

**IMPLEMENTASI METODE *DECISION TREE* UNTUK
PREDIKSI MAHASISWA BERPOTENSI NON-AKTIF
MENGUNAKAN ALGORITMA C4.5
(Studi Kasus: Program Studi Sistem Informasi
Universitas Cenderawasih)**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Komputer Program Studi Sistem Informasi**



**OLEH
DINI ABSHARI
NIM. 2020051074017**

**JURUSAN SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS CENDERAWASIH
JAYAPURA
2024**

ABSTRAK

Abshari, Dini. 2024. **Implementasi Metode *Decision Tree* Untuk Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Algoritma C4.5 (Studi Kasus: Program Studi Sistem Informasi Universitas Cenderawasih)**. Skripsi Program Studi Sistem Informasi Jurusan Sistem Informasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Cenderawasih Jayapura.

Universitas Cenderawasih merupakan salah satu perguruan tinggi terbesar di Papua yang memiliki 10 fakultas dan 83 program studi. Sistem Informasi adalah salah satu program studi di Universitas Cenderawasih yang memiliki kurang lebih 460 mahasiswa dan sebanyak 153 mahasiswa berstatus non-aktif. Mahasiswa non-aktif yaitu mahasiswa yang tidak mengikuti proses perkuliahan dan tidak membayar biaya administrasi kuliah dalam kurun waktu dua semester atau lebih. Dampak dari mahasiswa yang non-aktif akan berpengaruh terhadap mutu perguruan tinggi sehingga diperlukan tindakan untuk memprediksi faktor-faktor penyebab mahasiswa berpotensi non-aktif. Salah satu teknik yang digunakan untuk prediksi adalah teknik klasifikasi. Peneliti sebelumnya telah menunjukkan bahwa algoritma C4.5 dapat digunakan untuk proses pengklasifikasian data mahasiswa. Penelitian ini bertujuan untuk mengetahui pengimplementasian metode *decision tree* algoritma C4.5 untuk prediksi mahasiswa berpotensi non-aktif. Dataset yang digunakan pada penelitian ini yaitu data mahasiswa angkatan 2013-2021 yang berjumlah 330 data. Analisis data pada penelitian ini mengikuti tahapan-tahapan KDD (*Knowledge Discovery in Database*). Metode pengembangan sistem yang digunakan adalah metode *waterfall* dan bahasa pemrograman yang digunakan adalah PHP dengan framework codeigniter. Berdasarkan hasil pengujian, diperoleh bahwa sistem ini dapat berjalan dengan baik, dan dapat memprediksi mahasiswa apakah berstatus aktif atau non-aktif berdasarkan atribut-atribut yang diinput, serta memperoleh nilai akurasi terbaik sebesar 78,79% dengan persentase data training sebesar 60% dan data testing sebesar 40%.

Kata Kunci: Status Mahasiswa, Non-Aktif, Prediksi, *Decision Tree*, Algoritma C4.5, PHP, Codeigniter.

ABSTRACT

Abshari, Dini. 2024. ***Implementation of the Decision Tree Method for Predicting Potentially Non-Active Students Using the C4.5 Algorithm (Case Study: Cenderawasih University Information Systems Study Program)***. Thesis Information Systems Study Program, Department of Information Systems, Faculty of Mathematics and Natural Sciences, Cenderawasih University Jayapura.

Cenderawasih University is one of the largest universities in Papua which has 10 faculties and 83 study programs. Information Systems is one of the study programs at Cenderawasih University which has approximately 460 students and as many as 153 students are inactive. Inactive students are students who do not take part in the lecture process and do not pay tuition fees for a period of two semesters or more. The impact of inactive students will affect the quality of higher education so that action is needed to predict the factors that cause potential students to be inactive. One of the techniques used for prediction is classification techniques. Previous researchers have shown that the C4.5 algorithm can be used for the process of classifying student data. This study aims to determine the implementation of the C4.5 algorithm decision tree method for predicting potentially inactive students. The dataset used in this study is student data from 2013-2021 which amounts to 330 data. Data analysis in this study follows the stages of KDD (Knowledge Discovery in Database). The system development method used is the waterfall method and the programming language used is PHP with the codeigniter framework. Based on the test results, it is found that this system can run well, and can predict students whether they are active or inactive based on the attributes inputted, and get the best accuracy value of 78.79% with a percentage of training data of 60% and testing data of 40%.

Keywords: Student Status, Non-Active, Prediction, Decision Tree, C4.5 Algorithm, PHP, Codeigniter.

LEMBAR PERSETUJUAN

Skripsi dengan judul: **IMPLEMENTASI METODE *DECISION TREE* UNTUK PREDIKSI MAHASISWA BERPOTENSI NON-AKTIF MENGGUNAKAN ALGORITMA C4.5 (Studi Kasus: Program Studi Sistem Informasi Universitas Cenderawasih)** oleh Dini Abshari telah diperiksa dan disetujui untuk diuji.

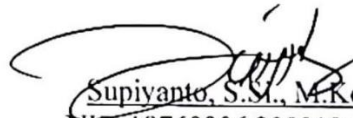
Jayapura, 20 Juni 2024

Pembimbing I



Agung Dwi Saputro, S.Kom., M.Kom
NIP. 19900323 201903 1 002

Pembimbing II



Supiyanto, S.Si., M.Kom.
NIP. 19760906 200212 1 003

LEMBAR PENGESAHAN

Skripsi dengan judul: **IMPLEMENTASI METODE *DECISION TREE* UNTUK PREDIKSI MAHASISWA BERPOTENSI NON-AKTIF MENGGUNAKAN ALGORITMA C4.5 (Studi Kasus: Program Studi Sistem Informasi Universitas Cenderawasih)** oleh Dini Abshari telah dipertahankan di depan Dewan Penguji pada hari Kamis tanggal 20 Juni 2024.

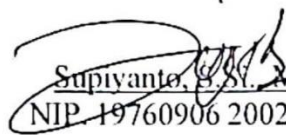
Dewan Penguji:

Nama	Jabatan	Tanda Tangan
1. <u>Remuz M.B. Kmurawak, ST., MT</u> NIP. 19840908 200812 1 003	Ketua	 (.....)
2. <u>Dr. Mingsep Rante Sampebua, ST., MT</u> NIP. 19750922 201404 1 001	Sekretaris	 (.....)
3. <u>Muhammad Asghar Nazal, S.Si., M.Cs</u> NIP. 19931107 202203 1 006	Anggota	 (.....)
4. <u>Agung Dwi Saputro, S.Kom., M.Kom</u> NIP. 19900323 201903 1 002	Anggota	 (.....)
5. <u>Supiyanto, S.Si., M.Kom</u> NIP. 19760906 200212 1 003	Anggota	 (.....)

Mengesahkan
Dekan Fakultas MIPA

Dr. Dirk Y. P. Runtuboi, M. Kes
NIP. 19760123 200112 1 003

Mengetahui
Ketua Jurusan Sistem Informasi


Supiyanto, S.Si., M.Kom
NIP. 19760906 200212 1 003

HALAMAN PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Kedua orang tua tercinta, Bapak Syamsu Alam dan Ibu Irianti.

Saudara-Saudaraku tersayang, Nadia Safitri, Ibnu Fairuz, dan Ahmad Kamil.

Teman-teman Sistem Informasi 2020.

PEDOMAN PENGGUNAAN SKRIPSI

Skripsi S1 FMIPA Uncen yang tidak dipublikasikan, terdaftar dan tersedia di Perpustakaan Universitas Cenderawasih dan terbuka untuk umum dengan ketentuan bahwa hak cipta ada pada penulis.

Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau peringkasan hanya dapat dilakukan seizin penulis, dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Memperbanyak atau menerbitkan sebagian atau seluruh skripsi haruslah seizin Rektor Universitas Cenderawasih.

Perpustakaan yang meminjam skripsi ini untuk keperluan anggotanya harus mengisi nama dan tanda tangan peminjam dan tanggal pinjam.

MOTTO

Finished what you started!

“Sesungguhnya bersama kesulitan ada kemudahan.”

(Q.S Al-Insyirah: 6)

“Keberanian dan ketekunanmu hari ini akan membawa hasil yang mengagumkan di masa depan. Teruslah berjuang, percayalah bahwa setiap langkah kecilmu hari ini adalah langkah menuju kesuksesan yang lebih besar.”

(Penulis)

PERNYATAAN PENULISAN SKRIPSI

Dengan ini menyatakan bahwa skripsi ini hasil karya sendiri yang belum pernah dipublikasikan pada perguruan tinggi atau lembaga manapun.

Jayapura, 20 Juni 2024

Dini Abshari

2020051074017

UCAPAN TERIMA KASIH

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Pengasih atas berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul “Implementasi Metode *Decision Tree* Untuk Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Algoritma C4.5”. Skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi Jurusan Sistem Informasi FMIPA Universitas Cenderawasih Jayapura.

Penulis dalam menyelesaikan skripsi ini mendapatkan banyak bantuan dari berbagai pihak baik berupa dukungan, pemikiran, motivasi, bimbingan, maupun informasi. Untuk itu, penulis mengucapkan terima kasih yang tulus kepada:

1. Bapak Dr. Dirk Y. P. Runtuboi, M. Kes, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Cenderawasih yang telah memberikan kesempatan kepada penulis untuk menjalani studi di Fakultas Matematika dan Ilmu Pengetahuan Alam.
2. Bapak Supiyanto, S.Si, M.Kom., selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Pengetahuan Alam Universitas Cenderawasih dan juga sebagai Dosen Pembimbing II yang telah membimbing penulis serta memberikan saran dan koreksi selama penulisan skripsi ini sehingga skripsi ini dapat terselesaikan dengan baik.
3. Bapak Agung Dwi Saputro, S.Kom, M.Kom., selaku Dosen Pembimbing I yang telah membimbing penulis, memberikan arahan, saran, dan nasehat selama penulisan skripsi sehingga skripsi ini dapat terselesaikan dengan baik.
4. Para Bapak/Ibu Dosen Program Studi Sistem Informasi yang telah memberikan segala ilmu dan pengetahuan.
5. Kedua orang tua tercinta, Bapak Syamsu Alam dan Ibu Irianti yang senantiasa mendoakan, mendukung, memberikan motivasi, dan kasih sayang kepada penulis sehingga penulis dapat menyelesaikan studi dengan baik.

6. Saudara-saudaraku tersayang, Nadia Safitri, Ahmad Kamil Maarif, dan Ibnu Fairuz Alqadri yang senantiasa menghibur, mendukung, dan memberi semangat kepada penulis.
7. Rahmatillah Haq, sebagai sahabat sekaligus teman seperjuangan selama kuliah, terima kasih telah berjuang bersama dari awal hingga akhir.
8. Sahabat-sahabat full senyum, Rachmaniah Maharany, Theonariya Kaban, Christine H Sinaga, Trifena Bulung, Arya Fira, dan Deliyan Padang. Terima kasih untuk kebersamaan yang telah dilalui bersama.
9. Teman-Teman Sistem Informasi Angkatan 2020, terima kasih untuk pengalaman yang berkesan selama kuliah.
10. Kak Sandika, sebagai kakak tingkat sekaligus staf jurusan yang telah memberikan banyak bantuan kepada penulis selama pengerjaan skripsi.
11. Terakhir, terima kasih kepada diri sendiri karena telah mampu berada sampai di titik ini. Terima kasih karena telah berjuang dan berusaha hingga sejauh ini. Terima kasih karena telah mampu melewati setiap tantangan dan rintangan. Terima kasih karena telah menyelesaikan skripsi ini dengan baik dan semaksimal mungkin. Terima kasih banyak diriku!

Harapan penulis, semoga semua pihak yang telah membantu penulis selama menempuh studi hingga penulisan skripsi diberikan kemudahan dalam setiap langkah dalam hidup.

Jayapura, 20 Juni 2024

Penulis

DAFTAR ISI

	Halaman
ABSTRAK	ii
ABSTRACT	iii
LEMBAR PERSETUJUAN	iv
LEMBAR PENGESAHAN	v
HALAMAN PERSEMBAHAN	vi
PEDOMAN PENGGUNAAN SKRIPSI	vii
MOTTO	viii
PERNYATAAN PENULISAN SKRIPSI	ix
UCAPAN TERIMA KASIH.....	x
DAFTAR ISI.....	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvii
DAFTAR LAMPIRAN.....	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI	5
2.1 Konsep Data	5
2.2 Konsep Data Mining	6
2.3 Klasifikasi.....	11
2.4 Pohon Keputusan (<i>Decision Tree</i>)	12
2.5 Algoritma C4.5	12
2.6 Knowledge Discovery in Database (KDD)	14
2.7 Confusion Matrix	16

2.8	Split Validation.....	17
2.9	RapidMiner.....	17
2.10	Web	17
2.11	UML	17
2.12	Flowchart.....	20
2.13	HTML.....	21
2.14	PHP.....	21
2.15	Codeigniter	21
2.16	Bootstrap	22
2.17	XAMPP	22
2.18	Metode Pengembangan Sistem	22
2.19	Metode Pengujian Sistem.....	24
BAB III METODOLOGI PENELITIAN		26
3.1	Lokasi Penelitian	26
3.2	Spesifikasi Kebutuhan	26
3.3	Jenis Penelitian dan Sumber Data	27
3.4	Teknik Pengumpulan Data	27
3.5	Variabel Penelitian	28
3.6	Metode Pengembangan Sistem	28
3.7	Teknik Pengujian Sistem.....	29
3.8	Tahapan Penelitian	30
BAB IV ANALISIS DAN PERANCANGAN		32
4.1	Analisis Sistem	32
4.2	Analisis Algoritma C4.5.....	35
4.3	Implementasi RapidMiner	45
4.4	Perancangan Sistem.....	49
4.5	Perancangan Form	53
BAB V IMPLEMENTASI DAN PENGUJIAN		58
5.1	Implementasi	58
5.2	Pengujian	65
BAB VI PENUTUP		76
6.1	Kesimpulan.....	76

6.2	Saran.....	76
DAFTAR PUSTAKA		77
LAMPIRAN.....		79

DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Model atau pengetahuan merupakan output data mining.....	6
Gambar 2. 2 Tugas data mining	7
Gambar 2. 3 Tahapan data mining	10
Gambar 2. 4 Tahapan proses KDD	14
Gambar 2. 5 Diagram UML	18
Gambar 2. 6 Model waterfall	22
Gambar 3. 1 Tahapan Penelitian	30
Gambar 4. 1 Proses pengujian data	45
Gambar 4. 2 Hasil decision tree	46
Gambar 4. 3 Rule decision tree	46
Gambar 4. 4 Nilai Akurasi	48
Gambar 4. 5 Use case diagram.....	49
Gambar 4. 6 Flowchart perancangan sistem	51
Gambar 4. 7 Flowchart Algoritma C4.5	52
Gambar 4. 8 Perancangan halaman utama	53
Gambar 4. 9 Perancangan halaman dataset prediksi kelompok	53
Gambar 4. 10 Perancangan halaman split data	54
Gambar 4. 11 Perancangan halaman hasil prediksi split data	54
Gambar 4. 12 Perancangan halaman import data.....	55
Gambar 4. 13 Perancangan halaman hasil prediksi import data	55
Gambar 4. 14 Perancangan halaman dataset prediksi tunggal	56
Gambar 4. 15 Perancangan halaman perhitungan.....	56
Gambar 4. 16 Perancangan halaman prediksi	57
Gambar 4. 17 Perancangan Modal About	57
Gambar 5. 1 Halaman Utama.....	58
Gambar 5. 2 Halaman Dataset Prediksi Kelompok	59
Gambar 5. 3 Halaman Split Data	59
Gambar 5. 4 Halaman Hasil Prediksi Split Data.....	60

Gambar 5. 5 Halaman Hasil Prediksi Split Data – Hasil Perhitungan Akurasi	60
Gambar 5. 6 Halaman Import Data	61
Gambar 5. 7 Halaman Hasil Prediksi Import Data	61
Gambar 5. 8 Halaman Hasil Prediksi Import Data – Hasil Perhitungan Akurasi ..	62
Gambar 5. 9 Halaman Dataset Prediksi Tunggal	62
Gambar 5. 10 Halaman Perhitungan – Entropy & Gain	63
Gambar 5. 11 Halaman Perhitungan – Pohon Keputusan.....	63
Gambar 5. 12 Halaman Prediksi – Form Input Data Mahasiswa.....	64
Gambar 5. 13 Halaman Prediksi – Hasil Prediksi.....	64
Gambar 5. 14 Halaman About	65
Gambar 5. 15 Pengujian dengan split data persentase data training 60%.....	66
Gambar 5. 16 Hasil uji akurasi dengan split data persentase data training 60% ...	66
Gambar 5. 17 Pengujian dengan split data persentase data training 70%.....	66
Gambar 5. 18 Hasil uji akurasi dengan split data persentase data training 70% ...	67
Gambar 5. 19 Pengujian dengan split data persentase data training 80%.....	67
Gambar 5. 20 Hasil uji akurasi dengan split data persentase data training 80% ...	68
Gambar 5. 21 Pengujian dengan import data persentase data training 60%	68
Gambar 5. 22 Hasil uji akurasi dengan import data persentase data training 60%	69
Gambar 5. 23 Pengujian dengan import data persentase data training 70%	69
Gambar 5. 24 Hasil uji akurasi dengan import data persentase data training 70%	70
Gambar 5. 25 Pengujian dengan import data persentase data training 80%	70
Gambar 5. 26 Hasil uji akurasi dengan import data persentase data training 80%	71
Gambar 5. 27 Grafik hasil pengujian akurasi dengan split data dan import data ..	72
Gambar 5. 28 Pengujian dengan import data menggunakan data testing baru	73
Gambar 5. 29 Hasil uji akurasi dengan data testing baru.....	74

DAFTAR TABEL

	Halaman
Tabel 2. 1 Contoh konsep dataset	15
Tabel 2. 2 Confusion Matrix	16
Tabel 2. 3 Diagram use case	19
Tabel 2. 4 Diagram flowchart	20
Tabel 3.1 Perangkat lunak yang digunakan	26
Tabel 3. 2 Perangkat keras yang digunakan.....	27
Tabel 4. 1 Data Mahasiswa	32
Tabel 4. 2 Data Mahasiswa Setelah Dilakukan Tahapan Seleksi	32
Tabel 4. 3 Sampel Data Training	33
Tabel 4. 4 Sampel Data Testing	33
Tabel 4. 5 Transformasi variabel target dan variabel prediktor	34
Tabel 4. 6 Hasil transformasi sampel data training.....	34
Tabel 4. 7 Hasil transformasi sampel data testing.....	34
Tabel 4. 8 Hasil perhitungan entropy	35
Tabel 4. 9 Perhitungan Node 1.....	38
Tabel 4. 10 Perhitungan Node 2 (IPS 2 – Rendah)	39
Tabel 4. 11 Perhitungan Node 3 (IPS 4 – Cukup)	39
Tabel 4. 12 Perhitungan Node 4 (IPS 4 – Rendah)	40
Tabel 4. 13 Perhitungan Node 5 (IPS 3 – Rendah)	40
Tabel 4. 14 Perhitungan Node 6 (IPS 1 – Cukup)	41
Tabel 4. 15 Perhitungan Node 7 (IPS 1 – Rendah)	41
Tabel 4. 16 Perhitungan Node 8 (IPS 1 – Tinggi)	42
Tabel 4. 17 Perhitungan Node 9 (IPS 4 – Tinggi)	42
Tabel 4. 18 Perhitungan Node 10 (IPS 2 – Tinggi)	43
Tabel 4. 19 Perhitungan Node 11 (IPS 3 – Rendah)	43
Tabel 4. 20 Perhitungan Node 12 (IPS 1 – Tinggi)	44
Tabel 4. 21 Perhitungan Node 13 (JK – L)	44

Tabel 4. 22 Hasil Prediksi Data Testing	48
Tabel 5. 1 Hasil pengujian akurasi dengan split data dan import data.....	71
Tabel 5. 2 Dataset mahasiswa yang digunakan sebagai data testing baru	72
Tabel 5. 3 Data testing baru yang telah ditransformasi.....	73
Tabel 5. 4 Pengujian Black Box.....	74

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Training Mahasiswa Hasil Transformasi	77
Lampiran 2. Data Testing Mahasiswa Hasil Transformasi	82