

**ANALISIS DAN IMPLEMENTASI PERBANDINGAN ALGORITMA C4.5
DAN NAIVE BAYES UNTUK PREDIKSI KELULUSAN MAHASISWA
(Studi Kasus: Universitas AMIKOM Yogyakarta)**

SKRIPSI



Disusun oleh

Elfrina Aryani

15.11.8547

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

**ANALISIS DAN IMPLEMENTASI PERBANDINGAN ALGORITMA C4.5
DAN NAIVE BAYES UNTUK PREDIKSI KELULUSAN MAHASISWA
(Studi Kasus: Universitas AMIKOM Yogyakarta)**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



**Disusun oleh
Elfrina Aryani
15.11.8547**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

PERSETUJUAN

SKRIPSI

ANALISIS DAN IMPLEMENTASI PERBANDINGAN ALGORITMA C4.5 DAN NAIVE BAYES UNTUK PREDIKSI KELULUSAN

MAHASISWA

(Studi Kasus: Universitas AMIKOM Yogyakarta)

yang dipersiapkan dan disusun oleh

Elfrina Aryani

15.11.8547

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 16 November 2018

Dosen Pembimbing,

Arif Dwi Laksito, M.Kom.

NIK. 190302150

PENGESAHAN

SKRIPSI

ANALISIS DAN IMPLEMENTASI PERBANDINGAN ALGORITMA C4.5 DAN NAIVE BAYES UNTUK PREDIKSI KELULUSAN MAHASISWA

(Studi Kasus: Universitas AMIKOM Yogyakarta)

yang dipersiapkan dan disusun oleh

Elfrina Aryani

15.11.8547

telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Juli 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Mardhiya Hayati, S.T., M.Kom

NIK. 190302108

Arif Dwi Laksito, M.Kom

NIK. 190302150

Norhikmah, M.Kom.

NIK. 190302245

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 3 Agustus 2022

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom, M.Kom.

NIK. 190302096

PERNYATAAN

PERNYATAAN

Saya yang bertanda tangan 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 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 tanggung jawab saya pribadi

Yogyakarta, 21 Juli 2022



Elfrina Aryani

15.11.8547

MOTTO

Hidup bagaikan pesawat kertas
Terbang dan pergi membawa impian
Sekuat tenaga dengan hembusan angin

Terus melaju terbang
Jangan bandingkan jarak terbangnya
Tapi bagaimana dan apa yang dilalui
Karena itulah satu hal yg penting
S'lalu sesuai kata hati

(Pesawat Kertas-JKT48)



PERSEMBAHAN

Dengan mengucapkan syukur Alhamdulillah atas nikmat dan karunia ALLAH SWT, skripsi ini saya persembahkan kepada :

1. Allah Subhana wata'ala yang telah melimpahkan segala rahmat dan karunia dalam bentuk apapun, sehingga dilancarkan dan diberikan kemudahan dalam segala urusan yang penulis hadapi, terutama dalam proses penyampaian naskah skripsi.
2. Kedua orangtua yang dengan ikhlas tanpa mengharapkan imbalan apapun selalu memberikan dukungan motivasi dan doa yang tiada henti untuk kesuksesan saya.
3. Bapak Arif Dwi Laksito selaku pembimbing saya yang sangat baik dan tulus ikhlas meluangkan waktunya untuk memberikan bimbingan dan pelajaran yang tiada ternilai harganya.
4. Taufik Firtriyadi, Arif Lutfi Kurniawan yang membantu, membagikan ilmu dalam proses penyusunan skripsi.
5. Rizal Ghozali, Cornelia Sundari, Mega Ayu Anggraeni, Alfian Rahman Wastita, Asa Hirasna, Rinda Mutiara yang sudah memberikan motivasi dan semangat dalam penyusunan skripsi.
6. Sahabat dan teman-temanku seperjuangan 15.S1-IF-01.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya dan shalawat serta salam juga tidak lupa penulis panjatkan kepada junjungan kita Nabi Muhammad SAW yang telah memberikan teladan mulia dalam menuntun ummatnya sehingga penulis dapat menyelesaikan skripsi ini. Skripsi yang berjudul “Analisis dan Implementasi Perbandingan Algoritma C.45 dan Naive Bayes untuk Prediksi Kelulusan Mahasiswa Studi Kasus : Universitas AMIKOM Yogyakarta” ini disusun sebagai salah satu syarat utama untuk menyelesaikan program sarjana pada Universitas AMIKOM Yogyakarta. Penyelesaian skripsi ini juga tidak lepas dari bantuan berbagai pihak, karena itu pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada:

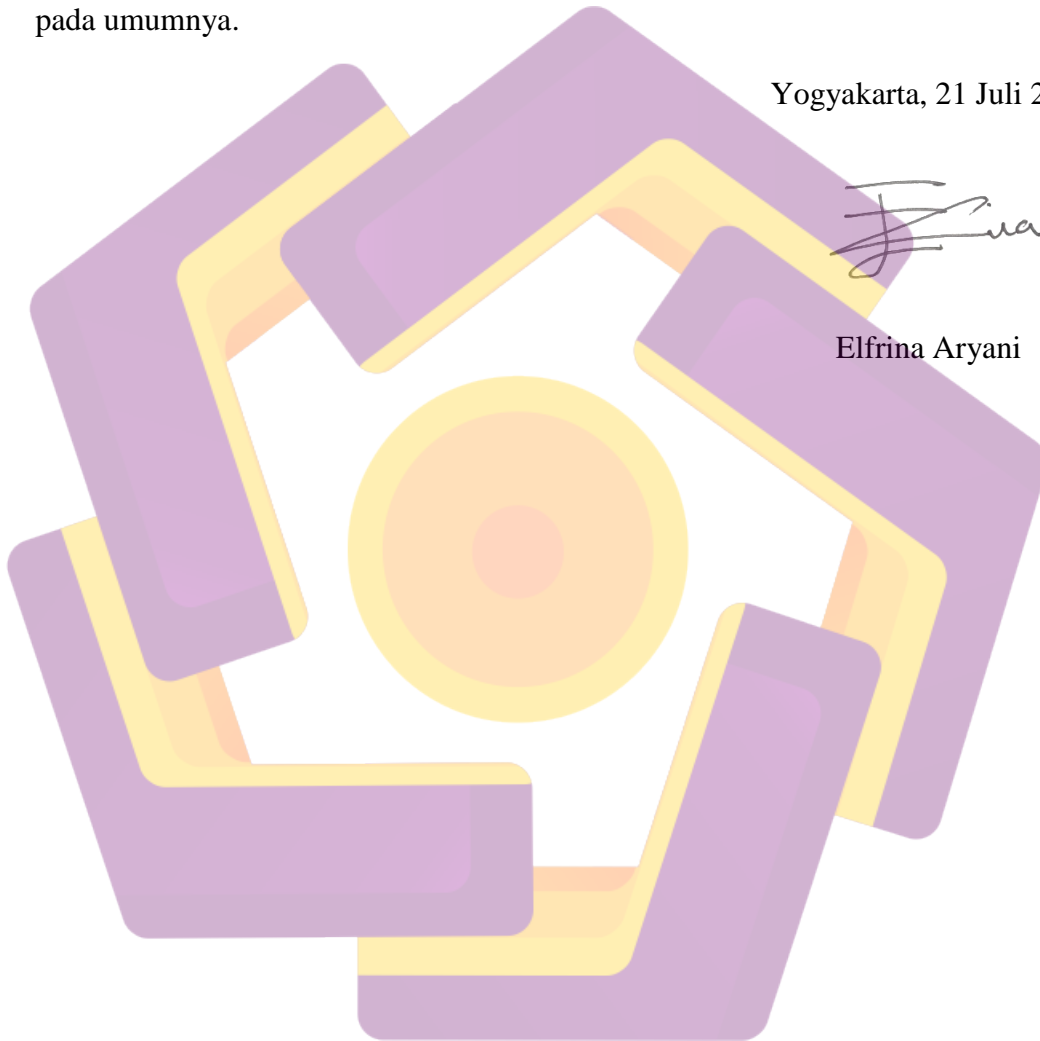
1. Prof. Dr. M. Suyanto, MM selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Arif Dwi Laksito, M.Kom selaku dosen pembimbing yang telah banyak memberikan arahan bagi penulis agar penulis menyelesaikan skripsi ini dengan baik.
3. Kedua orang tua, keluarga besar dan saudara-saudara yang telah memberikan dukungan, doa, kasih sayang kepada penulis.
4. Seluruh dosen dan staff Universitas AMIKOM Yogyakarta yang telah membantu dan membimbing selama proses perkuliahan.
5. Teman-teman kelas 15-S1IF-01 dan sahabat kampus terima kasih banyak semuanya dan semua pihak yang tidak dapat saya sebutkan satu persatu.

Penulis menyadari skripsi ini masih ada kekurangan, maka dari itu kritik dan saran yang membangun serta teguran dari semua pihak, penulis menerima dengan lapang dada untuk kesempurnaan karya selanjutnya. Semoga skripsi yang sederhana ini bermanfaat. Khususnya bagi penulis dan pembaca yang budiman pada umumnya.

Yogyakarta, 21 Juli 2022



Elfrina Aryani



DAFTAR ISI

JUDUL	i
PERSETUJUAN.....	ii
PENGESAHAN	iii
PERNYATAAN.....	iv
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	3
1.6 Metode Penelitian.....	3
1.6.1 Metode Pengumpulan Data	4
1.6.2 Preprocessing Data	4
1.6.3 Metode Implementasi dan Pengujian	4

1.6.4	Analisis.....	5
1.7	Sistematika Penulisan	5
BAB II	LANDASAN TEORI	7
2.1	Tinjauan Pustaka	7
2.2	Data Mining.....	15
2.3	Algoritma C 4.5.....	18
2.4	Algoritma Naive Bayes	20
2.5	Confussion Matrik.....	21
2.6	Rapidminer	22
BAB III	METODE PENELITIAN	24
3.1	Lokasi Penelitian	24
3.2	Bahan dan Data	24
3.3	Alat Penelitian	24
3.3.1	Perangkat Keras (Hardware).....	24
3.3.2	Perangkat Lunak (Software)	25
3.4	Metode Pengumpulan Data	25
3.4.1	Metode Studi Dokumentasi.....	25
3.4.2	Kepustakaan	26
3.5	Tahapan Data Mining.....	26
3.5.1	Pembersihan Data.....	26
3.5.2	Seleksi Data (data selection)	26
3.5.3	Transformasi Data.....	26
3.5.4	Proses Data Mining.....	27
3.6	Tahapan Penelitian	27
BAB IV	29

PEMBAHASAN	29
4.1 Hasil Pengumpulan Data	29
4.2 Tahapan Preprocessing	30
4.2.1 Cleaning Data	30
4.2.2 Seleksi	31
4.2.3 Transformasi	32
4.2.4 Data Akhir	33
4.3 Pengolahan Data	35
4.3.1 Metode Naive Bayes	39
4.3.2 Metode C4.5	41
4.4 Uji Hasil Hitung dengan Rapidminer	52
4.5 Pembahasan	54
4.5.1 Distribusi Data	54
4.5.2 Split Data	58
4.5.3 Hasil Metode Naive Bayes	59
4.5.4 Hasil Metode C4.5	61
4.5.5 Hasil Akhir Perbandingan	64
BAB V KESIMPULAN	65
5.1 Kesimpulan	65
5.2 Saran	65
DAFTAR PUSTAKA	67

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian.....	10
Tabel 4.1 Data Mahasiswa.....	29
Tabel 4. 2 Proses Data Mahasiswa.....	30
Tabel 4.3 Tabel Proses	31
Tabel 4.4 Aturan Transformasi IPK.....	32
Tabel 4.5 Aturan Transformasi SKS.....	32
Tabel 4.6 Aturan Transformasi Keterangan Lulus	33
Tabel 4.7 Hasil Akhir data	33
Tabel 4.8 Hasil Dataset Sample Perhitungan.....	35
Tabel 4.9 Hasil perhitungan Gain dan Entropy pertama.....	42
Tabel 4.10 Perhitungan Dataset Kedua.....	44
Tabel 4.11 Hasil perhitungan Gain dan Entropy Kedua	45
Tabel 4.12 Dataset untuk perhitungan ketiga.....	46
Tabel 4.13 Hasil Perhitungan Gain dan Entropy Ketiga.....	48
Tabel 4.14 Dataset baru hasil perhitungan ketiga	50
Tabel 4.15 Hasil perhitungan Informasi Gain dan Entropy ke empat.....	51
Tabel 4.16 Hasil Perbandingan Metode	64

DAFTAR GAMBAR

Gambar 2. 1 Pohon Keputusan Algoritma C.45	18
Gambar 3. 1 Tahapan Penelitian.....	27
Gambar 4.1 Pohon Keputusan Perhitungan Pertama.....	43
Gambar 4.2 Pohon keputusan perhitungan kedua.....	46
Gambar 4.3 Pohon keputusan perhitungan ketiga.....	49
Gambar 4.4 Pohon keputusan perhitungan keempat.....	51
Gambar 4.5 Skema Model uji metode Naive Bayes	52
Gambar 4. 6 Hasil uji perhitungan Naive Bayes.....	53
Gambar 4.7 Skema Model uji metode C.45	53
Gambar 4.8 Hasil pohon metode C4.5	54
Gambar 4.9 Hasil uji metode C4.5.....	54
Gambar 4.10 Distribusi Variabel jenis Kelamin	55
Gambar 4.11 Distribusi Variabel Asal Sekolah	56
Gambar 4.12 Distribusi Variabel IPK.....	56
Gambar 4.13 Distribusi Variabel SKS	57
Gambar 4.14 Distribusi Variabel Label Kelulusan	58
Gambar 4.15 Split data.....	59
Gambar 4.16 Skema Pengujian Metode Naive Bayes	60
Gambar 4. 17 Hasil Akurasi metode Naive Bayes.....	60
Gambar 4. 18 Skema Pengujian Metode C4.5	61
Gambar 4.19 Pohon Keputusan.....	62
Gambar 4.20 Grafik Pohon Keputusan	62
Gambar 4.21 Hasil Akurasi metode C.45	63

INTISARI

Universitas Amikom merupakan sebuah instansi yang bergerak dibidang akademik yang setiap tahun mahasiswanya semakin bertambah, namun tidak semua mahasiswa dapat lulus tepat waktu sesuai dengan masa studi yang ditempuh sehingga mengakibatkan penumpukan data mahasiswa yang semakin banyak. Karena permasalahan tersebut maka penulis membuat sebuah penelitian untuk menggali data yang dimiliki oleh prodi informatika Universitas Amikom Yogyakarta. Data yang akan dimanfaatkan dalam penelitian ini adalah data mahasiswa alumni berupa NIM, jenis kelamin, jumlah sks, indeks prestasi semester 1 sampai semester 6, asal sekolah, dan tahun kelulusan. Kemudian data tersebut akan digunakan sebagai data *training* dan data *testing* untuk memprediksi masa studi mahasiswa. Dari data masalah dan historis yang sudah terjadi dibutuhkan suatu teknik dan metode klasifikasi untuk menganalisa masalah tersebut. Algoritma C4.5 dan Naive Bayes adalah salah satu metode klasifikasi yang dapat di uji untuk memproses dan memprediksi kelulusan mahasiswa.

Dari penelitian yang dilakukan, terdapat tahapan preporcessing dari 200 data awal menghasilkan data yang siap digunakan sejumlah 175, kemudian berdasarkan tingkat akurasi setiap metode didapatkan hasil pengujian bahwa metode Naive Bayes menghasilkan tingkat akurasi 65,38% sedangkan metode C4.5 menghasilkan akurasi yang lebih tinggi dengan nilai 75% dan hasil akhir dari penelitian yang digunakan pada metode terbaik dengan menghasilkan nilai akurasi terbaik adalah metode C.45 dengan nilai akurasi yang cukup tinggi sebesar 75%.

Kata kunci: C4.5, Prediksi, Naive Bayes, Nilai Akurasi

ABSTRACT

Amikom University is an institution that is engaged in the academic field which every year the number of students is increasing, but not all students can graduate on time according to the study period taken, resulting in the accumulation of more student data. Because of these problems, the author made a study to explore the data owned by the informatics study program, Amikom University, Yogyakarta. The data that will be used in this study are student alumni data in the form of NIM, gender, number of credits, achievement index for semester 1 to semester 6, school origin, and year of graduation. Then the data will be used as training data and testing data to predict the student's study period. From the problem and historical data that has occurred, a technique and classification method is needed to analyze the problem. The C4.5 and Naive Bayes algorithms are one of the classification methods that can be tested to process and predict student graduation.

From the research and discussion conducted by the author, it can be concluded that the comparative testing process is carried out by preprocessing the initial 200 data, producing 175 ready-to-use data. accuracy is 65.38% while the C4.5 method produces a higher accuracy with a value of 75% and the final result of the research used in the best method by producing the best accuracy value is the C.45 method with a fairly high accuracy value of 75%.

Keywords: *C4.5, Prediction, Naive Bayes, Accuracy Value*