

TUGAS AKHIR

**IMPLEMENTASI ALGORITMA DECISION TREE UNTUK
KLASIFIKASI POLA SERANGAN PADA LOG FILE**



Miftarullah Firdaus

1410651012

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2019**

**IMPLEMENTASI ALGORITMA DECISION TREE UNTUK
KLASIFIKASI POLA SERANGAN PADA LOG FILE**

Tugas Akhir

**Ditulis dan diajukan untuk memenuhi persyaratan mendapatkan gelar
Sarjana Komputer pada Program Studi Teknik Informatika Fakultas
Teknik Universitas Muhammadiyah Jember**



Oleh:

MIFTARULLAH FIRDAUS

1410651012

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2019

HALAMAN PENGESAHAN

IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI POLA SERANGAN PADA LOG FILE

MIFTARULLAH FIRDAUS

1410651012

Telah mempertanggung jawabkan Tugas Akhirnya pada sidang Tugas Akhir tanggal 12 Maret 2019 sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui oleh:

Dosen Pembimbing

Triawan Adi Cahyanto, M.Kom

NPK. 12 03 719

Dosen Penguji I

Dosen Penguji II

Agung Nilogiri, S.T.,M.Kom

NIP. 19770330 200501 1 002

Ir. Dewi Lusiana, M.T

NPK. 07 12 086 702

Mengesahkan,

Dekan Fakultas Teknik

Mengetahui,

Ketua Program Studi Teknik Informatika

Ir. Suhartinah, M.T

NPK. 95 05 246

Yeni Dwi Rahayu, S. ST., M.Kom

NPK. 11 03 590

PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Miftarullah Firdaus

Nim : 1410651012

Program Studi : Teknik Informatika

Perguruan Tinggi: Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul “**Implementasi Algoritma Decision Tree Untuk Klasifikasi Pola Serangan Pada Log File**” adalah ASLI dan BELUM PERNAH dibuat orang lain, kecuali yang diacu dalam Daftar Pustaka pada Tugas Akhir ini.

Demikian Surat Pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak sesuai, penulis bersedia mendapatkan sanksi dari akademik.

Jember, 12 Maret 2019

Miftarullah Firdaus

NIM. 1410651012

HALAMAN PERSEMBAHAN

Atas kehadiran Allah SWT yang telah memberikan rahmat dan hidayahnya sehingga tugas akhir ini berhasil diselesaikan. Dalam penelitian ini penulis mempersembahkan karya ini untuk orang-orang yang sangat membantu dalam menyelesaikan penelitian ini antara lain:

1. Kepada Orang tua saya Sasyono dan Siti Lestari yang telah mendoakan dan menjadi penyemangat saya sehingga selalu istiqomah dalam mengerjakan skripsi. Dan juga kakak Nur Lailatul Hikmah dan Adik Riska Yunita Rahmawati yang tiada henti mengingatkan untuk tidak selalu menghabiskan waktu untuk bermain.
2. Kepada sahabat-sahabat saya yang ada di kosan putra Halmahera IV No 05 yang selama ini tidak pernah lelah memberikan semangat.
3. Kepada saudara-saudara terimakasih atas bantuan sarana, dukungan, nasihat yang di berikan selama ini, semoga Allah membalasnya dan memberikan barokah dan hidayahNya.
4. Teman-teman seperjuangan mahasiswa Program Studi Teknik Informatika angkatan 2014.
5. Universitas Muhammadiyah Jember dan Program Studi Teknik Informatika sehingga saya mendapatkan mendapatkan gelar Sarjana Komputer.

Saya menyadari bahwa tugas skripsi ini masih jauh dari kesempurnaan. Oleh karena itu saya menerima saran dan kritik yang membangun dari para pembaca demi sempurnanya skripsi ini. Melalui skripsi ini, saya berharap semoga skripsi yang sederhana ini memberi manfaat kepada kita semua.

MOTTO

“Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid”

(Albert Einstein)

Kerjakanlah hal yang bermanfaat untuk diri sendiri dan orang lain, karena hidup hanyalah sekali. Ingat hanya pada Allah, Kepada Dia-lah tempat meminta dan memohon.



UCAPAN TERIMAKASIH

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillah segala puji bagi Allah yang senantiasa memberikan nikmat yang sangat luar biasa dalam hidup ini kepada saya yang berupa nikmat untuk selalu dapat bersujud dan meminta ampunan kepada-Nya dengan iman dan islam. Dan berkat rahmad-Nya penulis diberikan kemudahan dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember.

Atas segala upaya, bimbingan dan arahan dari semua pihak, penulis mengucapkan terimakasih dan saya juga berterimakasih kepada pembimbing yang selama ini telah membantu saya dalam menyelesaikan tugas akhir ini.

Dalam kesempatan ini penulis mengucapkan terimakasih sebanyak-banyaknya kepada:

1. Allah SWT atas segala nikmat, rahmat dan hidayah-Nya. Terimakasih Allah atas kehendak-Mu telah memberikan kesempatan ini untuk mendapatkan gelar Sarjana Komputer.
2. Ibu Suhartinah, M.T. Selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Ibu Yeni Dwi Rahayu, ST., M.Kom. Selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Jember.
4. Bapak Triawan Adi Cahyanto, M.Kom selaku pembimbing yang telah meluangkan waktu untuk memberi arahan dan membimbing saya dalam menyelesaikan tugas akhir ini.
5. Bapak Agung Nilogiri S.T., M.Kom selaku penguji I dan Ibu Ir. Dewi Lusiana, MT selaku penguji II yang telah memberikan kritik dan saran dalam penyelesaian tugas akhir ini.
6. Kepada Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmunya kepada saya.

7. Ayahanda Sasyono dan Ibunda Siti Lestari yang telah mendoakan dan menjadi penyemangat saya sehingga selalu istiqomah dalam mengerjakan skripsi. Dan juga kakak Nur Lailatul Hikmah dan Adik Riska Yunita Rahmawati yang tiada henti mengingatkan untuk tidak selalu menghabiskan waktu untuk bermain.
8. Saudara saya Muhammad Mas'ud yang telah membantu untuk memberikan semangat, dukungan, bantuan, saran, dan doa terhadap penulis untuk menyelesaikan tugas akhir ini.
9. Kepada teman-teman Teknik Informatika Angkatan 2014 yang telah memberikan semangat dan doanya.

Penulis menyadari bahwa dalam penulisan skripsi ini masih jauh dari kesempurnaan. Karena itu, penulis memohon saran dan kritik yang bersifat membangun demi kesempurnanya dan semoga skripsi ini dapat berguna dan bermanfaat bagi kita semua. Aamiin.

Jember, 12 Maret 2019

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN KEASLIAN	iii
HALAMAN PERSEMBAHAN	iv
MOTTO	v
ABSTRAK	vi
ABSTRACT	vii
UCAPAN TERIMAKASIH	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 Batasan Masalah	3
1.5 Manfaat	3
BAB II LANDASAN TEORI	4
2.1 Literature Review	4
2.2 Sistem Deteksi Intrusi	6
2.3 Dataset	7
2.4 Data Mining	7
1. Data Selection	8
2. Pre-processing/Cleaning	8
3. Transformation	8
4. Data Mining	9
5. Interpretation Evaluation	9
2.4.1 Pengelompokan Data Mining	9
1. Deskripsi	9
2. Klasifikasi	9
3. Esimasi	9
4. Prediksi	10

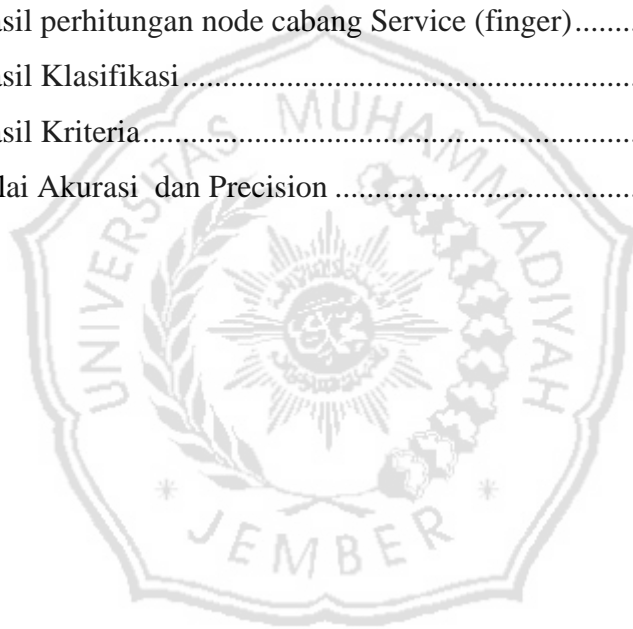
5. Klustering	10
6. Asosiasi	10
2.5 Klasifikasi	10
2.6 Algoritma Decision Tree	11
2.6.1 Pemilihan Atribut dan Pembentukan Tree	14
2.6.2 Algoritma-algoritma dalam Decision Tree	14
2.6.3 Algoritma C.45	14
2.6.4 Karakteristik Decision Tree	15
2.7 Weka	16
BAB III METODE PENELITIAN.....	18
3.1 Analisa Kebutuhan	18
3.1.1 Pengumpulan Data	18
3.1.2 Algoritma yang digunakan	18
3.1.3 Program yang digunakan	18
3.2 Desain Sistem	20
3.3 Perhitungan Algoritma C 4.5	21
3.3.1 Tahap Perhitungan	21
3.3.2 Perhitungan Node 1 (Akar/Root).....	23
3.3.3 Perhitungan Nilai Gain Atribut.....	26
3.3.4 Rule.....	29
3.3.5 Perhitungan Klasifikasi dan Kriteria	30
BAB IV HASIL DAN PEMBAHASAN	33
4.1 Analisis Data	33
4.2 Lingkungan Hasil	33
4.3 Perangkat Keras	33
4.4 Perangkat Lunak	34
4.5 Pengujian Algoritma C 4.5 Dengan Software WEKA	34
4.6 Evaluasi Hasil	41
BAB V KESIMPULAN DAN SARAN	42
5.1 Kesimpulan	42
5.2 Saran	42

DAFTAR PUSTAKA	44
LAMPIRAN 1.....	45
LAMPIRAN 2.....	57
LAMPIRAN 3.....	68
DAFTAR REVISI PENGUJI 1	
DAFTAR REVISI PENGUJI 2	
BIODATA PENULIS	



DAFTAR TABEL

Tabel 3.1 Fitur Utama dari Koneksi TCP individual	18
Tabel 3.2 Data Sampel	21
Tabel 3.3 Data training.....	22
Tabel 3.4 Hasil perhitungan Entropy dan Gain.....	25
Tabel 3.5 Hasil perhitungan node cabang Service (finger).....	28
Tabel 3.6 Hasil Klasifikasi.....	30
Tabel 3.7 Hasil Kriteria.....	31
Tabel 3.8 Nilai Akurasi dan Precision	31



DAFTAR GAMBAR

Gambar 2.1 <i>Data mining</i> sebagai tahapan dalam proses KDD	8
Gambar 2.2 Contoh Decision Tree.....	11
Gambar 3.1 Flowchart Penelitian dengan Menggunakan Weka	20
Gambar 3.2 Pohon Keputusan	29
Gambar 3.3 Rumus Prediksi	31
Gambar 4.1 Antar Muka Awal Weka Explore	35
Gambar 4.2 Antar Muka Weka Explore	35
Gambar 4.3 Antar Muka Praproses	36
Gambar 4.4 Memilih File	36
Gambar 4.5 Antar Muka Pra-proses Klasifikasi	37
Gambar 4.6 Memilih Algoritma C 4.5	37
Gambar 4.7 Memilih atribut target yang akan diproses	38
Gambar 4.8 Hasil pengujian dengan software Weka	39
Gambar 4.9 Hasil pohon keputusan menggunakan Weka	40
Gambar 4.10 Hasil klasifikasi menggunakan software	41

Daftar Pustaka

1. Basuki, Achmad dan Syarif, Iwan. 2003. “Modul Ajar Decision Tree”, Surabaya : PENS-ITS.
2. Budiman,Ade Surya.,Anty Adhi Parandani.2018.” Uji Akurasi Klasifikasi Dan Validasi Data Pada Penggunaan Metode Membership Function Dan Algoritma C4.5 Dalam Penilaian Penerima Beasiswa”. Jakarta : AMIK BSI.
3. Fayyad, U. (1996). Advances in Knowledge Discovery and Data Mining. MIT Press.
4. Hermanto,Bambang,Azhari SN, Fajri Profesio Putra. 2017” Analisis Perbandingan Algoritma ID3 Dan C4.5 Untuk Klasifikasi Penerima Hibah Pemasangan Air Minum Pada PDAM Kabupaten Kendal”.Lampung : Universitas Lampung.
5. J. E. Gewehr, M. Szugat, and R. Zimmer.2007.“BioWeka—Extending theWeka
6. Khaerani,Izza, dan Lekso Budi Handoko. 2018. ” Implementasi Dan Analisa Hasil Data Mining Untuk Klasifikasi Serangan Pada Intrusion Detection System (Ids) Dengan Algoritma C4.5”.Semarang: Universitas Dian Nuswantoro.
7. Larose, D. T. (2005). Discovering Knowledge in Data: An Introduction to Data mining. John Willey and Sons, Inc.
8. RHRDNWK.2016. Contoh Perhitungan Decision Tree dengan Algoritma C45. [online], (<https://www.ilmuskripsi.com/2016/05/contoh-perhitungan-decision-tree-dengan.html>).

9. Santosa, Iwan., Hammimatur Rosiyah, and Eza Rahmanita. 2018. Implementasi Algoritma Decision Tree C.45 Untuk Diagnosa Penyakit Tuberculosis (Tb). Madura : Universitas Trunojoyo.
10. Silberschatz, A. and Tuzhilin, A. 1995. On Subjective Measures of Interestingness in Knowledge Discovery. In *Proceedings of KDD-95: First International Conference on Knowledge Discovery and Data Mining*. USA : New York University.
11. S.J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. Chan. 1999. Cost-based Modeling for Fraud and Intrusion Detection: Results from the JAM Project. New York : Columbia University.
12. Tavallaee, Mahbod., Ebrahim Bagheri, Wei Lu, and Ali A. Ghorbani. 2009. A Detailed Analysis Of The Kdd Cup 99 Data Set. Proceeding Of the IEEE Symposium on Computational Intelligence.
13. Witten, I. dan Eibe Frank. 2006. *Data mining: Practical machine learning tools and Techniques*. USA : San Francisco.
14. Witten, I. dan Eibe Frank. 2005. *Data mining: Practical machine learning tools and Techniques*. USA : San Francisco.
15. Witten, I., Eibe Frank, Len Trigg, Mark Hall, Geoffrey Holmes, and Sally Jo Cunningham. 1999. *Weka: Practical machine learning tools and techniques with java implementations*.
16. Witten, I., Eibe Frank, Mark A. Hall. 2016. *Data Mining : Practical Machine Learning Tools and Techniques*”. Morgan Kaufmann: Fourth Edition.