

**IMPLEMENTASI ALGORITME C4.5 UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA**

SKRIPSI

**Diajukan Oleh :
Any Sawheri Gading
2011102441219**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
JULI 2024**

**IMPLEMENTASI ALGORITME C4.5 UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA**

SKRIPSI

Diajukan Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur

Diajukan oleh:

Any Sawheri Gading

2011102441219



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
JULI 2024**

LEMBAR PERSETUJUAN
IMPLEMENTASI ALGORITMA C4.5 UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA

SKRIPSI

Diajukan oleh:

Any Sawheri Gading

2011102441219

Disetujui untuk Diujikan

Pada Tanggal

30 Juni

Pembimbing



Rofilde Hasudungan, S.Kom, M.Sc

NIDN. 1107048601

Mengetahui,

Koordinator Tugas Akhir



Abdul Rahim, S.Kom., M.Cs.

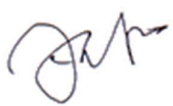

NIDN. 0009047901

LEMBAR PENGESAHAN
IMPLEMENTASI ALGORITME C4.5 UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA

SKRIPSI

Diajukan oleh:
Any Sawheri Gading
20111102441219

Diseminarkan dan Diujikan
Pada Tanggal 16 Juli 2024

Penguji I	Penguji II
	
<u>Wawan Joko Pranoto, S.Kom., M.Ti</u> NIDN. 0009047901	<u>Rofilde Hasudungan, S.Kom., M.Sc</u> NIDN. 1107048601

Mengetahui,
Ketua
Program Studi Teknik Informatika


Arbansyah, S.Kom., M.TI
NIDN. 1118019203

PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama : Any Sawheri Gading

NIM : 2011102441219

Program Studi : SI Teknik Informatika

Judul Penelitian : IMPLEMENTASI ALGORITME C4.5 UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA

menyatakan bahwa **skripsi** yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung risiko atau sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam **skripsi** saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini.

Samarinda, 30 Juni 2024

Yang membuat pernyataan



Any Sawheri Gading
NIM : 2011102441219

ABSTRAK

Menyelesaikan studi tepat waktu adalah suatu hal yang sangat diharapkan karena kelulusan merupakan tujuan utama oleh setiap mahasiswa. Selain itu tiap perguruan tinggi pasti memiliki keinginan agar mahasiswanya dapat lulus tepat waktu, karena hal ini dapat mendukung upaya dalam meningkatkan reputasi dan akreditasi. Salah satu cara yang dapat dilakukan dalam mengontrol tingkat kelulusan adalah dengan memprediksi kelulusan mahasiswa. Dalam upaya memprediksi tingkat kelulusan mahasiswa dapat menggunakan suatu model algoritme, salah satunya adalah algoritme C4.5. Algoritme C4.5 merupakan algoritme yang digunakan untuk melakukan proses klasifikasi data dengan menggunakan struktur pohon atau struktur berhierarki. Penelitian ini menggunakan data akademik 2017 – 2019 dalam pembuatan model algoritme C.45, kemudian menerapkan model tersebut ke dalam data akademik 2020 untuk memprediksi kelulusan. Evaluasi model algoritme C4.5 menunjukkan hasil akurasi sebesar 70,31% , dimana model ini memprediksi 118 mahasiswa yang lulus lambat dan 47 mahasiswa yang lulus tepat waktu. Karena nilai akurasi yang dapat dianggap cukup, maka perlu ditingkatkan dengan menambah lebih banyak data, menggunakan algoritme lain, atau menggunakan *cross-validation* untuk menguji model algoritme.

Kata Kunci : Prediksi Kelulusan Mahasiswa, *Data Mining*, *Decission Tree*, Algoritme C4.5.

ABSTRACT

Completing studies on time is highly expected because graduation is the main goal by every student. In addition, each university must have a desire for its students to graduate on time, because this can support efforts to improve reputation and accreditation. One way that can be done in controlling the graduation rate is to predict student graduation. In an effort to predict the graduation rate, students can use an algorithm model, one of which is the C4.5 algorithm. The C4.5 algorithm is an algorithm used to carry out the data classification process using a tree structure or hierarchical structure. This study uses academic data from 2017 – 2019 in the creation of a C.45 algorithm model, then applies the model to 2020 academic data to predict graduation. The evaluation of the C4.5 algorithm model showed an accuracy of 70,31%, where this model predicted 118 students who graduated late and 47 students who graduated on time. Since the accuracy value can be considered sufficient, it needs to be improved by adding more data, using other algorithms, or using cross-validation to test the algorithm model.

Keywords: *Student Graduation Prediction, Data Mining, Decission Tree, Algorithm C4.5.*

PRAKATA

Alhamdulillahirabbil'alamiin, segala puji dan syukur selalu dipanjatkan ke hadirat Allah SWT yang senantiasa melimpahkan rahmat, taufik, serta hidayah-Nya, sehingga penyusunan skripsi yang berjudul "IMPLEMENTASI ALGORITME C4.5 UNTUK MEMPREDIKSI KELULUSAN MAHASISWA" ini dapat terselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar S.Kom pada Prodi Teknik Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur Prodi Teknik Informatika. Penulis menyadari dalam penyusunan skripsi ini tidak luput dari bantuan dan dukungan berbagai pihak. Oleh karena itu pada kesempatan ini, penulis mengucapkan terima kasih yang sebesar besarnya kepada:

1. Alm. Sri Yunanik & Edyl selaku orang tua dari penulis serta saudari Asriyani Piqrah Gading selaku adik dari penulis yang sudah memotivasi penulis mulai dari awal proses perkuliahan di mulai, hingga penyusunan skripsi. Penulis mempersembahkan skripsi ini kepada Alm. Ibu tercinta sebagai tanda terima kasih dan syukur atas semua kepercayaan yang diberikan karena dianggap mampu mencapai gelar sarjana yang diamanatkan.
2. Seluruh keluarga besar penulis yang tidak dapat disebutkan satu persatu, baik yang dekat maupun yang jauh. Terima kasih untuk segala doa yang selalu dilambungkan untuk kesuksesan penulis.
3. Bapak Rofilde Hasudungan, S.Kom., M.Sc selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan dorongan yang luar biasa selama penyusunan skripsi ini berlangsung. Bimbingan dan masukan yang diberikan telah sangat amat berharga bagi penulis dalam mengembangkan ide dan mewujudkan tulisan ini.
4. Bapak Wawan Joko Pranoto, S.Kom., M.Ti selaku Dosen Penguji yang telah bersedia meluangkan waktu untuk membaca dan memberikan masukan yang konstruktif terhadap skripsi ini.
5. Saudara Fachrul Rasyid yang sudah bersedia selalu mendampingi setiap proses penulisan skripsi ini dari awal hingga selesai, serta senantiasa memberikan semangat dan motivasi untuk penulis agar terus optimis akan apa yang dijalani.
6. Semua sahabat seperjuangan penulis, terkhusus kepada saudari Aulia Khofifah Syamsuri, Khusnul Khotimah, Adia Lestari, serta Siti Patimah yang selalu ada dan memberikan doa, dukungan, dan bantuan di setiap langkah kegiatan perkuliahan penulis.
7. Segenap Staf Prodi Teknik Informatika UMKT dan semua pihak terkait yang telah membantu dalam berbagai hal, sehingga skripsi ini dapat terselesaikan dengan baik.
8. Dan terakhir, penulis ingin berterima kasih kepada diri penulis sendiri sebagai tanda apresiasi karena telah mampu bertahan untuk percaya saat menghadapi hal-hal yang sulit di saat tak ada siapapun yang mendampingi. *Fa inna ma'al-'usri yusroo, inna ma'al-'usri yusroo.*

Meskipun telah berusaha menyelesaikan skripsi ini sebaik mungkin, tentu saja penulis menyadari bahwa skripsi ini masih ada kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari para pembaca demi menyempurnakan segala kekurangan dalam penyusunan skripsi ini. Semoga skripsi ini dapat bermanfaat bagi para pembaca yang hendak mengembangkan pengetahuan di bidang Data Mining

Akhir kata, semoga Allah SWT senantiasa memberikan rahmat dan hidayah-Nya kepada kita semua. Amin.

Samarinda, 25 Juni 2024

Penulis,

A handwritten signature in black ink, appearing to read 'Any Sawheri Gading', written in a cursive style.

Any Sawheri Gading

DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN KEASLIAN PENELITIAN	iii
ABSTRAK	iv
ABSTRACT	vi
PRAKATA	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
Daftar TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
BAB II	5
METODE PENELITIAN	5
2.1 Objek Penelitian	5
2.2 Alat dan Bahan	5
2.3 Prosedur Penelitian	6
2.3.1 Pengumpulan Data	7
2.3.2 Pra-proses data	7

2.3.3	Data training dan data testing.....	8
2.3.4	Algoritme C4.5.....	8
2.3.5	Evaluasi Model	10
2.3.6	Prediksi kelulusan	10
BAB III		11
HASIL ANALISIS DAN PEMBAHASAN.....		11
3.1	Pengumpulan Data.....	11
3.2	Pra-proses data.....	13
3.2.1	Penggabungan data	13
3.2.2	Cleaning data	14
3.3	Data Training dan Data Testing	15
3.4	Algoritme C4.5	16
3.5	Evaluasi Model	20
3.6	Prediksi Kelulusan	21
BAB IV		23
4.1	Kesimpulan	23
4.2	Saran	23
DAFTAR RUJUKAN		25
LAMPIRAN.....		27
RIWAYAT HIDUP		44

DAFTAR GAMBAR

Gambar	Halaman
2.1 Alur Tahap Penelitian.....	6

DAFTAR TABEL

Tabel	Halaman
3.1 Data Yudisium Prodi Teknik Informatika Semester Genap TA 2020/2021	11
3.2 Data akademik 2017 Semester 1	12
3.3 Data akademik 2018 Semester 1	12
3.4 Data akademik 2019 Semester 1	12
3.5 Data akademik 2020 Semester 1	13
3.6 Data gabungan angkatan 2017-2019.....	13
3.7 Data angkatan 2020.....	14
3.8 Data final 2017-2019	14
3.9 Data final 2020.....	15
3.10 Data training.....	15
3.11 Data testing.....	15
3.12 Data Sampel.....	16
3.13 Frekuensi IPS 1	17
3.14 <i>Entropy</i> dan <i>IG</i> tiap atribut	18
3.15 Hasil prediksi model	20
3.16 Hasil akurasi.....	21
3.17 Prediksi kelulusan	21

DAFTAR LAMPIRAN

Lampiran	Halaman
L1. Data Yudisium Prodi TI Genap TA 2020/2021	27
L2. Data Yudisium Genap TA 2021/2022	28
L 3. Data Yudisium Genap TA 2022/2023	29
L 4. Data Yudisium Ganjil TA 2022/2023.....	30
L 5. Data final angkatan 2017-2019.....	31
L 6. Data final angkatan 2020.....	33
L7. Proses input data.....	35
L8. Pembagian data training dan data testing	35
L 9. Data training	36
L10. Data testing.....	38
L11. Pembuatan model algoritme C4.5	38
L12. Mengevaluasi model menggunakan data testing.....	39
L13. Hasil prediksi model algoritme C4.5.....	39
L14. Membuat rule Decission Tree.....	40
L15. Impor data mahasiswa angkatan 2020.....	41
L16. Menerapkan model prediksi ke dalam data mahasiswa angkatan 2020	41
L17. Hasil prediksi kelulusan mahasiswa angkatan 2020	42
L18. Surat Ijin Penelitian	44
L19. Kartu kendali bimbingan	45