

**ANALISIS KINERJA ALGORITMA C4.5 UNTUK PREDIKSI
KETEPATAN KELULUSAN MAHASISWA S1 INFORMATIKA
UNIVERSITAS AMIKOM YOGYAKARTA**

SKRIPSI



disusun oleh

Muhson Nawawi

17.11.1102

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2020**

**ANALISIS KINERJA ALGORITMA C4.5 UNTUK PREDIKSI
KETEPATAN KELULUSAN MAHASISWA S1 INFORMATIKA
UNIVERSITAS AMIKOM YOGYAKARTA
SKRIPSI**

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Muhson Nawawi

17.11.1102

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2020**

PERSETUJUAN

SKRIPSI

ANALISIS KINERJA ALGORITMA C4.5 UNTUK PREDIKSI KETEPATAN KELULUSAN MAHASISWA S1 INFORMATIKA UNIVERSITAS AMIKOM YOGYAKARTA

yang dipersiapkan dan disusun oleh

Muhson Nawawi

17.11.1102

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 20 Maret 2020

Dosen Pembimbing,

Asro Nasiri, Drs, M.Kom.

NIK. 190302152

PENGESAHAN

SKRIPSI

ANALISIS KINERJA ALGORITMA C4.5 UNTUK PREDIKSI KETEPATAN KELULUSAN MAHASISWA S1 INFORMATIKA UNIVERSITAS AMIKOM YOGYAKARTA

yang dipersiapkan dan disusun oleh

Muhson Nawawi

17.11.1102

telah dipertahankan di depan Dewan Penguji

pada tanggal 20 Oktober 2020

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Sharazita Dyah Anggita, M.Kom

NIK. 190302285

Rini Indrayani, ST, M.Eng

NIK. 190302417

Asro Nasiri, Drs, M.Kom

NIK. 190302152

Skripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 20 Oktober 2020

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si, MT

NIK. 190302038

PERNYATAAN KEASLIAN

PERNYATAAN KEASLIAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 20 Oktober 2020



Muhsan Nawawi

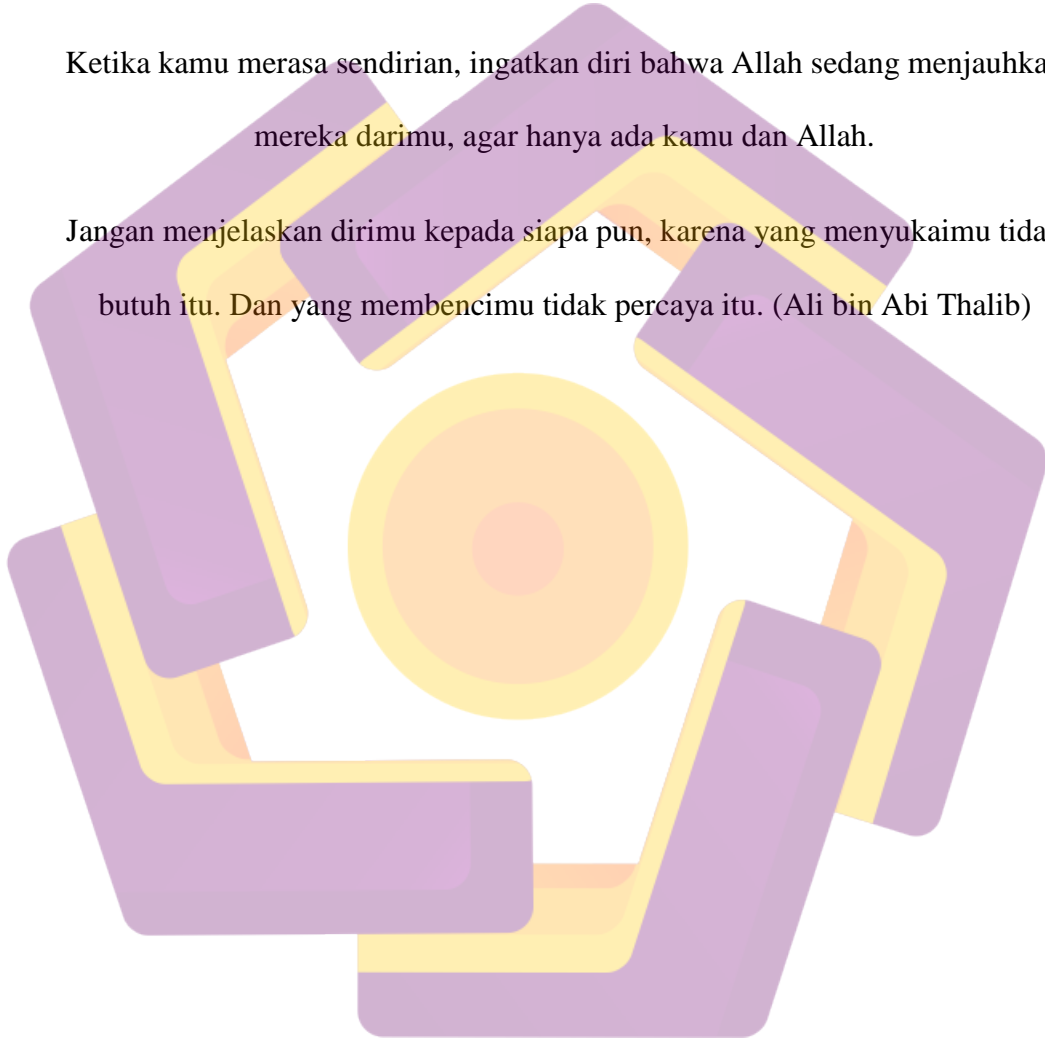
NIM. 17.11.1102

MOTTO

Bahagia adalah dia yang mampu menemukan hal sederhana dalam hidupnya dan tetap bersyukur.

Ketika kamu merasa sendirian, ingatkan diri bahwa Allah sedang menjauhkan mereka darimu, agar hanya ada kamu dan Allah.

Jangan menjelaskan dirimu kepada siapa pun, karena yang menyukaimu tidak butuh itu. Dan yang membencimu tidak percaya itu. (Ali bin Abi Thalib)



PERSEMBAHAN

Alhamdulillah segala puji syukur atas kehadiran Allah SWT yang telah memberikan limpahan rahmat dan karunia-Nya sehingga penelitian ini dapat dilakukan dan diselesaikan dengan sebaik-baiknya. Saya pribadi juga mengucapkan terimakasih untuk dukungan dan bantuan semua pihak yang membantu selesainya penelitian ini.

Skripsi ini saya persembahkan untuk saya pribadi dan kampus saya, Universitas AMIKOM Yogyakarta.

Untuk para pembaca, saya ucapkan terimakasih sebesar-besarnya. Semoga lewat perantara skripsi ini, akan menjadi ladang amal jariyah atas ilmu yang saya bagikan.

Untuk orangtua, keluarga, dan orang-orang yang saya sayangi saya ucapkan terimakasih atas motivasi, bimbingan dan doa-doanya.

Untuk bapak Asro Nasiri, yang sudah dengan sabar membimbing saya, dan banyak sekali kebaikan yang bapak berikan kepada saya. Saya ucapkan terimakasih sedalam-dalamnya dan semoga selalu dilimpahkan rezeki yang banyak dan berkah.

Dan untuk semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak dapat saya sebutkan satu persatu.

KATA PENGANTAR

Assalamualaikum Wr. Wb

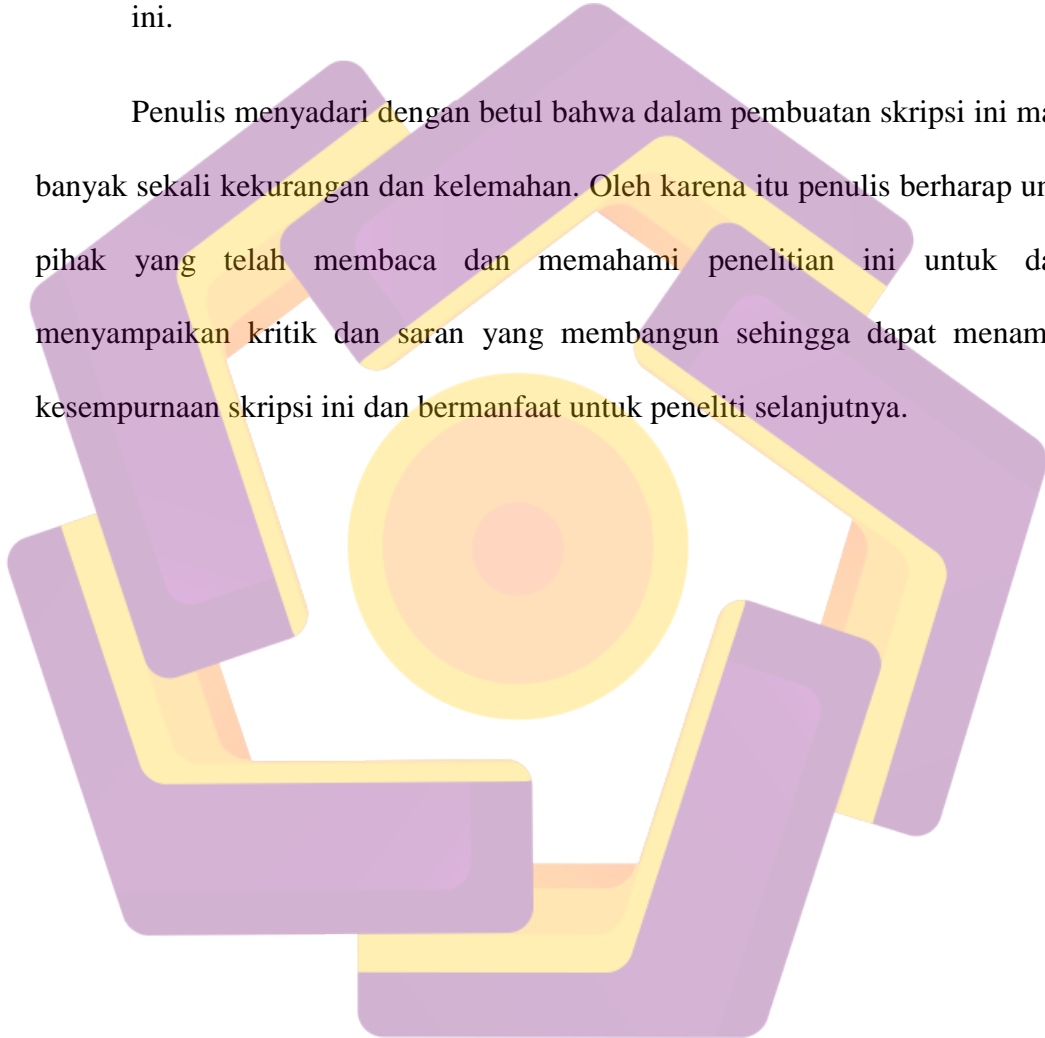
Puji syukur penulis haturkan atas kehadiran Allah SWT yang telah memberikan rahmat dan hidayah sehingga penulis dapat menyelesaikan skripsi ini dengan judul “Analisis Kinerja Algoritma C4.5 Untuk Prediksi Ketepatan Kelulusan Mahasiswa S1 Informatika Universitas Amikom Yogyakarta” dengan sebaik-baiknya. Tidak lupa sholawat serta salam penulis haturkan kepada junjungan nabi besar Muhammad SAW, yang mengantarkan manusia dari kegelapan ke zaman yang terang benderang.

Dengan selesainya skripsi ini, penulis mengucapkan banyak terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Sudarmawan, MT. selaku Dekan Fakultas Sains dan Teknologi, dan Ketua Program Studi S1 Informatika.
3. Drs. Asro Nasiri, M.kom. selaku pembimbing yang telah dengan sabar membimbing dan banyak memberikan kontribusi bagi penulis dalam pembuatan skripsi ini.
4. Para Dosen dan Staff Universitas AMIKOM Yogyakarta yang telah banyak memberi bantuan, pengalaman dan kontribusi lainnya selama penulis berkuliah di kampus ini hingga terselesaikannya skripsi ini.

5. Orang tua, dan keluarga penulis yang telah memberikan dukungan terbaiknya selama berkuliah.
6. Teman-teman IF-03 dan teman seperjuangan Chichen Noodle Lovers yang sudah menemani dan banyak membantu penulis selama berkuliah di kampus ini.

Penulis menyadari dengan betul bahwa dalam pembuatan skripsi ini masih banyak sekali kekurangan dan kelemahan. Oleh karena itu penulis berharap untuk pihak yang telah membaca dan memahami penelitian ini untuk dapat menyampaikan kritik dan saran yang membangun sehingga dapat menambah kesempurnaan skripsi ini dan bermanfaat untuk peneliti selanjutnya.



DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN KEASLIAN.....	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
INTISARI.....	xv
ABSTRACT.....	xvi
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	4
1.6 Metode Penelitian.....	4
1.6.1 Metode Pengumpulan Data.....	4
1.6.2 Metode Analisis	5
1.6.3 Metode Perancangan	5
1.6.4 Metode Pengujian.....	5
1.7 Sistematika Penulisan.....	5
BAB II.....	7
LANDASAN TEORI.....	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	10
2.2.1 Data Mining	10

2.2.2	Algoritma C4.5.....	13
2.2.3	Confusion Matrix	15
2.3	Peralatan dan Perancangan Web	16
2.3.1	Aplikasi Web.....	16
2.3.2	HTML (Hypertext Multi Language)	17
2.3.3	CSS (Cascading Style Sheet)	18
2.3.4	PHP (Hypertext Preprocessor File)	18
2.3.5	Bootstrap	18
2.3.6	Web Server.....	19
2.3.7	MYSQL.....	19
2.4	Konsep Pemodelan Data	20
2.4.1	UML (<i>Unifed Modeling Language</i>).....	20
2.4.1.1	Use Case Diagram	20
2.4.1.2	Activity Diagram.....	23
2.5	Konsep Dasar Basis Data	24
2.6	Flowchart.....	28
BAB III		30
METODE PENELITIAN.....		30
3.1	Alat dan Bahan Penelitian	30
3.1.1	Perangkat Keras	30
3.1.2	Perangkat Lunak.....	30
3.2	Alur Penelitian.....	31
3.2.1	Studi Literatur	32
3.2.2	Pengumpulan Dataset.....	33
3.2.3	Pembersihan Data.....	49
3.2.4	Pembuatan Aplikasi	51
3.2.5	Pengujian Aplikasi	83
3.2.6	Analisis Hasil	99
BAB IV		106
IMPLEMENTASI.....		106
4.1	Deskripsi Implementasi	106
4.2	Pembuatan Database.....	106

4.2.1	Tabel Data Training	107
4.2.2	Tabel Data Testing	108
4.3	Pembuatan Tampilan Sistem	109
4.3.1	Tampilan Halaman Dashboard.....	109
4.3.2	Tampilan Data Training	110
4.3.3	Tampilan Data Testing	111
4.3.4	Tampilan Hasil Prediksi	112
4.3.5	Tampilan Pohon Keputusan	113
4.3.6	Tampilan Prediksi Manual	114
4.3.7	Tampilan Akurasi	115
4.4	Implementasi Program	116
4.5	Implementasi Data Uji Coba	167
4.5.1	Uji ke 1	167
4.5.2	Uji ke 2	168
4.5.3	Uji ke 3	169
4.5.4	Uji ke 4	170
4.5.5	Uji ke 5	170
4.6	Pencocokan Hasil Uji Algoritma Program Dengan Tools	172
4.6.1	Hasil Uji Algoritma Dengan Program	172
4.6.2	Hasil Uji Algoritma Dengan Tools	173
BAB V	174
PENUTUP	174
5.1	Kesimpulan.....	174
5.2	Saran	174
DAFTAR PUSTAKA	176

DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Penelitian.....	8
Tabel 2.2 Tabel Simbol Use Case Diagram.....	21
Tabel 2.3 Tabel Simbol Activity Diagram.....	23
Tabel 2.4 Notasi Entity Relationship Diagram.....	25
Tabel 2.5 Tabel Symbol Flowchart.....	28
Tabel 3.1 Dataset.....	33
Tabel 3.2 Use Case Sistem Analisis Kinerja Algoritma C4.5.....	52
Tabel 3.3 Use Case Dashboard.....	53
Tabel 3.4 Use Case Data Training.....	56
Tabel 3.5 Use Case Data Testing.....	58
Tabel 3.6 Use Case Hasil Prediksi.....	59
Tabel 3.7 Use Case Pohon Keputusan.....	60
Tabel 3.8 Use Case Prediksi Manual.....	60
Tabel 3.9 Use Case Akurasi.....	62
Tabel 3.10 Tabel data training.....	69
Tabel 3.11 Tabel data testing.....	71
Tabel 3.12 Data Training Uji ke 1.....	83
Tabel 3.13 Data Training Uji ke 2.....	87
Tabel 3.14 Data Training Uji ke 3, 4 dan 5.....	91
Tabel 3.15 Data Testing Uji ke 1, 2 dan 3.....	97
Tabel 3.16 Data Testing Uji ke 4.....	98
Tabel 3.17 Data Testing Uji ke 5.....	98

DAFTAR GAMBAR

Gambar 3.1 Diagram Alur Penelitian.....	32
Gambar 3.2 Use Case Sistem Analisis Kinerja Algoritma C4.5.....	52
Gambar 3.3 Use Case Dashboard	53
Gambar 3.4 Use Case Data Training	55
Gambar 3.5 Use Case Data Testing	57
Gambar 3.6 Use Case Hasil Prediksi	58
Gambar 3.7 Use Case Pohon Keputusan	59
Gambar 3.8 Use Case Prediksi Manual.....	60
Gambar 3.9 Use Case Akurasi	61
Gambar 3.10 Activity Diagram Data Training	63
Gambar 3.11 Activity Diagram Data Testing	65
Gambar 3.12 Activity Diagram Hasil Prediksi	66
Gambar 3.13 Activity Diagram Pohon Keputusan	67
Gambar 3.14 Activity Diagram Prediksi Manual	68
Gambar 3.15 Activity Diagram Akurasi	69
Gambar 3.16 Perancangan Antar Muka Dashboard.....	73
Gambar 3.17 Perancangan Antar Muka Data Training.....	74
Gambar 3.18 Perancangan Antar Muka Data Testing	75
Gambar 3.19 Perancangan Antar Muka Tambah Data	76
Gambar 3.20 Perancangan Antar Muka Import Data.....	77
Gambar 3.21 Perancangan Antar Muka Hapus Semua Data	78
Gambar 3.22 Perancangan Antar Muka Filter Data.....	79
Gambar 3.23 Perancangan Antar Muka Hasil Prediksi	80
Gambar 3.24 Perancangan Antar Muka Detail	81
Gambar 3.25 Perancangan Antar Muka Pohon Keputusan.....	81
Gambar 3.26 Perancangan Antar Muka Prediksi Manual.....	82
Gambar 3.27 Perancangan Antar Muka Akurasi	83
Gambar 3.28 Data Keputusan Prediksi Kelulusan.....	99
Gambar 3.29 Perhitungan Node 1	100
Gambar 3.30 Pohon Keputusan Hasil Perhitungan Node 1	101
Gambar 3.31 Perhitungan Node 1.1	102
Gambar 3.32 Pohon Keputusan Hasil Perhitungan Node 1.1	103
Gambar 3.33 Perhitungan Node 1.1.2	103
Gambar 3.34 Pohon Keputusan Hasil Perhitungan Node 1.1.2	104
Gambar 3.35 Data Testing Prediksi	104
Gambar 3.36 Hasil Prediksi Dari Data Testing.....	105
Gambar 4.1 Query Pembuatan Database	106
Gambar 4.2 Tabel Data Training	108
Gambar 4.3 Tabel Data Testing	109

Gambar 4.4 Halaman Dashboard	110
Gambar 4.5 Halaman Data Training	111
Gambar 4.6 Halaman Data Testing	112
Gambar 4.7 Halaman Hasil Prediksi	113
Gambar 4.8 Halaman Pohon Keputusan	114
Gambar 4.9 Halaman Prediksi Manual	115
Gambar 4.10 Halaman Akurasi	116
Gambar 4.11 Tampilan Import Data Training	125
Gambar 4.12 Tampilan Import Data Training Berhasil	126
Gambar 4.13 Tampilan Import Data Testing Berhasil	134
Gambar 4.14 Tampilan Hasil Prediksi	145
Gambar 4.15 Tampilan Pohon Keputusan	151
Gambar 4.16 Tampilan Prediksi Manual	159
Gambar 4.17 Tampilan Akurasi	166
Gambar 4.18 Implementasi Uji ke 1	167
Gambar 4.19 Implementasi Uji ke 2	168
Gambar 4.20 Implementasi Uji ke 3	169
Gambar 4.21 Implementasi Uji ke 4	170
Gambar 4.22 Implementasi Uji ke 5	171
Gambar 4.23 Hasil Uji Algoritma dengan Program	172
Gambar 4.24 Hasil Uji Algoritma dengan Tools	173

INTISARI

Data mining merupakan suatu proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan disajikan dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar untuk pengambilan keputusan. Beberapa algoritma data mining yang sering digunakan antara lain naïve bayes, c4.5, neural network dan masih banyak lagi. Peneliti melihat dari berbagai penelitian lain tentang beberapa perbedaan kinerja algoritma klasifikasi data mining, algoritma c4.5 mempunyai akurasi yang lebih tinggi, jadi peneliti akan meneliti tentang bagaimana kinerja algoritma algoritma C4.5 untuk kasus dalam memprediksi ketepatan kelulusan Mahasiswa S1 Informatika Universitas Amikom Yogyakarta.

Dalam penelitian ini, peneliti menerapkan metode penelitian beberapa yang dipersiapkan misalnya metode pengumpulan data, analisis, perancangan, pengujian, dan menganalisis data hasil dari pengujian yang telah dilakukan. Penelitian ini mengambil sampel data alumni di Program Studi Informatika Universitas Amikom Yogyakarta. Variabel yang digunakan dalam penelitian ini antara lain: jenis kelamin, asal sekolah, konsentrasi, ipk dan lama masa studi.

Dalam hasil akhir, peneliti menyimpulkan bahwa dalam percobaan yang telah dilakukan sebanyak 5 kali uji coba dengan rincian data yang berbeda, menghasilkan kesimpulan Kinerja dari algoritma c4.5 ketika diterapkan dalam program menghasilkan akurasi baik, dan akurasinya sama dengan menggunakan bantuan tools rapidminer, Berdasarkan dari uji data, data terakhir menunjukkan akurasi terbaik dalam percobaan yaitu menggunakan percobaan dengan spesifikasi data training sebanyak 168 data, dan data testing sebanyak 6 data dan menghasilkan Akurasi sebesar 83,33%.

Kata Kunci : Data Mining, Performa, Algoritma C4.5.

ABSTRACT

Data mining is a process of finding relationships in data that are unknown to the user and presented in an understandable way so that these relationships can be the basis for decision making. Some of the data mining algorithms that are often used include naïve Bayes, C4.5, neural networks and many more. Researchers see from various other studies about some differences in the performance of the data mining classification algorithms, the c4.5 algorithm has higher accuracy, so researchers will examine how the performance of the C4.5 algorithm for cases in predicting the accuracy of graduation of Undergraduate students of Informatics at Amikom University Yogyakarta.

In this study, the researcher applied several prepared research methods, for example the method of data collection, analysis, design, testing, and analysis of data resulting from tests that had been carried out. This study took a sample of alumni data at the Informatics Study Program at Amikom University Yogyakarta. The variables used in this study include: gender, school origin, concentration, ipk and length of study period.

In the final results, the researcher concludes that in an experiment that has been carried out 5 times with different details of data, resulting in a conclusion that the performance of the C4.5 algorithm when applied to the program produces good accuracy, and the accuracy is the same as using the rapidminer tools. Test data, the latest data shows the best accuracy in the experiment, namely using experiments with 168 training data specifications, and testing data as many as 6 data and produces an accuracy of 83.33%.

Keywords: *Data Mining, Performance, C4.5 Algorithm.*