunakan tepat untuk memperoleh margin terbaik dalam memprediksi mahasiswa drop out[6].

Penelitian yang dilakukan oleh Khafizh Hastuti, 2015, Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif. Mahasiswa non aktif adalah mahasiswa yang berhenti studi dan tidak melakukan registrasi administratif. Mahasiswa yang memiliki status non aktif memiliki kecenderungan untuk drop out. Tingginya persentase mahasiswa dengan status non aktif mempengaruhi nilai akreditasi universitas. Perlu diketahui faktor-faktor penyebab mahasiswa memiliki status non aktif. Teknik klasifikasi data mining dapat digunakan untuk prediksi mahasiswa non aktif. Banyak algoritma klasifikasi data mining yang dapat digunakan, sehingga perlu dilakukan komparasi untuk mengetahui tingkat akurasi dari masing-masing algoritma. Algoritma yang digunakan adalah logistic regression, decision tree, naïve bayes dan neural network. Data yang digunakan sebanyak 3861 mahasiswa program studi Teknik Informatika, Sistem Informasi dan Desain Komunikasi Visual Universitas Dian Nuswantoro. Hasil dari proses klasifikasi dievaluasi dengan menggunakan cross validation, confusion matrix, ROC Curve dan T-Test untuk mengetahui algoritma klasifikasi data mining yang paling akurat untuk prediksi mahasiswa non aktif [7].

Penelitian yang dilakukan oleh Dwi Untari, 2015, Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5. Keberhasilan dan kegagalan studi mahasiswa merupakan cerminan kualitas suatu perguruan tinggi. Mahasiswa non-aktif adalah mahasiswa yang tidak melakukan registrasi pada awal semester atau tidak mengikuti perkuliahan sekurang-kurangnya satu semester. Keberadaan mahasiswa non-aktif ini tentunya berpengaruh terhadap mahasiswa lulus tidak tepat waktu, semakin banyak mahasiswa non-aktif maka semakin banyak juga mahasiswa lulus tidak tepat waktu selain itu mahasiswa non-aktif dapat meningkatkan jumlah mahasiswa berpotensi drop out. Lulus tepat waktu merupakan elemen penilaian akreditasi perguruan tinggi. Oleh karena itu untuk mengatasi masalah tersebut diterapkan data mining dengan metode C.4.5 untuk mencari karakteristik mahasiswa berpotensi non-aktif. Desain penelitian ini menggunakan CRISP-DM dan penelitian menggunakan data mahasiswa S1 Fakultas Ilmu Komputer Universitas Dian Nuswantoro. Proses validasi yang digunakan adalah split validation, dan untuk pengujian model menggunakan confusion matrix. Hasil menunjukkan nilai akurasi terbaik adalah 97.60% dengan ratio data training 90%. Berdasarkan percobaan yang dilakukan, mahasiswa dengan IPS Semester 3 adalah dibawah 2.60, SKS semester 3 yang diambil adalah <20 sks, SKS semester 4 adalah kosong dan IPS semester 4 adalah kosong lebih besar berpotensi menjadi mahasiswa non-aktif [8].

3. Hasil dan Pembahasan

Pada bagian ini akan dijelaskan mengenai tahapan penelitian untuk memperoleh hasil klasifikasi mahasiswa yang berpotensi Drop out menggunakan metode C4.5. adapun tahapan tersebut adalah tahap persiapan data set, penentuan akar, dan penentuan rule. dan klasifikasi. Data set yang digunakan adalah data mahasiswa FEKON UNISAN Gorontalo yang diperoleh dari BAAK UNISAN Gorontalo tahun akademik 2016/2017.

1. Persiapan Data Set

Data set yang digunakan dalam penelitian ini dikomposisikan sebagai berikut:

Tabel 1. Kelompok Dataset

Kelas target	Data yang digunakan			
	Data training 50%	Data validasi 25%	Data testing 25%	Jumlah data
Non DO	120	50	50	220
DO	100	50	50	200



Jumlah	220	100	100	420
--------	-----	-----	-----	-----

Selanjutnya data disiapkan dengan melakukan seleksi dan transformasi data set (lampiran a). Seleksi data bertujuan untuk mengidentifikasi variable-variabel relevan yang digunakan dalam penelitian ini. Dan transformasi data digunakan untuk mengubah dataset sehingga konten informasi terbaik diambil dan dimasukkan pada tool mining yang digunakan dalam format yang tepat.

Tahapan berikutnya yang dilakukan pada penelitian ini yaitu mengikuti cara kerja metode C4.5 mulai dari training data, menghitung entropi hingga menghasilkan rule yang dijadikan sebagai pola untuk penentuan mahasiswa yang berpotensi Drop Out. Atribut adalah bagian data, yang mewakili karakteristik atau fitur dari objek data. Atribut, dimensi, fitur dan variabel sering digunakan secara bergantian dalam literatur. Atribut yang digunakan dalam penelitian ini diambil berdasarkan referensi dari penelitian sebelumnya.

Tabel 2. Atribut Klasifikasi Potensi mahasiswa Drop Out

No	Atribut	Nilai Atribut
1	Jenis Kelamin	Laki-laki Perempuan
2	Umur	<20 Tahun 20 s/d 25 Tahun >25 Tahun
3	Agama	Muslim Non Muslim
4	Asal Daerah	Dalam Kota Gorontalo Luar Kota Gorontalo Luar Provinsi
5	Kelas	Kelas Reguler Kelas Karyawan
6	Nilai IPS Semester I	<2,75 2,75 - 3,00 >3,00
7	Nilai IPS Semester II	<2,75 2,75 - 3,00 >3,00
8	Nilai IPS Semester III	<2,75 2,75 - 3,00 >3,00
9	Nilai IPS Semester IV	<2,75 2,75 - 3,00 >3,00

* Ket: Nilai IPS (Nilai mata pelajaran Ilmu Pengetahuan Sosial)
Sembilan (9) atribut tersebut kemudian diolah menggunakan algoritma C4.5 berdasarkan dataset yang telah di seleksi (lampiran a).

2. Penentuan Akar (Node)

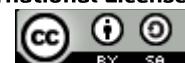
Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Nilai entropi yang diperoleh berdasarkan persamaan (1) adalah sebagai berikut :

$$Entropi(S) = \left(-\left(\frac{41}{98}\right) x \log_2 \left(\frac{41}{98}\right) \right) + \left(-\left(\frac{57}{98}\right) x \log_2 \left(\frac{57}{98}\right) \right) = 0.9806857$$

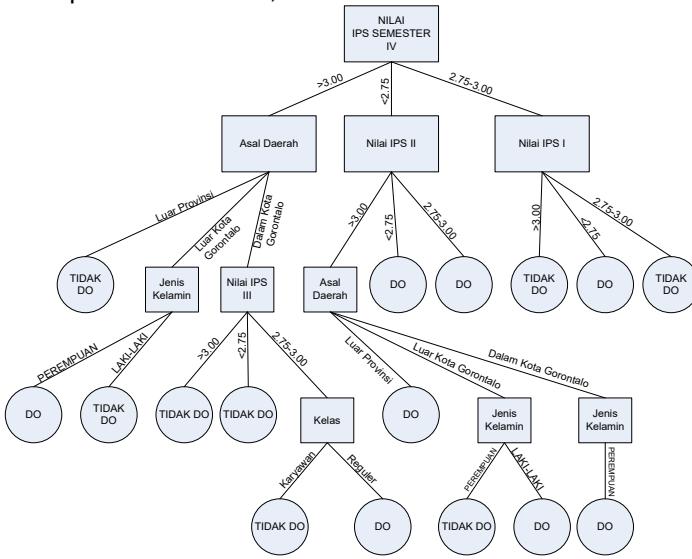
Tabel 3.Hasil Perhitungan pada Dataset

Total Kasus	Sum(DO)	Sum (TIDAK DO)	EntropiTatal
98	41	57	0.9806857

Setelah nilai entropi didapatkan dari keseluruhan kasus,dilakukan analisis pada setiap atribut dan nilai-nilainya. Hasil perhitungan nilai entropi setiap atribut dapat dilihat pada (lampiran b). Nilai Gain berdasarkan persamaan (2) dihitung untuk setiap atribut yang digunakan (lampiran b).



Selanjutnya dibuat pohon keputusan berdasarkan node yang telah ditentukan, hingga menghasilkan node keputusan terakhir seperti di bawah ini;



Gambar 3. Node Keputusan Terakhir

Pada node keputusan terakhir dapat dilihat bahwa atribut agama tidak lagi dijadikan sebagai Node sebab nilai gain yang diperoleh untuk atribut agama pada perhitungan gain Node ke 2 diperoleh nilai gain atribut agama sebesar -4.53.

<< 200 dari dataset

<< 98 data yang telah di cleaning

3. Klasifikasi (Rules)

Selanjutnya hasil dari pohon keputusan tersebut di ekstrasi menjadi rule seperti pada tabel di bawah ini :

Tabel 4. Rule berdasarkan node akhir

Id	Rule
1	JIKA Nilai IPS IV = 2.75 - 3.00 DAN Nilai IPS I = 2.75 - 3.00 MAKA Klasifikasi =TIDAK DO
2	JIKA Nilai IPS IV = 2.75 - 3.00 DAN Nilai IPS I = <2.75 MAKA Klasifikasi =DO
3	JIKA Nilai IPS IV = 2.75 - 3.00 DAN Nilai IPS I = >3.00 MAKA Klasifikasi =TIDAK DO
4	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = 2.75 - 3.00 MAKA Klasifikasi =DO
5	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = <2.75 MAKA Klasifikasi =DO
6	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Luar Provinsi MAKA Klasifikasi =TIDAK DO
7	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = >3.00 DAN Asal Daerah = Luar Provinsi MAKA Klasifikasi = DO
8	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Nilai IPS III = <2.75 MAKA Klasifikasi = TIDAK DO
9	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Nilai IPS III = >3.00 MAKA Klasifikasi = TIDAK DO
10	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Luar Kota Gorontalo DAN Jenis Kelamin = Laki-Laki MAKA Klasifikasi = TIDAK DO
11	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Luar Kota Gorontalo DAN Jenis Kelamin = Perempuan MAKA Klasifikasi = DO
12	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Jenis Kelamin = Laki-Laki MAKA Klasifikasi = TIDAK DO
13	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Jenis Kelamin = Perempuan MAKA Klasifikasi = DO
14	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = >3.00 DAN Asal Daerah = Luar Kota Gorontalo DAN Jenis Kelamin = Laki-Laki MAKA Klasifikasi = DO
15	JIKA Nilai IPS IV = <2.75 DAN Nilai IPS II = >3.00 DAN Asal Daerah = Luar Kota

	Gorontalo DAN Jenis Kelamin = Perempuan MAKA Klasifikasi = TIDAK DO
16	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Nilai IPS III = 2.75 - 3.00 DAN Kelas = Karyawan MAKA Klasifikasi = TIDAK DO
17	JIKA Nilai IPS IV = >3.00 DAN Asal Daerah = Dalam Kota Gorontalo DAN Nilai IPS III = 2.75 - 3.00 DAN Kelas = Reguler MAKA Klasifikasi = DO

Tujuh belas (17) rule tersebut merupakan 17 kelas yang dihasilkan dalam mengelompokkan mahasiswa yang berpotensi Drop Out. Selanjutnya ke-17 kelas (rules) tersebut dapat dijadikan sebagai kelas/kelompok dalam menentukan mahasiswa yang berpotensi drop out sebelum diterima menjadi mahasiswa di FEKON UNISAN Gorontalo. Akurasi dan tingkat error tidak dihitung dikarenakan hanya melihat hasil klasifikasi mahasiswa yang Drop Out dan tidak Drop Out

4. Kesimpulan dan Saran

4.1 Kesimpulan

Adapun kesimpulan yang diperoleh dari penelitian yang dilakukan adalah sebagai berikut :

1. Metode C4.5 yang digunakan dalam klasifikasi mahasiswa yang berpotensi drop out menghasilkan tujuh belas (17) rule .
2. 17 rule tersebut dapat dijadikan sebagai pola dalam menentukan mahasiswa yang berpotensi mengalami drop out.

4.2 Saran

Penelitian yang dilakukan masih jauh dari kesempurnaan, sehingga dibutuhkan beberapa saran diantaranya:

1. Perlunya Pengumpulan Sampel yang lebih banyak yang dapat mewakili populasi, valid, dan tidak biasa, terutama pada data testing agar dapat menghasilkan model yang efisien dan efektif untuk digunakan dalam melakukan klasifikasi penentuan mahasiswa berpotensi drop out.
2. Pola yang ditemukan sebaiknya dibuatkan aplikasi sehingga dapat memudahkan FEKON UNISAN dalam mengolah data mahasiswa.

Daftar Pustaka

- [1] BAAK Unisan, 2016
- [2] Panduan Akreditasi Fakultas Ekonomi UNISAN Gorontalo,2014
- [3] Witten, I. H., Frank, E., & Hall, M. A., "Data Mining: Practical Machine Learning and Tools" Burlington : Morgan Kaufmann Publisher,2011.
- [4] Kusrini, "Konsep dan Aplikasi", Yogyakarta, Andi,2009.
- [5] Kusrini&Luthfi,E.T., "Algoritma Data Mining",Yogyakarta,Andi Publishing, 2009.
- [6] Siti Nurhayati,"Prediksi Mahasiswa Drop Out Menggunakan Metode Support Vector Machine", 2016.
- [7] Khafizh Hastuti," Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif",2015.
- [8] Dwi Untari,"Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5",2015.
- [9] panduan akademik UNISAN,2011.
- [10] Han, J., & Kamber, M., "Data Mining: Concepts and Techniques. Morgan Kaufmann", 2006.
- [11] Gorunescu, Florin,"Data Mining: Concepts, Models, and Techniques",Verlag Berlin Heidelberg : Springer, . 2011.

